

Quadro 2 - Fichamento dos artigos do Mapeamento sistemático.

Nº	TÍTULO	CITAÇÃO	OBJETIVO	CITAÇÃO INDIRETA	RESUMO
01	Overview of Early <i>ChaGPT</i> 's Presence in Medical Literature: Insights From a Hybrid Literature Review by <i>ChaGPT</i> and Human Experts Visão Geral da Presença Inicial do <i>ChaGPT</i> na Literatura Médica: Percepções a Partir de uma Revisão Bibliográfica Híbrida Realizada por <i>ChaGPT</i> e Especialistas Humanos	Temsah et al. (2023)	Resumir e sintetizar o conhecimento atual do <i>ChaGPT</i> na literatura médica indexada durante seus quatro meses iniciais.	Temsah et al. (2023), conduziram uma revisão narrativa híbrida, realizada colaborativamente por autores humanos e <i>ChaGPT</i> , com o objetivo de resumir e sintetizar o conhecimento atual do <i>ChaGPT</i> na literatura médica indexada durante seus quatro meses iniciais. A busca nas bases de dados PubMed e EuropePMC resultou em 65 e 110 artigos, respectivamente, que abordavam o impacto do <i>ChaGPT</i> na educação médica, pesquisa científica, redação médica, considerações éticas, tomada de decisões diagnósticas, potencial de automação e críticas. As descobertas apontam para um crescente corpo de literatura sobre as aplicações e implicações do <i>ChaGPT</i> na saúde, ressaltando a necessidade de mais pesquisas para avaliar sua eficácia e abordar preocupações éticas emergentes.	ChaGPT, um <i>Chatbot</i> de inteligência artificial, rapidamente ganhou destaque em diversos domínios, incluindo a educação médica e a literatura em saúde. Esta revisão narrativa híbrida, conduzida de forma colaborativa por autores humanos e <i>ChaGPT</i> , tem como objetivo resumir e sintetizar o conhecimento atual sobre o <i>ChaGPT</i> na literatura médica indexada durante seus primeiros quatro meses. Uma estratégia de busca foi empregada nas bases de dados PubMed e EuropePMC, resultando em 65 e 110 artigos, respectivamente. Esses artigos focaram no impacto do <i>ChaGPT</i> na educação médica, pesquisa científica, escrita médica, considerações éticas, tomada de decisão diagnóstica, potencial de automação e críticas. Os resultados indicam um crescente corpo de literatura sobre as aplicações e implicações do <i>ChaGPT</i> na área da saúde, destacando a necessidade de mais pesquisas para avaliar sua eficácia e preocupações éticas.
02	Opportunities, Challenges, and Future Directions of Generative Artificial Intelligence in Medical Education: Scoping Review Oportunidades, Desafios e Direções Futuras da Inteligência Artificial Generativa na Educação Médica: Revisão Abrangente	Preiksaitis e Rose (2023)	Sintetizar as potenciais oportunidades e limitações da IA generativa na educação médica.	Preiksaitis e Rose (2023), conduziram um estudo com o objetivo de sintetizar as potenciais oportunidades e limitações da Inteligência Artificial generativa na educação médica. A análise temática revelou várias aplicações promissoras, como aprendizagem autodirigida, cenários de simulação e assistência por escrito. Contudo, a literatura também destacou desafios significativos, incluindo problemas de integridade acadêmica, precisão dos dados e possíveis impactos negativos na aprendizagem. Com base nesses temas e no estado atual da literatura, os autores propõem três áreas principais de investigação: desenvolvimento	Contexto: As tecnologias de inteligência artificial (IA) generativa estão sendo cada vez mais utilizadas em diversos campos, gerando considerável interesse e preocupação em relação à sua aplicação potencial na educação médica. Essas tecnologias, como Chat GPT e Bard, podem gerar novos conteúdos e possuem uma ampla gama de possíveis aplicações. Objetivo: Este estudo teve como objetivo sintetizar as oportunidades e limitações potenciais da IA generativa na educação médica. Procurou identificar temas predominantes na literatura recente sobre as aplicações e desafios potenciais da IA generativa na educação médica e usar esses temas para orientar futuras áreas de exploração. Métodos: Realizamos uma revisão de escopo, seguindo o framework de Arksey e O'Malley, de artigos em inglês publicados a partir de 2022 que discutiam IA generativa no contexto da educação médica. Uma busca na literatura foi realizada utilizando as bases de dados PubMed, Web of Science e Google Scholar. Selecionamos artigos para inclusão, extraímos dados de estudos relevantes e realizamos uma síntese quantitativa e qualitativa dos dados. Resultados: A análise temática revelou diversas aplicações potenciais para a IA generativa na educação médica, incluindo aprendizagem autodirigida, cenários de simulação e assistência na escrita. No

				<p>da capacidade dos <i>discentes</i>, habilidades para avaliar criticamente a IA, e a necessidade de repensar a metodologia de avaliação, além de estudar as interações humano-IA. Essas propostas visam orientar futuras pesquisas e explorações no campo da educação médica com IA generativa.</p>	<p>entanto, a literatura também destacou desafios significativos, como questões de integridade acadêmica, precisão dos dados e possíveis prejuízos ao aprendizado. Com base nesses temas e no estado atual da literatura, propomos as seguintes 3 áreas-chave para investigação: desenvolvimento de habilidades dos <i>discentes</i> para avaliar a IA criticamente, repensar a metodologia de avaliação e estudar as interações humano-IA. Conclusões: A integração da IA generativa na educação médica apresenta oportunidades empolgantes, juntamente com desafios consideráveis. Há a necessidade de desenvolver novas habilidades e competências relacionadas à IA, bem como abordagens ponderadas e nuançadas para examinar o uso crescente da IA generativa na educação médica.</p>
03	<p>Medical Student Experiences and Perceptions of ChaGPT and Artificial Intelligence: Cross-Sectional Study</p> <p>Experiências e Percepções de Discentes de Medicina em Relação ao ChaGPT e à Inteligência Artificial: Estudo Transversal</p>	<p>Alkhaaldi et al. (2023)</p>	<p>Avaliar as experiências e perspectivas dos discentes de medicina sobre ChaGPT e IA em sua formação e carreiras futuras.</p>	<p>Alkhaaldi et al. (2023), conduziram um estudo para avaliar as experiências e perspectivas dos discentes de medicina em relação ao ChaGPT e à Inteligência Artificial (IA). Dos 325 candidatos aos programas de residência, 265 participaram da pesquisa, representando uma taxa de resposta de 81,5%. A maioria dos entrevistados não utilizou o ChaGPT durante a faculdade de medicina, com 20,4% usando-o para avaliações escritas e 9,4% na prática clínica. Para a residência, a maioria planejava usar a tecnologia para explorar novos tópicos e pesquisas médicas (63,4%) e preparação para exames (57%). Houve uma associação significativa entre o sexo masculino e a crença na melhoria da precisão do diagnóstico, redução de erros médicos e aprimoramento do atendimento ao paciente pela IA. A experiência anterior com IA correlacionou-se positivamente com a percepção favorável em relação à IA em termos de atendimento ao paciente, redução de erros médicos e diagnósticos equivocados, e aumento da precisão diagnóstica. Esses resultados oferecem insights sobre as atitudes e expectativas dos discentes de medicina em relação à incorporação da IA em suas práticas futuras.</p>	<p>Contexto: A inteligência artificial (IA) tem o potencial de revolucionar a forma como a medicina é aprendida, ensinada e praticada, e a educação médica deve preparar os aprendizes para essas mudanças inevitáveis. No entanto, a medicina acadêmica tem sido lenta em adotar os avanços recentes da IA. Desde seu lançamento em novembro de 2022, o ChaGPT emergiu como um modelo de linguagem grande, rápido e fácil de usar, que pode auxiliar profissionais de saúde, educadores médicos, discentes, residentes e pacientes. Enquanto muitos estudos focam nas capacidades, potencialidades e riscos da tecnologia, há uma lacuna na análise da perspectiva dos usuários finais. Objetivo: O objetivo deste estudo foi avaliar as experiências e perspectivas dos discentes de medicina recém-formados sobre o ChaGPT e a IA em sua formação e futuras carreiras. Métodos: Foi realizada uma pesquisa transversal online com discentes de medicina recém-formados em um centro médico acadêmico internacional entre 5 de maio de 2023 e 13 de junho de 2023. Estatísticas descritivas foram usadas para tabular as frequências das variáveis. Resultados: Dos 325 candidatos aos programas de residência, 265 completaram a pesquisa (uma taxa de resposta de 81,5%). A grande maioria dos respondentes negou ter usado o ChaGPT na faculdade de medicina, com 20,4% (n=54) usando-o para ajudar na conclusão de avaliações escritas e apenas 9,4% usando a tecnologia em seu trabalho clínico (n=25). Mais discentes planejavam usá-lo durante a residência, principalmente para explorar novos tópicos médicos e pesquisa (n=168, 63,4%) e preparação para exames (n=151, 57%). Discentes do sexo masculino eram significativamente mais propensos a acreditar que a IA melhorará a precisão diagnóstica (n=47, 51,7% vs n=69, 39,7%; P=.001), reduzirá o erro médico (n=53, 58,2% vs n=71, 40,8%; P=.002) e melhorará o cuidado ao paciente (n=60, 65,9% vs n=95, 54,6%; P=.007). Experiência prévia com IA estava significativamente associada a uma percepção positiva da IA em termos de melhoria do cuidado ao paciente, diminuição de erros médicos e diagnósticos equivocados, e aumento da precisão dos diagnósticos (P=.001, P<.001, P=.008, respectivamente). Conclusões: Os discentes de medicina pesquisados tinham pouca experiência formal e informal com ferramentas de IA e percepções limitadas sobre os usos potenciais da IA na</p>

					saúde, mas tinham visões geralmente positivas sobre o ChatGPT e a IA e eram otimistas quanto ao futuro da IA na educação médica e na saúde. Currículos estruturados e políticas e diretrizes formais são necessários para preparar adequadamente os discentes de medicina para a integração iminente da IA na medicina.
04	AI-Powered Chatbots in Medical Education: Potential Applications and Implications <i>Chatbots</i> Impulsionados por Inteligência Artificial na Educação Médica: Aplicações Potenciais e Implicações	Ghorashi et al. (2023)	Explorar as aplicações e implicações potenciais dos Chatbots na educação médica, com um foco particular em como eles podem melhorar a aprendizagem e a pesquisa.	Ghorashi et al. (2023) discutiram as potenciais aplicações dos Chatbots na educação médica, enfatizando sua capacidade de simplificar conceitos complexos, automatizar a criação de auxiliares de memória e atuar como tutores interativos. Apesar das vantagens da integração de Chatbots alimentados por IA, os autores salientaram a importância de os <i>discentes</i> usarem essas ferramentas como suporte, não dependendo exclusivamente delas. Destacaram a necessidade de programar os Chatbots para referenciar recursos médicos baseados em evidências, produzindo conteúdo preciso e confiável em conformidade com padrões científicos e considerações éticas, visando otimizar os benefícios educacionais.	Resumo: Espera-se que a inteligência artificial (IA) tenha um impacto considerável na prática rotineira da medicina, abrangendo desde a educação médica até a prática clínica em diversas especialidades e, em última instância, o cuidado ao paciente. Com a iminente adoção generalizada da IA na prática médica, é imperativo que as escolas de medicina se adaptem ao uso dessas tecnologias avançadas em seus currículos para formar futuros profissionais de saúde que possam integrar essas ferramentas de forma harmoniosa em sua prática. Chatbots, sistemas de IA programados para processar e gerar linguagem humana, estão atualmente sendo avaliados para diversas tarefas na educação médica. Este artigo explora as potenciais aplicações e implicações dos Chatbots na educação médica, especificamente em aprendizado e pesquisa. Com sua capacidade de resumir, simplificar conceitos complexos, automatizar a criação de auxiliares de memória e atuar como tutor interativo e referência médica no ponto de atendimento, os Chatbots têm o potencial de aprimorar a compreensão, retenção e aplicação do conhecimento médico dos discentes em tempo real. Embora a integração de Chatbots alimentados por IA na educação médica apresente inúmeras vantagens, é crucial que os discentes utilizem essas ferramentas como auxiliares, em vez de depender inteiramente delas. Os Chatbots devem ser programados para referenciar recursos médicos baseados em evidências e produzir conteúdo preciso e confiável que esteja em conformidade com os padrões da ciência médica, diretrizes de redação científica e considerações éticas.
05	Using ChatGPT as a Learning Tool in Acupuncture Education: Comparative Study Utilizando o ChatGPT como Ferramenta de Aprendizado na Educação em Acupuntura: Estudo Comparativo	Lee (2023)	Avaliar o potencial do ChatGPT como ferramenta educacional em programas universitários de acupuntura, com foco em sua capacidade de apoiar os <i>discentes</i> no aprendizado da seleção de pontos de acupuntura, planejamento de tratamento e tomada de decisões.	Lee (2023), realizou um estudo para avaliar o potencial do ChatGPT como ferramenta educacional em programas universitários de acupuntura, concentrando-se na capacidade de apoiar os <i>discentes</i> na aprendizagem da seleção de pontos de acupuntura, planejamento de tratamento e tomada de decisões. Analisando 14 dos 21 estudos de caso, observou-se que o ChatGPT-4 gerou significativamente mais pontos de acupuntura (média de 9,0, DP 1,1) em comparação com o ChatGPT-3,5 (média de 5,6, DP 0,6; P<0,001). As proporções de sobreposição de pontos de acupuntura para o ChatGPT-3,5 (média de 0,40, DP 0,28) e ChatGPT-4 (média de 0,34, DP 0,27; P = 0,67) não apresentaram diferença significativa. Esses resultados sugerem a	Resumo: Contexto: O ChatGPT (OpenAI) é um modelo de inteligência artificial de última geração com potenciais aplicações nos campos médicos da prática clínica, pesquisa e educação. Objetivo: Este estudo teve como objetivo avaliar o potencial do ChatGPT como ferramenta educacional em programas universitários de acupuntura, focando em sua capacidade de apoiar os <i>discentes</i> na seleção de pontos de acupuntura, planejamento de tratamentos e tomada de decisões. Métodos: Coletamos estudos de caso publicados na revista Acupuncture in Medicine entre junho de 2022 e maio de 2023. Tanto o ChatGPT-3,5 quanto o ChatGPT-4 foram utilizados para gerar sugestões de pontos de acupuntura com base nas apresentações dos casos. Um teste de postos sinalizados de Wilcoxon foi realizado para comparar o número de pontos de acupuntura gerados pelo ChatGPT-3,5 e pelo ChatGPT-4, e a razão de sobreposição dos pontos de acupuntura foi calculada. Resultados: Entre os 21 estudos de caso, 14 foram incluídos para análise. O ChatGPT-4 gerou significativamente mais pontos de acupuntura (9,0, DP 1,1) em comparação com o ChatGPT-3,5 (5,6, DP 0,6; P<.001). As razões de sobreposição dos pontos de acupuntura para o ChatGPT-3,5 (0,40, DP 0,28) e o ChatGPT-4

				eficácia potencial do <i>ChaGPT</i> como uma ferramenta de apoio educacional na prática da acupuntura.	(0,34, DP 0,27; P=.67) não foram significativamente diferentes. Conclusões: O <i>ChatGPT</i> pode ser uma ferramenta educacional útil para discentes de acupuntura, proporcionando insights valiosos sobre planos de tratamento personalizados. No entanto, ele não pode substituir completamente os métodos de diagnóstico tradicionais, e mais estudos são necessários para garantir sua implementação segura e eficaz na educação em acupuntura.
06	Reshaping medical education: Performance of <i>ChaGPT</i> on a PES medical examination Reformulando a Educação Médica: Desempenho do <i>ChaGPT</i> em um Exame Médico do PES (Progress Test of Medical Education)	<u>Wójcik, Rulkiewicz e Pruszczyk (2023)</u>	Verificar o desempenho do <i>ChaGPT</i> nos exames da faculdade de medicina e destacar como ele pode mudar a educação e a prática médica.	Wójcik, Rulkiewicz e Pruszczyk (2023) conduziram um estudo para verificar o desempenho do <i>ChaGPT</i> nos exames da faculdade de medicina, utilizando o modelo mais atualizado na época da publicação (GPT-4). O <i>ChaGPT</i> foi testado em um exame de licenciamento médico na Polônia (PES) em junho de 2023, respondendo a um conjunto de 120 perguntas. Os resultados revelaram avanços notáveis no processamento de linguagem natural, com o <i>ChaGPT</i> atingindo uma taxa de acerto de 67,1%, respondendo com precisão a 80 das 120 perguntas. Esses resultados destacam o potencial do <i>ChaGPT</i> em tarefas de resposta a perguntas médicas e sugerem implicações significativas para a educação e prática médica.	Resumo Contexto: Atualmente, estamos vivenciando uma terceira revolução digital impulsionada pela inteligência artificial (IA), e o surgimento do novo transformador gerativo pré-treinado de chat (<i>ChaGPT</i>) representa um avanço tecnológico significativo com profundas implicações para a sociedade global, especialmente no campo da educação. Métodos: O objetivo deste estudo foi avaliar o desempenho do <i>ChaGPT</i> em exames de escolas médicas e destacar como ele pode mudar a educação e a prática médica. Recentemente, o <i>ChaGPT</i> da OpenAI (OpenAI, San Francisco; versão GPT-4 de 24 de maio) foi testado contra um importante exame de licenciamento de especialização médica polonesa (PES), e os resultados estão disponíveis. A versão do <i>ChaGPT</i> -4 utilizada neste estudo era o modelo mais atualizado na época da publicação (GPT-4). O <i>ChaGPT</i> respondeu às perguntas de 28 de junho de 2023 a 30 de junho de 2023. Resultados: O <i>ChaGPT</i> demonstra avanços notáveis em modelos de processamento de linguagem natural nas tarefas de responder a perguntas médicas. Em junho de 2023, o desempenho do <i>ChaGPT</i> foi avaliado com base em sua capacidade de responder a um conjunto de 120 perguntas, onde alcançou uma taxa de acerto de 67,1%, respondendo corretamente a 80 perguntas. Conclusões: O <i>ChaGPT</i> pode ser usado como uma ferramenta auxiliar na educação médica. Embora o <i>ChaGPT</i> possa servir como uma ferramenta valiosa na educação médica, ele não pode substituir completamente a expertise e o conhecimento humanos devido às suas limitações inerentes. (Cardiol J). Palavras-chave: <i>ChaGPT</i> , inovações, inteligência artificial, IA na medicina, TI em saúde, educação médica, processamento de linguagem, assistente virtual de ensino
07	An exploratory survey about using <i>ChaGPT</i> in education, healthcare, and research Uma Pesquisa Exploratória sobre o Uso do <i>ChaGPT</i> na Educação, Saúde e Pesquisa	Hosseini et al. (2023)	Explorar como o <i>ChatGPT</i> está sendo integrado, destacando as incertezas e os potenciais benefícios e desafios dessa tecnologia.	Hosseini et al. (2023) destacaram que o <i>ChaGPT</i> , sendo o primeiro grande modelo de linguagem a alcançar uma ampla adoção pelo público em geral, tem gerado discussões significativas sobre sua integração aceitável e ótima em áreas como educação, pesquisa e saúde. Essas perspectivas foram abordadas em um painel de discussão híbrido, envolvendo participantes tanto virtualmente quanto presencialmente.	Objetivo: O <i>ChaGPT</i> é o primeiro modelo de linguagem grande (LLM) a alcançar um grande público mainstream. Sua rápida adoção e exploração pela população em geral geraram uma ampla gama de discussões sobre sua integração aceitável e ideal em diferentes áreas. Em um evento de painel de discussão híbrido (virtual e presencial), examinamos várias perspectivas sobre o uso do <i>ChaGPT</i> na educação, pesquisa e saúde. Materiais e Métodos: Pesquisamos os participantes presenciais e online usando uma plataforma de interação com o público (Slido). Analisamos quantitativamente as respostas recebidas sobre perguntas relacionadas ao uso do <i>ChaGPT</i> em vários contextos. Comparamos grupos categóricos par a par com o teste exato de Fisher. Além disso, utilizamos métodos qualitativos para analisar e codificar as discussões. Resultados: Recebemos 420 respostas de um estimado de 844 participantes (taxa de resposta de 49,7%). Apenas 40% do público havia experimentado o <i>ChaGPT</i> . Mais trainees haviam experimentado o <i>ChaGPT</i> em comparação com os

					professores. Aqueles que usaram o ChaGPT estavam mais interessados em utilizá-lo em uma gama mais ampla de contextos no futuro. Dos três contextos discutidos, a maior incerteza foi mostrada em relação ao uso do ChaGPT na educação. Prós e contras foram levantados durante a discussão sobre o uso dessa tecnologia na educação, pesquisa e saúde. Discussão: Houve uma variedade de perspectivas sobre os usos do ChaGPT na educação, pesquisa e saúde, com ainda muita incerteza sobre sua aceitabilidade e usos ideais. Houve diferentes perspectivas dos respondentes de diferentes funções (trainee vs professor vs staff). Mais discussão é necessária para explorar as percepções sobre o uso de <i>LLMs</i> como o ChaGPT em setores vitais como educação, saúde e pesquisa. Dado os riscos envolvidos e desafios imprevistos, adotar uma abordagem ponderada e medida reduziria a probabilidade de danos.
08	Assessment of the capacity of <i>ChaGPT</i> as a self-learning tool in medical pharmacology: a study using MCQs Avaliação da Capacidade do <i>ChaGPT</i> como Ferramenta de Autoaprendizagem em Farmacologia Médica: Um Estudo Utilizando Questões de Múltipla Escolha (MCQs)	Choi, (2023)	Avaliar o potencial do <i>ChaGPT</i> como uma ferramenta autônoma de autoaprendizagem, com atenção específica à sua eficácia na resposta a questões de múltipla escolha (MCQs) e no fornecimento de justificativas confiáveis para suas respostas.	Choi (2023), investigou o potencial do <i>ChaGPT</i> como ferramenta autônoma de autoaprendizagem, focando na eficácia do modelo ao responder a questões de múltipla escolha (MCQs) e fornecer justificativas confiáveis. O <i>ChaGPT</i> apresentou uma precisão geral de 76,0%, com desempenho inferior em questões de resolução de problemas em comparação com recordação e interpretação. As justificativas corretas foram fornecidas em 77,8% das respostas, com erros decorrentes principalmente de informações incorretas e raciocínios falhos. No aspecto de referências, o <i>ChaGPT</i> forneceu citações incorretas para 69,7% das respostas, embora 77,0% dessas fossem consideradas pertinentes e precisas em relação ao gabarito, embora a veracidade dos parágrafos de referência não tenha sido verificada diretamente	Contexto: O <i>ChaGPT</i> é um modelo de linguagem grande desenvolvido pela OpenAI que demonstra uma notável capacidade de simular a fala humana. Esta investigação tenta avaliar o potencial do <i>ChaGPT</i> como uma ferramenta autônoma de autoaprendizagem, com atenção específica à sua eficácia em responder a perguntas de múltipla escolha (MCQs) e fornecer justificativas credíveis para suas respostas. Métodos: O estudo utilizou 78 itens de teste do Exame Abrangente de Ciências Médicas Básicas Coreanas (K-CBMSE) dos anos de 2019 a 2021. Os 78 itens de teste traduzidos do coreano para o inglês, com quatro prompts iniciais por item, resultaram em um total de 312 MCQs. As MCQs foram submetidas ao <i>ChaGPT</i> e as respostas foram analisadas quanto à correção, consistência e relevância. Resultados: O <i>ChaGPT</i> respondeu com uma precisão geral de 76,0%. Comparado ao seu desempenho em perguntas de recordação e interpretação, o modelo teve um desempenho inferior em perguntas de resolução de problemas. O <i>ChaGPT</i> ofereceu justificativas corretas para 77,8% (182/234) das respostas, com erros principalmente decorrentes de informações incorretas e raciocínios falhos. Em termos de referências, o <i>ChaGPT</i> forneceu citações incorretas para 69,7% (191/274) das respostas. Embora a veracidade dos parágrafos de referência não pudesse ser verificada, 77,0% (47/61) foram considerados pertinentes e precisos em relação ao gabarito. Conclusão: A versão atual do <i>ChaGPT</i> tem limitações em responder com precisão às MCQs e gerar justificativas corretas e relevantes, especialmente no que se refere às referências. Para evitar possíveis ameaças, como a disseminação de imprecisões e a diminuição das habilidades de pensamento crítico, o <i>ChaGPT</i> deve ser utilizado com supervisão.
09	Large Language Models in Medical Education: Opportunities, Challenges, and Future Directions Modelos de Linguagem de Grande Escala na	Abd-Alrazaq et al. (2023)	Explorar as oportunidades, desafios e direções futuras do uso de Large Language Models (<i>LLMs</i>) na educação médica, utilizando o GPT-4 como estudo de caso. Eles propuseram que os <i>LLMs</i> têm o potencial de revolucionar o desenvolvimento do currículo médico,	Abd-Alrazaq et al. (2023) abordaram as oportunidades, desafios e direções futuras do uso de Large Language Models (<i>LLMs</i>) na educação médica, utilizando o GPT-4 como estudo de caso. Eles propuseram que os <i>LLMs</i> têm o potencial de revolucionar o desenvolvimento do currículo médico,	Resumo: A integração de modelos de linguagem grande (<i>LLMs</i>), como aqueles da série Generative Pre-trained Transformers (GPT), na educação médica tem o potencial de transformar as experiências de aprendizagem dos discentes e elevar seu conhecimento, habilidades e competência. Baseando-se em uma vasta experiência profissional e acadêmica, propomos que os <i>LLMs</i> têm a promessa de revolucionar o desenvolvimento do currículo médico, metodologias de ensino, planos de estudo personalizados e materiais de aprendizagem, avaliações de discentes e muito mais. No entanto, também examinamos criticamente os desafios

	Educação Médica: Oportunidades, Desafios e Direções Futuras			metodologias de ensino, planos de estudo, materiais de aprendizagem personalizados e avaliações de <i>discentes</i> . No entanto, o artigo também destacou desafios críticos, incluindo preconceito algorítmico, confiança excessiva, plágio, desinformação, desigualdade, privacidade e preocupações com direitos autorais. Ao navegar na transição de um paradigma educacional orientado pela informação para um orientado pela inteligência artificial (IA), enfatizaram a importância de compreender tanto o potencial quanto as armadilhas dos <i>LLMs</i> na educação médica. O estudo ofereceu perspectivas fundamentais para orientar futuras recomendações e melhores práticas, promovendo o uso responsável e eficaz de tecnologias de IA na educação médica."	que essa integração pode apresentar, abordando questões de viés algorítmico, dependência excessiva, plágio, desinformação, desigualdade, privacidade e preocupações com direitos autorais na educação médica. À medida que navegamos na transição de um paradigma educacional orientado por informações para um paradigma educacional impulsionado pela inteligência artificial (IA), argumentamos que é fundamental compreender tanto o potencial quanto as armadilhas dos <i>LLMs</i> na educação médica. Este artigo, portanto, oferece nossa perspectiva sobre as oportunidades e desafios do uso de <i>LLMs</i> nesse contexto. Acreditamos que os insights obtidos desta análise servirão como base para futuras recomendações e melhores práticas no campo, promovendo o uso responsável e eficaz das tecnologias de IA na educação médica.
10	Efficacy and limitations of <i>ChaGPT</i> as a biostatistical problem-solving tool in medical education in Serbia: a descriptive study. Eficiência e Limitações do <i>ChaGPT</i> como Ferramenta para Resolução de Problemas Bioestatísticos na Educação Médica na Sérvia: Um Estudo Descritivo.	Ignjatović e Stevanović (2023)	Avaliar o desempenho do <i>ChaGPT</i> (GPT-3.5 e GPT-4) como ferramenta de estudo na resolução de problemas bioestatísticos e identificar quaisquer potenciais desvantagens que possam surgir do uso do <i>ChaGPT</i> na educação médica, particularmente na resolução de problemas bioestatísticos práticos.	Ignjatović e Stevanović (2023), conduziram um estudo para avaliar o desempenho do <i>ChaGPT</i> (GPT-3.5 e GPT-4) como ferramenta de estudo na resolução de problemas bioestatísticos. O GPT-3.5 apresentou sucesso na resolução de 5 problemas práticos na primeira tentativa, abrangendo temas como dados categóricos, estudo transversal, medição de confiabilidade, propriedades de probabilidade e teste t. No entanto, enfrentou dificuldades em questões relacionadas à análise de variância, teste do qui-quadrado e tamanho da amostra em três tentativas. Por outro lado, o GPT-4 demonstrou desempenho aprimorado, resolvendo uma tarefa relacionada ao intervalo de confiança na primeira tentativa e todas as questões em 3 tentativas, com orientação e monitoramento precisos.	Objetivo: Este estudo teve como objetivo avaliar o desempenho do <i>ChaGPT</i> (GPT-3.5 e GPT-4) como ferramenta de estudo na resolução de problemas bioestatísticos e identificar possíveis desvantagens que possam surgir do uso do <i>ChaGPT</i> na educação médica, particularmente na resolução de problemas práticos de bioestatística. Métodos: O <i>ChaGPT</i> foi testado para avaliar sua capacidade de resolver problemas bioestatísticos do <i>Handbook of Medical Statistics</i> de Peacock e Peacock neste estudo descritivo. As tabelas dos problemas foram transformadas em perguntas textuais. Dez problemas bioestatísticos foram escolhidos aleatoriamente e usados como entrada baseada em texto para a conversa com o <i>ChaGPT</i> (versões 3.5 e 4). Resultados: O GPT-3.5 resolveu 5 problemas práticos na primeira tentativa, relacionados a dados categóricos, estudo transversal, medição de confiabilidade, propriedades de probabilidade e o teste t. O GPT-3.5 falhou em fornecer respostas corretas sobre análise de variância, teste qui-quadrado e tamanho amostral dentro de 3 tentativas. O GPT-4 também resolveu uma tarefa relacionada ao intervalo de confiança na primeira tentativa e resolveu todas as perguntas dentro de 3 tentativas, com orientação e monitoramento precisos. Conclusão: A avaliação do desempenho de ambas as versões do <i>ChaGPT</i> em 10 problemas bioestatísticos revelou que o desempenho do GPT-3.5 e 4 ficou abaixo da média, com taxas de resposta correta de 5 e 6 em 10 na primeira tentativa. O GPT-4 conseguiu fornecer todas as respostas corretas dentro de 3 tentativas. Esses achados indicam que os discentes devem estar cientes de que essa ferramenta, mesmo ao fornecer e calcular diferentes análises estatísticas, pode

					estar errada, e devem estar atentos às limitações do ChaGPT e ser cautelosos ao incorporar este modelo na educação médica.
11	A systematic review and meta-analysis on <i>ChaGPT</i> and its utilization in medical and dental research. Uma revisão sistemática e meta-análise sobre o <i>ChaGPT</i> e sua utilização na pesquisa médica e odontológica.	Bagde et al. (2023)	Oferecer uma avaliação completa e baseada em fatos do potencial do <i>ChaGPT</i> como uma ferramenta para pesquisas médicas e odontológicas, que poderia direcionar pesquisas subsequentes e influenciar as práticas clínicas.	Bagde, Dhopre et al. (2023) realizaram uma revisão abrangente com o objetivo de avaliar o potencial do <i>ChaGPT</i> como ferramenta para pesquisas médicas e odontológicas. Analisando 11 estudos descritivos, que abrangiam diferentes domínios, incluindo revisões sistemáticas, câncer, doenças hepáticas, diagnóstico por imagem, educação e vacinação contra COVID-19, a revisão revelou uma variabilidade significativa na precisão do <i>ChaGPT</i> , com intervalos de 18,3% a 100%. A meta-análise destacou uma razão de chances (OR) de 2,25 e um risco relativo (RR) de 1,47 com intervalo de confiança (IC) de 95%, indicando uma significativa superioridade na precisão do <i>ChaGPT</i> em fornecer respostas corretas em comparação com o total de respostas para consultas. No entanto, a presença de heterogeneidade significativa entre os estudos sugere uma variabilidade considerável nos tamanhos dos efeitos entre os estudos incluídos.	Resumo: Desde o seu lançamento, <i>ChaGPT</i> tem causado um grande impacto ao ser utilizado em várias áreas da vida. O principal objetivo desta revisão foi oferecer uma avaliação completa e baseada em fatos do potencial do <i>ChaGPT</i> como uma ferramenta para pesquisa médica e odontológica, o que poderia direcionar pesquisas subsequentes e influenciar práticas clínicas. Métodos: Diferentes bases de dados online foram examinadas em busca de artigos relevantes que estivessem de acordo com os objetivos do estudo. Uma equipe de revisores foi montada para elaborar um framework metodológico adequado para inclusão de artigos e meta-análise. Resultados: Onze estudos descritivos foram considerados para esta revisão, que avaliaram a precisão do <i>ChaGPT</i> em responder a consultas médicas relacionadas a diferentes áreas como revisões sistemáticas, câncer, doenças hepáticas, imagens diagnósticas, educação e vacinação contra COVID-19. Os estudos relataram faixas de precisão diferentes, de 18,3% a 100%, em diversos conjuntos de dados e especialidades. A meta-análise mostrou um odds ratio (OR) de 2,25 e um risco relativo (RR) de 1,47 com um intervalo de confiança (IC) de 95%, indicando que a precisão do <i>ChaGPT</i> em fornecer respostas corretas foi significativamente maior em comparação com o total de respostas para as consultas. No entanto, significativa heterogeneidade estava presente entre os estudos, sugerindo variabilidade considerável nos tamanhos dos efeitos entre os estudos incluídos. Conclusão: As observações indicam que <i>ChaGPT</i> tem a capacidade de fornecer soluções apropriadas para questões nas áreas médica e odontológica, mas pesquisadores e médicos devem avaliar com cautela suas respostas, pois estas nem sempre são confiáveis. Em geral, a importância deste estudo reside em destacar a precisão do <i>ChaGPT</i> nos campos médico e odontológico e enfatizar a necessidade de investigações adicionais para melhorar seu desempenho. © 2017 Elsevier Inc. Todos os direitos reservados.
12	The potential role of <i>ChaGPT</i> and artificial intelligence in anatomy education: a conversation with <i>ChaGPT</i> . Direcionamentos Futuros e Recomendações para a Integração do <i>ChaGPT</i> na Educação Médica.	Totlis et al. (2023)	Determinar suas capacidades e potencial para uso no ensino de anatomia (AE) e na pesquisa de anatomia (AR).	Totlis et al. (2023), realizaram uma entrevista com o <i>ChaGPT</i> para avaliar suas capacidades no ensino de anatomia (AE) e pesquisa de anatomia (AR). O <i>ChaGPT</i> demonstrou habilidades impressionantes ao fornecer descrições anatômicas precisas e bem estruturadas, destacando relações clínicas e estruturais. O <i>Chatbot</i> também ofereceu resumos concisos de capítulos e conselhos sobre terminologia anatômica, mesmo em termos complexos. No entanto, revelou limitações ao abordar variantes anatômicas e seu significado clínico, a menos que essas variantes fossem sistematicamente	Resumo: Um estudo recente publicado no JMIR Med Educ Journal explorou o impacto potencial do Generative Pre-Train (<i>ChaGPT</i>), um modelo de linguagem gerativa, na educação, pesquisa e prática médica. No presente estudo, foi realizada uma entrevista com <i>ChaGPT</i> para determinar suas capacidades e potencial de uso na educação anatômica (AE) e pesquisa anatômica (AR). Métodos: O estudo envolveu 18 perguntas feitas ao <i>ChaGPT</i> após a obtenção de uma assinatura online da 4ª edição. As perguntas foram selecionadas aleatoriamente e avaliadas com base na precisão, relevância e abrangência. Resultados: <i>ChaGPT</i> forneceu descrições anatômicas precisas e bem estruturadas, incluindo relevância clínica e relações entre estruturas. O <i>Chatbot</i> também ofereceu resumos concisos de capítulos e conselhos úteis sobre terminologia anatômica, mesmo com termos complexos. No entanto, quando se tratava de variantes anatômicas e sua significância clínica, as respostas do <i>Chatbot</i> foram inadequadas, a menos que as variantes fossem classificadas sistematicamente em tipos. <i>ChaGPT</i> -4 gerou quizzes de múltipla escolha e questões de correspondência

				<p>classificadas. Além disso, o <i>ChaGPT-4</i> mostrou-se capaz de gerar questionários de múltipla escolha, questões de diversos níveis de dificuldade e resumos de artigos, embora tenha reconhecido suas próprias limitações em termos de precisão, conforme apontado pelos autores do estudo.</p>	<p>de diferentes níveis de dificuldade, além de resumos de artigos quando apresentados com texto. No entanto, o <i>Chatbot</i> reconheceu suas limitações em termos de precisão, assim como os autores do presente estudo. Conclusão: <i>ChaGPT-4</i> pode ser uma ferramenta educacional interativa valiosa para discentes no campo da anatomia, incentivando o engajamento e novas perguntas. No entanto, não pode substituir o papel crítico dos educadores e deve ser usado como uma ferramenta complementar. Pesquisas futuras devem estabelecer diretrizes para o uso e aplicação ótimos do <i>ChaGPT</i> na educação médica.</p>
13	<p>The Utilization of <i>ChaGPT</i> in Reshaping Future Medical Education and Learning Perspectives: A Curse or a Blessing?</p> <p>A Utilização do <i>ChaGPT</i> na Reformulação das Perspectivas Futuras da Educação Médica e do Aprendizado: Uma Maldição ou uma Bênção?</p>	Breeding et al. (2023)	<p>Avaliar como discentes de medicina e leigos avaliam as informações produzidas pelo <i>ChaGPT</i> em comparação com um recurso baseado em evidências em relação ao diagnóstico e tratamento de 5 condições cirúrgicas comuns.</p>	<p>Breeding et al. (2023), conduziram um estudo para avaliar as percepções de discentes de medicina e leigos sobre as informações produzidas pelo <i>ChaGPT</i> em comparação com um recurso baseado em evidências em relação ao diagnóstico e tratamento de 5 condições cirúrgicas comuns. Dos 56 participantes, 50,9% eram discentes de medicina dos EUA, e 49,1% eram da população em geral. Os discentes de medicina relataram que os artigos do <i>ChaGPT</i> foram percebidos como mais claros e melhor organizados do que a fonte baseada em evidências para todas as 5 condições avaliadas, incluindo apendicite, diverticulite, obstrução do intestino delgado, sangramento gastrointestinal superior e SBO. No entanto, apesar dessas percepções positivas, os discentes de medicina ainda consideraram as passagens baseadas em evidências mais confiáveis para todas as condições.</p>	<p>Contexto: <i>ChaGPT</i> tem o potencial de impactar significativamente a educação médica. Nosso estudo tem como objetivo avaliar como discentes de medicina e leigos percebem as informações geradas pelo <i>ChaGPT</i> em comparação com recursos baseados em evidências sobre o diagnóstico e manejo de cinco condições cirúrgicas comuns. Métodos: Realizamos uma pesquisa online anônima com 60 perguntas entre discentes de medicina dos terceiro e quarto anos dos EUA e leigos. Os participantes avaliaram artigos produzidos pelo <i>ChaGPT</i> e fontes baseadas em evidências quanto à clareza, relevância, confiabilidade, validade, organização e abrangência. Cada participante recebeu dois artigos cegos (um de cada fonte) para cada condição cirúrgica. Testes t de amostras pareadas foram utilizados para comparar as avaliações entre as duas fontes. Resultados: Dos 56 participantes da pesquisa, 50,9% (n = 28) eram discentes de medicina dos EUA e 49,1% (n = 27) eram da população em geral. Os discentes de medicina classificaram os artigos do <i>ChaGPT</i> significativamente mais altos em clareza (apendicite: 4,39 vs 3,89, P = 0,020; diverticulite: 4,54 vs 3,68, P < 0,001; obstrução intestinal: 4,43 vs 3,79, P = 0,003; sangramento gastrointestinal: 4,36 vs 3,93, P = 0,020) e organização (diverticulite: 4,36 vs 3,68, P = 0,021; obstrução intestinal: 4,39 vs 3,82, P = 0,033) em comparação com as fontes baseadas em evidências. No entanto, os discentes de medicina consistentemente classificaram as passagens baseadas em evidências como mais abrangentes em todas as cinco condições (colecistite: 4,04 vs 3,36, P = 0,009; apendicite: 4,07 vs 3,36, P = 0,015; diverticulite: 4,07 vs 3,36, P = 0,015; obstrução intestinal: 4,11 vs 3,54, P = 0,030; sangramento gastrointestinal superior: 4,11 vs 3,29, P = 0,003). Conclusão: Os discentes de medicina percebem que os artigos gerados pelo <i>ChaGPT</i> são mais claros e melhor organizados do que as fontes baseadas em evidências sobre a patogênese, diagnóstico e manejo de condições cirúrgicas comuns. No entanto, eles consideram os artigos baseados em evidências significativamente mais abrangentes.</p>
14	Transforming Medical Education: Assessing the Integration of <i>ChaGPT</i> Into Faculty Workflows at a Caribbean Medical School	Cross et al. (2023)	<p>Obter informações sobre a extensão e o tipo de uso do <i>ChaGPT</i> em uma escola de medicina do Caribe, a American University</p>	<p>Cross et al. (2023), investigaram o uso do <i>ChaGPT</i> na American University of Antigua College of Medicine (AUA), buscando compreender a extensão e os tipos de aplicação do <i>ChaGPT</i>. Os resultados revelaram que aproximadamente 33% dos professores na instituição atualmente utilizam o <i>ChaGPT</i>. Houve uma aceitação geral do programa por</p>	<p>Introdução: <i>ChaGPT</i> é um Modelo de Linguagem Grande (LLM) que permite processamento de linguagem natural e interações com usuários em estilo conversacional. Desde seu lançamento em 2022, ele teve um impacto significativo em diversos campos ocupacionais, incluindo a educação médica. Nosso objetivo foi obter insights sobre o tipo e a extensão do uso do <i>ChaGPT</i> na Escola de Medicina da Universidade Americana de Antígua (AUA). Métodos: Administramos um questionário por e-mail a 87 docentes em tempo integral da escola. Utilizamos o software Qualtrics Experience Management (QualtricsXM,</p>

	Transformando a Educação Médica: Avaliando a Integração do <i>ChaGPT</i> nos Fluxos de Trabalho do Corpo Docente em uma Escola de Medicina no Caribe.	of Antigua College of Medicine (AUA).	parte dos usuários, e a maioria dos professores acreditava que o <i>ChaGPT</i> deveria ser uma opção para os <i>discentes</i> . A principal aplicação identificada foi a geração de perguntas de múltipla escolha (MCQ), embora tenha sido observada a preocupação sobre a inclusão de informações incorretas nas saídas do <i>ChaGPT</i> .	Qualtrics, Provo, UT) para quantificar e representar graficamente os resultados. Os resultados da pesquisa foram investigados por meio de comparações de gráficos de barras para números absolutos e porcentagens de várias categorias relacionadas ao uso do <i>ChaGPT</i> , além de estatísticas descritivas para perguntas na escala Likert. Resultados: Descobrimos que aproximadamente 33% dos docentes estavam usando atualmente o <i>ChaGPT</i> . Houve uma aceitação ampla do programa por parte daqueles que o utilizavam, e a maioria acreditava que deveria ser uma opção para os <i>discentes</i> . A principal tarefa para a qual o <i>ChaGPT</i> estava sendo utilizado era a geração de questões de múltipla escolha (QME). A principal preocupação dos docentes era a inclusão de informações incorretas na saída do <i>ChaGPT</i> . Conclusão: <i>ChaGPT</i> foi adotado rapidamente por um subconjunto do corpo docente da faculdade, demonstrando sua crescente aceitação. Dado o nível de aprovação expresso sobre o programa, acreditamos que o <i>ChaGPT</i> continuará a formar uma parte importante e em expansão dos fluxos de trabalho do corpo docente na AUA e na educação médica em geral.
15	Artificial Intelligence in Medical Education: Comparative Analysis of <i>ChaGPT</i> , Bing, and Medical Students in Germany Inteligência Artificial na Educação Médica: Análise Comparativa entre o <i>ChaGPT</i> , Bing e Discentes de Medicina na Alemanha.	Roos et al. (2023)	Avaliar e comparar o desempenho de 3 <i>LLMs</i> , GPT-4, Bing e GPT-3.5-Turbo, nos Exames do Estado Médico Alemão de 2022, além de avaliar como se saíram em comparação com discentes de medicina. Os resultados indicaram que o GPT-4 teve o melhor desempenho global, respondendo corretamente a 88,1% das questões, seguido pelo Bing (86,0%) e GPT-3.5-Turbo (65,7%). Em média, os discentes acertaram 74,6% das questões. Ambos GPT-4 e Bing superaram significativamente os discentes em ambos os exames. Excluindo as questões de mídia, o Bing obteve o melhor desempenho (90,7%), seguido de perto pelo GPT-4 (90,4%), enquanto o GPT-3.5-Turbo ficou para trás (68,2%). Foi observado um declínio significativo no desempenho do GPT-4 e do Bing no exame do outono de 2022, atribuído a uma maior proporção de questões relacionadas à mídia e possivelmente ao aumento da dificuldade das questões	Contexto: Modelos de linguagem grandes (<i>LLMs</i> , na sigla em inglês) têm demonstrado um potencial significativo em diversos domínios, incluindo a medicina. No entanto, há escassez de estudos que examinem seu desempenho em exames médicos, especialmente aqueles realizados em idiomas diferentes do inglês, e em comparação direta com discentes de medicina. Analisar o desempenho dos <i>LLMs</i> nos exames médicos estaduais pode fornecer insights sobre suas capacidades e limitações, além de avaliar seu papel potencial na educação médica e na preparação para exames. Objetivo: Este estudo teve como objetivo avaliar e comparar o desempenho de três <i>LLMs</i> - GPT-4, Bing e GPT-3.5-Turbo - nos Exames Médicos Estaduais da Alemanha de 2022 e avaliar seu desempenho em relação ao dos discentes de medicina. Métodos: Os <i>LLMs</i> foram avaliados em um total de 630 questões dos Exames Médicos Estaduais da Alemanha na primavera e outono de 2022. O desempenho foi avaliado com e sem questões relacionadas a mídia. As análises estatísticas incluíram ANOVA de uma via e testes t independentes para comparações em pares. A força relativa dos <i>LLMs</i> em comparação com a dos discentes também foi avaliada. Resultados: GPT-4 obteve o melhor desempenho geral, respondendo corretamente a 88,1% das questões, seguido de perto por Bing (86,0%) e GPT-3.5-Turbo (65,7%). Os discentes tiveram uma taxa média de respostas corretas de 74,6%. Tanto GPT-4 quanto Bing superaram significativamente os discentes em ambos os exames. Quando questões relacionadas a mídia foram excluídas, Bing alcançou o melhor desempenho de 90,7%, seguido de perto por GPT-4 (90,4%), enquanto GPT-3.5-Turbo ficou atrás (68,2%). Houve uma queda significativa no desempenho de GPT-4 e Bing no exame de outono de 2022, atribuída a uma maior proporção de questões relacionadas a mídia e um potencial aumento na dificuldade das questões. Conclusões: Os <i>LLMs</i> , especialmente GPT-4 e Bing, demonstram potencial como ferramentas valiosas na educação médica e na pré-testagem de questões de exame. Seu alto desempenho, mesmo em comparação com os discentes de medicina,

					indica perspectivas promissoras para desenvolvimentos futuros e integração no cenário educacional e clínico.
16	<p>The Potential and Concerns of Using AI in Scientific Research: <i>ChaGPT</i> Performance Evaluation.</p> <p>O Potencial e as Preocupações do Uso de IA na Pesquisa Científica: Avaliação do Desempenho do <i>ChaGPT</i>.</p>	Khlaif et al. (2023)	<p>Examinar o potencial e as preocupações do uso de IA na pesquisa científica.</p>	<p>Khlaif et al. (2023), realizaram um estudo para examinar o potencial e as preocupações do uso de Inteligência Artificial (IA) na pesquisa científica. Ao gerar artigos de pesquisa de alto impacto, o <i>ChaGPT</i> demonstrou habilidade na produção de relatórios de qualidade quando fornecidas instruções detalhadas e contexto do estudo. Contudo, o impacto do <i>ChaGPT</i> foi menor no desenvolvimento do quadro de investigação e na análise de dados, sendo identificada a necessidade de melhorias no desenvolvimento da revisão da literatura. Além disso, revisores expressaram preocupações em relação à propriedade e integridade da pesquisa ao utilizar texto gerado por IA. Apesar dessas preocupações, reconhece-se que o <i>ChaGPT</i> possui um forte potencial para aumentar a produtividade humana na pesquisa e pode ser uma ferramenta útil na escrita acadêmica.</p>	<p>Contexto: A inteligência artificial (IA) tem muitas aplicações em vários aspectos de nossa vida diária, incluindo saúde, direito criminal, educação, direito civil, negócios e responsabilidade legal. Um aspecto da IA que tem recebido muita atenção é o processamento de linguagem natural (NLP), que se refere à capacidade dos computadores de entender e gerar linguagem humana. Objetivo: Este estudo tem como objetivo examinar o potencial e as preocupações do uso da IA na pesquisa científica. Para isso, foram gerados artigos de pesquisa de alto impacto analisando a qualidade dos relatórios gerados pelo ChaGPT e avaliando o impacto da aplicação no quadro de pesquisa, análise de dados e revisão da literatura. O estudo também explorou preocupações em torno da propriedade e integridade da pesquisa ao usar texto gerado por IA. Métodos: Um total de 4 artigos foram gerados usando o ChaGPT e posteriormente avaliados por 23 revisores. Os pesquisadores desenvolveram um formulário de avaliação para avaliar a qualidade dos artigos gerados. Além disso, foram gerados 50 resumos usando ChaGPT e sua qualidade foi avaliada. Os dados foram submetidos a ANOVA e análise temática para analisar os dados qualitativos fornecidos pelos revisores. Resultados: Quando utilizados prompts detalhados e fornecendo o contexto do estudo, o ChaGPT gerou pesquisas de alta qualidade que poderiam ser publicadas em periódicos de alto impacto. No entanto, o ChaGPT teve um impacto menor no desenvolvimento do quadro de pesquisa e na análise de dados. A área principal que precisa de melhoria foi o desenvolvimento da revisão da literatura. Além disso, os revisores expressaram preocupações com a propriedade e a integridade da pesquisa ao usar texto gerado por IA. Ainda assim, o ChaGPT tem um forte potencial para aumentar a produtividade humana na pesquisa e pode ser usado na escrita acadêmica. Conclusões: O texto gerado por IA tem o potencial de melhorar a qualidade de artigos de pesquisa de alto impacto. Os achados deste estudo sugerem que os tomadores de decisão e pesquisadores devem focar mais na metodologia da pesquisa, que inclui o design de pesquisa, o desenvolvimento de ferramentas de pesquisa e a análise profunda de dados, para tirar implicações teóricas e práticas fortes, estabelecendo assim uma revolução na pesquisa científica na era da IA. As implicações práticas deste estudo podem ser aplicadas em diferentes campos, como a educação médica, para fornecer materiais que desenvolvam as competências básicas tanto para discentes de medicina quanto para membros do corpo docente.</p>

17	<p>Artificial Intelligence in Medicine: Cross-Sectional Study Among Medical Students on Application, Education, and Ethical Aspects.</p> <p>Inteligência Artificial na Medicina: Estudo Transversal entre Discentes de Medicina sobre Aplicações, Educação e Aspectos Éticos.</p>	Weidener e Fischer (2024)	<p>Explorar como discentes de medicina da Alemanha, Áustria e Suíça percebem o uso da IA na medicina e o ensino da IA e da ética da IA na educação médica, de acordo com o uso de aplicativos de bate-papo baseados em IA, como o <i>ChaGPT</i>.</p> <p>Weidener e Fischer (2024), conduziram um estudo para explorar a percepção dos discentes de medicina da Alemanha, Áustria e Suíça sobre o uso da inteligência artificial (IA) na medicina e o ensino de IA e ética de IA na educação médica, com foco em aplicativos de bate-papo baseados em IA, como o <i>ChaGPT</i>. A pesquisa envolveu 487 discentes e revelou uma educação formal limitada sobre IA e ética de IA nos currículos médicos. Apesar disso, 38,8% dos participantes tinham experiência anterior com aplicativos de chat baseados em IA. A maioria (71,7%) previu um impacto positivo da IA na medicina. Houve um consenso (74,9%) sobre a necessidade de instrução sobre IA e ética de IA na educação médica, apesar de considerarem as ofertas atuais inadequadas. Todos os temas propostos para a educação ética em IA foram considerados altamente relevantes."</p>		<p>Contexto: O uso de inteligência artificial (IA) na medicina não apenas implica diretamente a profissão médica, mas também está cada vez mais associado a vários aspectos éticos potenciais. Além disso, a expansão do uso de IA e de aplicativos baseados em IA, como o <i>ChaGPT</i>, demanda uma mudança correspondente na educação médica para preparar adequadamente os futuros profissionais para o uso eficaz dessas ferramentas e abordar os desafios éticos associados que elas apresentam.</p> <p>Objetivo: Este estudo visa explorar como os discentes de medicina da Alemanha, Áustria e Suíça percebem o uso de IA na medicina e o ensino de IA e ética em IA na educação médica, de acordo com seu uso de aplicativos de chat baseados em IA, como o <i>ChaGPT</i>.</p> <p>Métodos: Este estudo transversal, realizado de 15 de junho a 15 de julho de 2023, entrevistou discentes de medicina nos três países por meio de uma pesquisa baseada na web. O objetivo foi avaliar as percepções dos discentes sobre IA na medicina e a integração de IA e ética em IA na educação médica. A pesquisa incluiu 53 itens distribuídos por 6 seções, foi desenvolvida e pré-testada. A análise dos dados utilizou estatísticas descritivas (mediana, moda, intervalo interquartil, número total e porcentagens) e testes qui-quadrado ou U de Mann-Whitney, conforme apropriado.</p> <p>Resultados: A pesquisa com 487 discentes de medicina na Alemanha, Áustria e Suíça revelou uma educação formal limitada em IA ou ética em IA nos currículos médicos, embora 38,8% (189/487) tivessem experiência prévia com aplicativos de chat baseados em IA, como o <i>ChaGPT</i>. Apesar das exposições prévias variadas, 71,7% (349/487) anteciparam um impacto positivo da IA na medicina. Houve um consenso generalizado (385/487, 74,9%) sobre a necessidade de instrução em IA e ética em IA na educação médica, embora as ofertas atuais fossem consideradas inadequadas. Em relação ao conteúdo de ética em IA, todos os tópicos propostos foram considerados altamente relevantes.</p> <p>Conclusões: Este estudo revelou uma discrepância pronunciada entre o uso de aplicativos de chat baseados em IA, como o <i>ChaGPT</i>, entre discentes de medicina na Alemanha, Áustria e Suíça, e o ensino de IA na educação médica. Para preparar adequadamente os futuros profissionais médicos, há uma necessidade urgente de integrar o ensino de IA e ética em IA nos currículos médicos.</p> <p>Palavras-chave: ética em IA; tecnologia em IA; inteligência artificial; ética; currículo médico; educação médica; faculdade de medicina; medicina.</p>
18	<p>AI-Enabled Medical Education: Threads of Change, Promising Futures, and Risky Realities Across Four Potential Future Worlds.</p>	Knopp et al. (2023)	<p>Planejar a interseção dos cuidados de saúde e da educação médica no futuro.</p>		<p>Knopp et al. (2023), propuseram uma perspectiva para planejar a interseção dos cuidados de saúde e da educação médica no futuro, utilizando o modelo GPT da OpenAI. Criaram quatro cenários distintos: 'Harmonia de IA', 'Conflito de IA', 'O mundo do equilíbrio ecológico' e 'Risco existencial'. Os riscos</p>

	Educação Médica Habilida por Inteligência Artificial: Fios de Mudança, Futuros Promissores e Realidades Arriscadas em Quatro Potenciais Mundos Futuros.			associados incluem desinformação, perda de privacidade, aumento da desigualdade, erosão da autonomia humana e dilemas éticos. Por outro lado, os benefícios potenciais envolvem maior eficiência, intervenções personalizadas, maior colaboração, detecção precoce e investigação acelerada.	cuidados de saúde e educação médica no futuro. Métodos: Utilizamos o GPT-4 e técnicas de planejamento estratégico baseadas em cenários para elaborar quatro mundos hipotéticos influenciados pela integração da IA nos cuidados de saúde e na educação médica. Este método, usado por organizações como a Shell e o Conselho de Acreditação para Educação Médica de Pós-Graduação, avalia a prontidão para futuros alternativos e gerencia efetivamente a incerteza, o risco e a oportunidade. Os cenários detalhados oferecem insights sobre os ambientes potenciais que a profissão médica pode enfrentar e estabelecem a base para a geração de hipóteses e construção de ideias sobre a implementação responsável de IA. Resultados: Foram criados os seguintes quatro mundos usando o modelo GPT da OpenAI: Harmonia de IA, Conflito de IA, O mundo do Equilíbrio Ecológico e Risco Existencial. Os riscos incluem desinformação e má informação, perda de privacidade, aumento da desigualdade, erosão da autonomia humana e dilemas éticos. Os benefícios envolvem melhoria da eficiência, intervenções personalizadas, colaboração aprimorada, detecção precoce e pesquisa acelerada. Conclusões: Para garantir o uso responsável da IA, os autores sugerem focar em três áreas-chave: desenvolver um robusto quadro ético, promover a colaboração interdisciplinar e investir em educação e treinamento. Um forte quadro ético enfatiza a segurança do paciente, privacidade e autonomia, ao mesmo tempo que promove equidade e inclusividade. A colaboração interdisciplinar incentiva a cooperação entre diversos especialistas no desenvolvimento e implementação de tecnologias de IA, garantindo que elas abordem as necessidades complexas e desafios nos cuidados de saúde e na educação médica. Investir em educação e treinamento prepara profissionais e trainees com habilidades e conhecimentos necessários para usar e avaliar criticamente as tecnologias de IA de forma eficaz. A integração da IA nos cuidados de saúde e na educação médica representa um ponto crítico entre avanços transformadores e riscos significativos. Ao trabalharmos juntos para abordar riscos e consequências imediatas e de longo prazo, podemos garantir que a integração da IA conduza a um futuro mais equitativo, sustentável e próspero para os cuidados de saúde e a educação médica. A medida que nos envolvemos com tecnologias de IA, nossas ações coletivas determinarão o estado futuro dos cuidados de saúde e da educação médica para aproveitar o poder da IA enquanto garantimos a segurança e o bem-estar da humanidade. Palavras-chave: ChaGPT; GPT-4; OpenAI; inteligência artificial; autônomo; autonomia; ética; ética em IA; futuro; futuro dos cuidados de saúde; educação médica; privacidade; cenário; planejamento estratégico.
19	Medical students' patterns of using ChaGPT as a feedback tool and perceptions of	Park (2023)	Analizar padrões de uso do ChaGPT antes e depois de atividades em grupo e explorar as	No estudo conduzido por Park (2023), os resultados indicaram que o uso do ChaGPT como ferramenta de feedback em atividades de grupo teve impactos positivos no aprendizado dos discentes de medicina. A análise revelou	Objetivo: Este estudo teve como objetivo analisar os padrões de uso do ChaGPT antes e depois das atividades em grupo, além de explorar as percepções dos discentes de medicina sobre o ChaGPT como uma ferramenta de feedback na sala de aula. Métodos: O estudo envolveu 99 discentes do segundo ano de pré-medicina que participaram de um curso de "Liderança e Comunicação" de março a junho de

	<p><i>ChaGPT in a Leadership and Communication course in Korea: a cross-sectional study.</i></p> <p>Padrões de Uso do <i>ChaGPT</i> como Ferramenta de Feedback e Percepções de Discentes de Medicina em um Curso de Liderança e Comunicação na Coreia: Um Estudo Transversal.</p>	<p>habilidades de discentes de medicina. percepções do <i>ChaGPT</i> como ferramenta de feedback em sala de aula.</p>	<p>que os <i>discentes</i> consideraram o feedback do <i>ChaGPT</i> útil, resultando em revisões e reenvios das respostas do grupo. A satisfação com o feedback, incluindo correção, utilidade e ética, mostrou diferenças significativas dependendo do uso do <i>ChaGPT</i> durante as aulas, enquanto não foram observadas diferenças significativas com base no gênero ou na experiência prévia dos <i>discentes</i> com o <i>ChaGPT</i>. As vantagens mais destacadas foram relacionadas à capacidade do <i>ChaGPT</i> em fornecer respostas às perguntas e resumir informações, embora houvesse preocupações sobre a produção de informações sem evidências de apoio. Esses resultados sugerem que a integração do <i>ChaGPT</i> pode ser benéfica no contexto educacional, especialmente quando aplicado após a primeira rodada de discussão para revisões.</p>	<p>2023. Os <i>discentes</i> participaram de atividades individuais e em grupo relacionadas a estratégias de negociação. O <i>ChaGPT</i> foi utilizado para fornecer feedback sobre suas soluções. Uma pesquisa foi administrada para avaliar as percepções dos discentes sobre o feedback do <i>ChaGPT</i>, seu uso na sala de aula, e os pontos fortes e desafios do <i>ChaGPT</i> de 17 a 19 de maio de 2023. Resultados: Os discentes indicaram que o feedback do <i>ChaGPT</i> foi útil, revisando e reenviando suas respostas de grupo de várias maneiras após receberem o feedback. A maioria dos respondentes concordou com o uso do <i>ChaGPT</i> durante a aula. A resposta mais comum sobre o contexto apropriado para usar o feedback do <i>ChaGPT</i> foi "após a primeira rodada de discussão, para revisões". Houve diferença significativa na satisfação com o feedback do <i>ChaGPT</i>, incluindo correção, utilidade e ética, dependendo se o <i>ChaGPT</i> foi usado ou não durante a aula, mas não houve diferença significativa conforme o gênero ou se os discentes tinham experiência prévia com o <i>ChaGPT</i>. As maiores vantagens foram "fornecer respostas para questões" e "sumarizar informações", e a maior desvantagem foi "produzir informações sem evidências de suporte". Conclusão: Os discentes estavam cientes das vantagens e desvantagens do <i>ChaGPT</i>, e tiveram uma atitude positiva em relação ao uso do <i>ChaGPT</i> na sala de aula. Palavras-chave: Inteligência artificial; Feedback; Liderança; Discentes de medicina; República da Coreia.</p>
20	<p>Evaluating <i>ChaGPT</i>'s Ability to Solve Higher-Order Questions on the Competency-Based Medical Education Curriculum in Medical Biochemistry.</p> <p>Avaliando a Capacidade do <i>ChaGPT</i> em Resolver Questões de Ordem Superior no Currículo de Educação Médica Baseada em Competências em Bioquímica Médica.</p>	Ghosh e Bir, 2023	<p>Determinar se o <i>ChaGPT</i> pode resolver problemas de ordem superior relacionados à bioquímica médica.</p>	<p>No estudo de Ghosh e Bir (2023), o <i>ChaGPT</i> foi avaliado quanto à sua capacidade de resolver questões de ordem superior em bioquímica médica. Os resultados indicaram uma pontuação média de 4,0, atribuída por dois especialistas em bioquímica usando uma escala de zero a cinco. Embora a pontuação tenha sido inferior ao máximo hipotético de cinco ($p=0,001$) e comparável a quatro ($p=0,16$) em um teste de classificação sinalizada de Wilcoxon, os resultados sugerem que o <i>ChaGPT</i> demonstrou habilidades promissoras na resolução de problemas complexos de bioquímica médica. Não houve diferença significativa nas respostas entre diferentes módulos do currículo de Educação Médica Baseada em Competências (CBME) em bioquímica médica (Kruskal-Wallis $p=0,39$), e a confiabilidade entre os avaliadores foi excelente ($ICC=0,926$; IC 95%: 0,814-0,971; $F=19$; $p=0,001$). Esses achados indicam uma consistência notável nas</p> <p>Contexto: A inteligência artificial (IA) relacionada à saúde está em desenvolvimento. A capacidade do sistema de realizar processos cognitivos sofisticados, como resolução de problemas, tomada de decisões, raciocínio e percepção, é referida como pensamento cognitivo superior na IA. Este tipo de pensamento exige mais do que apenas processamento de fatos; também envolve compreender e trabalhar com ideias abstratas, avaliar e aplicar dados relevantes ao contexto, e gerar novos insights com base em aprendizados anteriores e experiência. <i>ChaGPT</i> é um software conversacional baseado em inteligência artificial que pode interagir com pessoas para responder perguntas e usa modelos de processamento de linguagem natural. A plataforma tem gerado grande interesse mundial e continua estabelecendo uma tendência contínua na resolução de muitos problemas complexos em várias dimensões. No entanto, a capacidade do <i>ChaGPT</i> de responder corretamente a questões que exigem pensamento de nível superior em bioquímica médica ainda não foi investigada. Portanto, esta pesquisa teve como objetivo avaliar a aptidão do <i>ChaGPT</i> para responder a perguntas de ordem superior em bioquímica médica. Objetivo: Neste estudo, nosso objetivo foi determinar se o <i>ChaGPT</i> pode abordar problemas de ordem superior relacionados à bioquímica médica. Métodos: Este estudo transversal foi realizado online, conversando com a versão atual do <i>ChaGPT</i> (14 de março de 2023, que é atualmente gratuita para usuários registrados). Foram apresentadas 200 questões de raciocínio em bioquímica médica que requerem pensamento de ordem superior. Essas perguntas foram escolhidas aleatoriamente do banco de questões da</p>

			<p>avaliações do <i>ChaGPT</i> por diferentes especialistas.</p>	<p>instituição e classificadas de acordo com os módulos de competências do currículo de Educação Médica Baseada em Competências (CBME). As respostas foram coletadas e arquivadas para pesquisas subsequentes. Dois acadêmicos especialistas em bioquímica examinaram as respostas em uma escala de zero a cinco. A precisão das pontuações foi determinada por um teste de classificação de sinais de Wilcoxon de amostra única usando valores hipotéticos. Resultado: O software de IA respondeu a 200 questões que exigiam pensamento de ordem superior com uma pontuação mediana de 4,0 (Q1=3,50, Q3=4,50). Usando um teste de classificação de sinais de Wilcoxon de amostra única, o resultado foi menor que o máximo hipotético de cinco ($p=0,001$) e comparável a quatro ($p=0,16$). Não houve diferença nas respostas a questões de diferentes módulos CBME em bioquímica médica (Kruskal-Wallis $p=0,39$). A confiabilidade interavaliadora das pontuações atribuídas por dois membros do corpo docente de bioquímica foi excelente ($ICC=0,926$ (IC 95%: 0,814-0,971); $F=19$; $p=0,001$). Conclusão: Os resultados desta pesquisa indicam que o <i>ChaGPT</i> tem o potencial de ser uma ferramenta eficaz para responder a perguntas que exigem pensamento de ordem superior em bioquímica médica, com uma pontuação mediana de quatro em cinco. No entanto, o treinamento contínuo e o desenvolvimento com dados de avanços recentes são essenciais para melhorar o desempenho e torná-lo funcional para o campo em constante crescimento do uso acadêmico médico. Palavras-chave: bioquímica, <i>ChaGPT</i>, MCQs, resolução de questões de múltipla escolha, habilidades cognitivas de ordem superior, educação médica baseada em competências, bioquímica médica, medicina laboratorial, inteligência artificial, educação médica.</p>	
21	<p>Assessment of Resident and AI <i>Chatbot</i> Performance on the University of Toronto Family Medicine Residency Progress Test: Comparative Study.</p> <p>Avaliação do Desempenho de Residentes e <i>Chatbot</i> de IA no Teste de Progresso da Residência em Medicina de Família da Universidade de</p>	Huang et al. (2023)	<p>Comparar quantitativa e qualitativamente o desempenho dos residentes de Medicina Geral e Familiar (GPT-3.5, GPT-4 e Medicina de Família) em um teste de conhecimentos médicos de múltipla escolha adequado ao nível de residente de Medicina de Família.</p>	<p>No estudo de Huang et al. (2023), o GPT-4 demonstrou um desempenho significativamente superior ao GPT-3.5 ao responder a questões de conhecimentos médicos para residentes de Medicina Geral e Familiar. O GPT-4 alcançou uma precisão de 82,4%, em comparação com 57,4% do GPT-3.5, representando uma diferença de 25,0%. Além disso, o GPT-4 mostrou pontuações mais altas em todas as 11 categorias de conhecimento em Medicina de Família. Na análise qualitativa, o GPT-4 destacou-se ao fornecer justificativas para suas escolhas em 86,1% das respostas, enquanto o GPT-3.5 teve apenas 16,7%. Ambos os modelos cometem erros lógicos comumente, sendo esses os mais frequentes, enquanto erros aritméticos foram menos comuns. Em comparação com a média de desempenho dos residentes de Medicina</p>	<p>Contexto: <i>Chatbots</i> baseados em modelos de linguagem grandes (LLM), exemplificados pelo <i>ChaGPT</i> (especificamente GPT-3.5 e GPT-4), evoluíram rapidamente em suas capacidades para tarefas de propósito geral e aplicações educacionais, incluindo testes de conhecimento médico. Objetivo: Este estudo teve como objetivo comparar o desempenho do GPT-3.5, GPT-4 e residentes de Medicina de Família em um teste de conhecimento médico de múltipla escolha projetado para o nível de um residente de Medicina de Família, tanto de forma quantitativa quanto qualitativa. Métodos: O Teste de Progresso do Departamento de Medicina de Família e Comunidade da Universidade de Toronto, composto por questões de múltipla escolha, foi administrado ao GPT-3.5 e GPT-4. As respostas dos <i>Chatbots</i> de IA foram avaliadas quanto à correção, extensão da resposta, tempo gasto, fornecimento de justificativa e tipos de erros. Os resultados foram comparados com um grupo de residentes de Medicina de Família que realizaram o mesmo teste simultaneamente. Resultados: O GPT-4 superou significativamente o GPT-3.5, respondendo corretamente a 82,4% (89/108) das perguntas em comparação com 57,4% (62/108) para o GPT-3.5 (diferença de 25,0%, IC 95% 16,3%-32,8%; teste de McNemar: $P<0,001$). O GPT-4 também obteve pontuações mais altas em todas as 11 categorias de conhecimento em Medicina de Família. Em 86,1% das respostas, o GPT-4 forneceu uma justificativa para a resposta</p>

	Toronto: Estudo Comparativo.			Geral e Familiar (56,9%), o GPT-3.5 teve um desempenho semelhante ($P = 0,16$), mas o GPT-4 superou significativamente até mesmo o residente com melhor desempenho em Medicina de Família ($P < 0,001$). Esses resultados indicam que o GPT-4 possui potencial como uma ferramenta eficaz para avaliação de conhecimentos médicos para residentes em Medicina de Família.	escolhida, em comparação com 16,7% para o GPT-3.5. Qualitativamente, erros lógicos foram os mais comuns tanto nas respostas do GPT-3.5 quanto do GPT-4. Conclusões: O GPT-4 demonstra um desempenho superior tanto ao GPT-3.5 quanto aos residentes de Medicina de Família em um teste de conhecimento médico de múltipla escolha. Além de alcançar alta precisão, o GPT-4 fornece justificações coerentes para suas respostas, sendo uma ferramenta promissora para a educação médica, incluindo criação de questões e disseminação de conhecimento. Palavras-chave: IA; GPT-3.5; GPT-4; LLM; processamento de linguagem natural; inteligência artificial; exame educacional; medicina de família; modelo de linguagem grande; educação médica; exame de conhecimento médico.
22	Examining the Threat of <i>ChaGPT</i> to the Validity of Short Answer Assessments in an Undergraduate Medical Program Analizando a Ameaça do <i>ChaGPT</i> para a Validade de Avaliações de Resposta Curta em um Programa Médico de Graduação.	Morjaria et al. (2023)	Avaliar a extensão da ameaça representada pelo <i>ChaGPT</i> à validade dos problemas de avaliação de respostas curtas usados para examinar discentes de medicina pré-estagiários em nosso programa de graduação em medicina.	No estudo de Morjaria et al. (2023), as respostas geradas pelo <i>ChaGPT</i> para problemas de avaliação de respostas curtas alcançaram uma pontuação média de 3,29 em 5, superando significativamente os <i>discentes</i> com notas mínimas de aprovação, que obtiveram uma média de 2,38. No entanto, o desempenho do <i>ChaGPT</i> foi inferior às pontuações médias históricas da turma (média 3,67) quando consideradas todas as respostas anteriores, independentemente do desempenho dos <i>discentes</i> . Não foram identificadas tendências estatisticamente significativas no desempenho entre os domínios da Taxonomia de Bloom. Esses resultados destacam a necessidade de avaliar cuidadosamente a validade das respostas geradas pelo <i>ChaGPT</i> em contextos específicos de avaliação médica.	Objetivos O <i>ChaGPT</i> é um modelo de inteligência artificial capaz de interpretar prompts de texto livre e fornecer respostas detalhadas e humanas em uma ampla gama de temas. Este estudo avaliou o grau de ameaça representado pelo <i>ChaGPT</i> para a validade dos problemas de avaliação de resposta curta utilizados para examinar discentes de medicina pré-clínica em nosso programa de educação médica de graduação. Métodos: Quarenta problemas usados em avaliações anteriores de discentes foram recuperados e estratificados pelos níveis da Taxonomia de Bloom. Trinta desses problemas foram submetidos ao <i>ChaGPT</i> -3.5. Para os 10 problemas restantes, foram recuperadas respostas de discentes que obtiveram apenas a nota mínima de aprovação no passado. Seis tutores avaliaram cada uma das 40 respostas. A comparação do desempenho entre as respostas geradas pelo <i>discentes</i> e as geradas pelo <i>ChaGPT</i> , agregadas como um todo e agrupadas pelos níveis de raciocínio cognitivo da Taxonomia de Bloom, foi feita usando testes t, ANOVA, alfa de Cronbach e d de Cohen. As pontuações das respostas geradas pelo <i>ChaGPT</i> também foram comparadas com o desempenho médio histórico da turma. Resultados: As respostas geradas pelo <i>ChaGPT</i> receberam uma pontuação média de 3,29 de 5 ($n = 30$, IC 95% 2,93-3,65) em comparação com 2,38 para um grupo de discentes que atingiram as notas mínimas de aprovação ($n = 10$, IC 95% 1,94-2,82), representando um desempenho superior ($P = 0,008$, $\eta^2 = 0,169$), mas foram superadas pelas pontuações médias históricas da turma nos mesmos 30 problemas (média 3,67, $P = 0,018$), incluindo todas as respostas passadas independentemente do nível de desempenho dos <i>discentes</i> . Não houve tendência estatisticamente significativa no desempenho entre os domínios da Taxonomia de Bloom. Conclusão: Embora o <i>ChaGPT</i> tenha conseguido passar nos problemas de avaliação de resposta curta que abrangem o currículo pré-clínico, ele superou apenas os discentes com desempenho abaixo da média. Observamos que, em vários casos, os tutores estavam convencidos de que as respostas produzidas pelo <i>ChaGPT</i> foram produzidas pelos <i>discentes</i> . Os riscos para a validade da avaliação incluem a incerteza na identificação de discentes com dificuldades e a incapacidade de intervir de maneira oportuna. O desempenho do <i>ChaGPT</i> em problemas que exigem maior demanda de raciocínio cognitivo merece mais pesquisas. Palavras-chave: <i>ChaGPT</i> ; inteligência artificial; discente de medicina; pré-clínico; avaliação de resposta curta.

23	The Significance of Artificial Intelligence Platforms in Anatomy Education: An Experience With <i>ChaGPT</i> and Google Bard. A Significância das Plataformas de Inteligência Artificial na Educação Anatômica: Uma Experiência com <i>ChaGPT</i> e Google Bard.	Ilgaz e Çelik, 2023	Avaliar o uso de dois grandes modelos de linguagem (<i>LLMs</i>), <i>ChaGPT</i> e Google Bard, no ensino de	No estudo conduzido por Ilgaz e Çelik (2023), que avaliou o uso de dois grandes modelos de linguagem (<i>LLMs</i>), <i>ChaGPT</i> e Google Bard, no ensino de anatomia, observou-se que ambos os modelos conseguiram realizar tarefas como responder perguntas e gerar questões de múltipla escolha com diversos graus de precisão. Não houve diferença significativa entre <i>ChaGPT</i> e Google Bard na resposta a perguntas, e ambos apresentaram habilidade para gerar questões de múltipla escolha com alto grau de precisão. No entanto, o desempenho na redação de artigos ainda não atingiu um nível satisfatório. O estudo enfatizou a necessidade de cautela ao utilizar <i>LLMs</i> na educação médica devido ao fato de estarem em constante desenvolvimento, podendo gerar informações imprecisas ou enganosas em algumas situações.	Este estudo avaliou o uso de dois modelos de linguagem grandes (<i>LLMs</i>), <i>ChaGPT</i> e Google Bard, no ensino de anatomia. Os modelos foram solicitados a responder perguntas, gerar questões de múltipla escolha e escrever artigos sobre temas de anatomia. Os resultados mostraram que os modelos foram capazes de realizar essas tarefas com diferentes graus de precisão. <i>ChaGPT</i> e Google Bard não apresentaram diferenças significativas na capacidade de responder perguntas. Ambos os modelos conseguiram gerar questões de múltipla escolha com alto grau de precisão. No entanto, o desempenho dos modelos na escrita de artigos ainda não estava em um nível suficiente. O estudo também encontrou que o uso de <i>LLMs</i> na educação médica requer cautela, pois esses modelos ainda estão em desenvolvimento e, às vezes, podem gerar informações imprecisas ou enganosas. É crucial avaliar cuidadosamente os resultados dos <i>LLMs</i> antes de utilizá-los em ambientes educacionais. Em resumo, o estudo concluiu que os <i>LLMs</i> têm o potencial de ser ferramentas valiosas para o ensino de anatomia. No entanto, são necessárias mais pesquisas para melhorar a precisão dos modelos e compreender melhor como podem ser empregados de maneira eficaz em contextos educacionais.
24	Analysing the Applicability of <i>ChaGPT</i> , Bard, and Bing to Generate Reasoning-Based Multiple-Choice Questions in Medical Physiology. Analizando a Aplicabilidade do <i>ChaGPT</i> , Bard e Bing para Gerar Questões de Múltipla Escolha Baseadas em Raciocínio em Fisiologia Médica.	Agarwal, Sharma e Goswami (2023)	Avaliar e comparar a aplicabilidade de <i>ChaGPT</i> , Bard e Bing na geração de MCQs baseados em raciocínio para discentes de graduação do MBBS (Bacharelado em Medicina, Bacharelado em Cirurgia) na área de fisiologia. .	No estudo de Agarwal, Sharma e Goswami (2023), que buscou avaliar e comparar a eficácia do <i>ChaGPT</i> , Bard e Bing na geração de questões de múltipla escolha (MCQs) baseadas em raciocínio para discentes de graduação do MBBS na área de fisiologia, observou-se que <i>ChaGPT</i> e Bard geraram 110 MCQs, enquanto o Bing produziu apenas 100. A validade das MCQs foi significativamente diferente entre os modelos, com pontuações medianas de 3 para <i>ChaGPT</i> , 3 para Bard e 3 para Bing ($p<0,001$). A dificuldade das questões também apresentou diferença significativa, sendo classificada como 1 para <i>ChaGPT</i> , 1 para Bard e 1 para Bing ($p=0,006$). Contudo, a capacidade de raciocínio necessária para responder às questões não diferiu significativamente entre os modelos ($p=0,235$), com classificações de 1 para <i>ChaGPT</i> , 1 para Bard e 1 para Bing. O coeficiente Kappa (K) foi $\geq 0,8$ para todos os parâmetros em todos os modelos de IA, indicando uma concordância substancial.	Contexto: A inteligência artificial (IA) está evoluindo no sistema de educação médica. <i>ChaGPT</i> , Google Bard e Microsoft Bing são modelos baseados em IA que podem resolver problemas na educação médica. No entanto, a aplicabilidade da IA na criação de questões de múltipla escolha baseadas em raciocínio no campo da fisiologia médica ainda precisa ser explorada. Objetivo: Avaliar e comparar a aplicabilidade de <i>ChaGPT</i> , Bard e Bing na geração de questões de múltipla escolha baseadas em raciocínio para discentes de graduação em MBBS (Bachelor of Medicine, Bachelor of Surgery) no tema de fisiologia. Métodos: A Comissão Médica Nacional da Índia desenvolveu um currículo de fisiologia com 11 módulos e várias competências. Dois fisiologistas escolheram independentemente uma competência de cada módulo. O terceiro fisiologista solicitou que as três IA gerassem cinco questões de múltipla escolha para cada competência escolhida. Os dois fisiologistas que forneceram as competências classificaram as questões de múltipla escolha geradas pelas IA em uma escala de 0 a 3 para validade, dificuldade e capacidade de raciocínio necessária para respondê-las. A média dos dois escores foi analisada usando o teste de Kruskal-Wallis para comparar a distribuição das respostas totais e por módulo, seguido por um teste post-hoc para comparações pareadas. O coeficiente Kappa de Cohen (K) foi usado para avaliar a concordância nos escores entre os dois avaliadores. Os dados foram expressos como mediana com intervalo interquartílico, e a significância estatística foi determinada por um valor de $p < 0,05$. Resultados: <i>ChaGPT</i> e Bard geraram 110 questões de múltipla escolha para as competências escolhidas, enquanto Bing forneceu apenas 100 questões, falhando em gerar para duas competências. A

				Esses achados destacam as disparidades na qualidade e validade das questões geradas pelos diferentes modelos de IA na área de fisiologia.	validade das questões de múltipla escolha foi avaliada como 3 (3-3) para ChaGPT, 3 (1,5-3) para Bard e 3 (1,5-3) para Bing, mostrando uma diferença significativa ($p < 0,001$) entre os modelos. A dificuldade das questões foi avaliada como 1 (0-1) para ChaGPT, 1 (1-2) para Bard e 1 (1-2) para Bing, com uma diferença significativa ($p = 0,006$). A capacidade de raciocínio necessária para responder às questões foi avaliada como 1 (1-2) para ChaGPT, 1 (1-2) para Bard e 1 (1-2) para Bing, sem diferença significativa ($p = 0,235$). O coeficiente Kappa foi $\geq 0,8$ para todos os três parâmetros em todos os três modelos de IA. Conclusão: A IA ainda precisa evoluir para gerar questões de múltipla escolha baseadas em raciocínio na fisiologia médica. ChaGPT, Bard e Bing mostraram certas limitações. Bing gerou as questões de múltipla escolha menos válidas significativamente, enquanto ChaGPT gerou as questões menos difíceis significativamente.
25	Assessment of Artificial Intelligence Performance on the Otolaryngology Residency In-Service Exam. Avaliação do Desempenho da Inteligência Artificial no Exame de Residência em Otorrinolaringologia.	Mahajan et al. (2023)	Determinar o uso potencial e a confiabilidade de um grande modelo de aprendizagem de idiomas para responder a perguntas práticas em otorrinolaringologia - cirurgia de cabeça e pescoço, e avaliar sua eficácia para estagiários e <i>discentes</i> de cirurgia, foi constatado que o <i>ChaGPT</i> apresentou uma taxa de respostas corretas de 53% e uma taxa de explicações corretas de 54%. Os resultados indicam uma tendência de diminuição na precisão das respostas e explicações à medida que a dificuldade das perguntas aumenta. Essas descobertas ressaltam a importância de considerar o nível de complexidade das questões ao avaliar o desempenho do <i>ChaGPT</i> nesta área específica da medicina.	No estudo conduzido por Mahajan et al. (2023), que buscou determinar o uso potencial e a confiabilidade de um grande modelo de aprendizagem de idiomas para responder a perguntas práticas em otorrinolaringologia - cirurgia de cabeça e pescoço, e avaliar sua eficácia para estagiários e <i>discentes</i> de cirurgia, foi constatado que o <i>ChaGPT</i> apresentou uma taxa de respostas corretas de 53% e uma taxa de explicações corretas de 54%. Os resultados indicam uma tendência de diminuição na precisão das respostas e explicações à medida que a dificuldade das perguntas aumenta. Essas descobertas ressaltam a importância de considerar o nível de complexidade das questões ao avaliar o desempenho do <i>ChaGPT</i> nesta área específica da medicina.	Objetivos: Este estudo visa determinar o potencial uso e confiabilidade de um modelo de aprendizado de linguagem ampla para responder questões em uma área subespecializada da medicina, especificamente em questões de exames práticos em cirurgia de otorrinolaringologia e cabeça e pescoço, e avaliar sua eficácia atual para residentes e aprendizes cirúrgicos. Desenho do Estudo e Configuração: Todas as questões disponíveis de um banco de questões de acesso pago ao público foram inseridas manualmente no <i>ChaGPT</i> . Métodos: As saídas do <i>ChaGPT</i> foram comparadas com o padrão das respostas e explicações do banco de questões. As questões foram avaliadas em dois domínios: precisão e abrangência das explicações. Resultados: No geral, nosso estudo demonstra uma taxa de resposta correta pelo <i>ChaGPT</i> de 53% e uma taxa de explicação correta de 54%. Observamos que, com o aumento da dificuldade das questões, há uma diminuição na taxa de precisão das respostas e explicações. Conclusão: Atualmente, plataformas de aprendizado impulsionadas por inteligência artificial não são robustas o suficiente para serem recursos confiáveis de educação médica para auxiliar aprendizes em cenários específicos de tomada de decisão de pacientes em subespecialidades. Palavras-chave: BoardVitals; <i>ChaGPT</i> ; inteligência artificial; exames internos; modelos de linguagem ampla; treinamento de residência em otorrinolaringologia.
26	<i>ChaGPT</i> versus human in generating medical graduate exam multiple choice questions-A multinational prospective study (Hong Kong S.A.R., Singapore,	Cheung et al. (2023)	Avaliar a qualidade das questões de múltipla escolha (MCQs) produzidas pelo <i>ChaGPT</i> para uso em exames médicos de pós-graduação, em comparação com exames médicos de pós-graduação, em comparação com	No estudo de Cheung et al. (2023), o <i>ChaGPT</i> demonstrou eficiência ao gerar 50 questões de múltipla escolha (MCQs) para exames médicos de pós-graduação em um tempo significativamente menor em comparação com examinadores humanos. Embora tenha sido observada uma pontuação inferior do <i>ChaGPT</i> em relação à relevância das questões em comparação com as elaboradas por humanos	Introdução: Modelos de linguagem ampla, especialmente o <i>ChaGPT</i> , têm demonstrado notáveis capacidades de processamento de linguagem. Diante da carga substancial de trabalho do corpo docente universitário de medicina, este estudo visa avaliar a qualidade de questões de múltipla escolha (QME) produzidas pelo <i>ChaGPT</i> para uso em exames de pós-graduação em medicina, em comparação com questões elaboradas por professores universitários com base em livros-texto médicos padrão. Métodos: Foram geradas 50 QME pelo <i>ChaGPT</i> com referência a dois livros-texto médicos de graduação (Harrison's e Bailey & Love's). Outras 50 QME foram elaboradas por dois professores universitários

	<p>Ireland, and the United Kingdom).</p> <p><i>ChaGPT</i> versus humano na geração de questões de múltipla escolha para exame de graduação em medicina - Um estudo prospectivo multinacional (Hong Kong S.A.R., Singapura, Irlanda e Reino Unido)</p>	<p>questões escritas por equipes de professores universitários com base em livros médicos padrão.</p>	<p>em um domínio específico, não houve diferença significativa na qualidade global, considerando a pontuação total e outros domínios. Destaca-se que as questões geradas pelo <i>ChaGPT</i> apresentaram uma gama mais ampla de pontuações, enquanto as elaboradas por humanos foram mais consistentes. Esses resultados ressaltam a eficiência do <i>ChaGPT</i> na geração de MCQs, com algumas variações em comparação com as elaboradas por humanos.</p>	<p>usando os mesmos livros-texto médicos. Todas as 100 QME foram numeradas individualmente, randomizadas e enviadas a cinco avaliadores internacionais independentes para avaliação da qualidade das QME utilizando um escore padronizado em cinco domínios de avaliação: adequação da questão, clareza e especificidade, relevância, poder discriminatório das alternativas e adequação para exame de pós-graduação em medicina. Resultados: O tempo total necessário para o <i>ChaGPT</i> criar as 50 questões foi de 20 minutos e 25 segundos, enquanto dois examinadores humanos levaram um total de 211 minutos e 33 segundos para redigir as 50 questões. Ao comparar a média do escore entre as questões construídas pela inteligência artificial e as elaboradas por humanos, apenas no domínio de relevância que a inteligência artificial foi inferior aos humanos (IA: 7,56 +/- 0,94 vs humano: 7,88 +/- 0,52; p = 0,04). Não houve diferença significativa na qualidade das questões entre as questões elaboradas pela IA versus as elaboradas por humanos, tanto no escore total de avaliação quanto nos outros domínios. As questões geradas pela IA apresentaram uma faixa mais ampla de escores, enquanto as criadas por humanos foram consistentes e dentro de uma faixa mais estreita. Conclusão: O <i>ChaGPT</i> tem o potencial de gerar QME de qualidade comparável para exames de pós-graduação em medicina em um tempo significativamente menor.</p>	
27	<p>Success of <i>ChaGPT</i>, an AI language model, in taking the French language version of the European Board of Ophthalmology examination: A novel approach to medical knowledge assessment.</p> <p>Êxito do <i>ChaGPT</i>, um modelo de linguagem de IA, ao realizar a versão em francês do exame do Conselho Europeu de Oftalmologia: Uma abordagem inovadora</p>	<p>Panthier e Gatinel (2023)</p>	<p>Avaliar o desempenho do <i>ChaGPT</i>, um modelo de linguagem de inteligência artificial (IA) de ponta desenvolvido pela OpenAI, na conclusão com êxito da versão em francês do exame do Conselho Europeu de Oftalmologia (EBO) e avaliar seu potencial.</p>	<p>No estudo realizado por Panthier e Gatinel (2023), que teve como objetivo avaliar o desempenho do <i>ChaGPT</i> no exame do Conselho Europeu de Oftalmologia (EBO) em francês, observou-se que o <i>ChaGPT</i> alcançou uma taxa de sucesso de 91%. Esses resultados indicam um alto nível de competência do modelo de linguagem de inteligência artificial no conhecimento e aplicação da oftalmologia. A IA demonstrou precisão ao fornecer respostas corretas em todas as categorias de perguntas, abrangendo ciências básicas, conhecimento clínico e gestão clínica. Além disso, o <i>ChaGPT</i> respondeu às perguntas de forma rápida, exigindo apenas uma fração do tempo necessário em comparação com os participantes humanos. Esses achados destacam o potencial do <i>ChaGPT</i> como uma ferramenta eficaz na avaliação do conhecimento oftalmológico e na educação médica.</p>	<p>Resumo: O objetivo deste estudo foi avaliar o desempenho do <i>ChaGPT</i>, um modelo de linguagem artificial de ponta desenvolvido pela OpenAI, na conclusão bem-sucedida da versão em francês do exame da European Board of Ophthalmology (EBO) e avaliar seu papel potencial na educação médica e avaliação do conhecimento. Métodos: O <i>ChaGPT</i>, baseado na arquitetura GPT-4, foi exposto a uma série de questões do exame da EBO em francês, abrangendo vários aspectos da oftalmologia. O desempenho do AI foi avaliado comparando suas respostas com as respostas corretas fornecidas por especialistas em oftalmologia. Além disso, o estudo avaliou o tempo que o <i>ChaGPT</i> levou para responder a cada pergunta como medida de eficiência. Resultados: O <i>ChaGPT</i> alcançou uma taxa de sucesso de 91% no exame da EBO, demonstrando um alto nível de competência em conhecimento e aplicação da oftalmologia. O AI forneceu respostas corretas em todas as categorias de questões, indicando um forte entendimento das ciências básicas, conhecimento clínico e gestão clínica. O modelo de IA também respondeu rapidamente às perguntas, levando apenas uma fração do tempo necessário pelos candidatos humanos. Conclusão: O desempenho do <i>ChaGPT</i> na versão em francês do exame da EBO demonstra seu potencial para ser uma ferramenta valiosa na educação médica e avaliação do conhecimento. Pesquisas adicionais são necessárias para explorar formas ótimas de implementar modelos de linguagem de IA na educação médica e abordar as preocupações éticas e práticas associadas.</p>

	para a avaliação do conhecimento médico.				
28	The Intersection of <i>ChaGPT</i> , Clinical Medicine, and Medical Education. A Interseção entre o <i>ChaGPT</i> , Medicina Clínica e Educação Médica.	Wong (2023)	Apresentar a perspectiva dos autores sobre a integração de modelos de IA como o <i>ChaGPT</i> na medicina clínica e na educação médica	Wong et al. (2023), delineiam sua perspectiva sobre a incorporação de modelos de inteligência artificial, como o <i>ChaGPT</i> , na medicina clínica e na educação médica. Destacando a capacidade excepcional do <i>ChaGPT</i> em gerar respostas semelhantes às humanas, aprimoradas por meio da Aprendizagem por Reforço com Feedback Humano, o artigo sugere que essa abordagem pode remodelar significativamente as metodologias pedagógicas na formação médica. A revisão abrangente e as experiências pessoais dos autores elucidam os aspectos positivos e negativos, bem como as considerações éticas do uso do <i>ChaGPT</i> na prática clínica, com ênfase nas implicações para a educação médica. Em um contexto de transformação impulsionada pela IA, o artigo destaca a necessidade de explorar como essa tecnologia pode potencialmente aumentar a capacidade humana na criação e disseminação de conhecimento, revolucionando a educação médica e a prática clínica. A importância de manter a integridade acadêmica e os padrões profissionais é ressaltada, juntamente com a necessidade crucial de estabelecer diretrizes claras para o uso ético e responsável das tecnologias de IA na medicina clínica e na educação médica."	As avançadas tecnologias de inteligência artificial (IA), especificamente modelos gerativos de linguagem como o <i>ChaGPT</i> (OpenAI), continuam a se desenvolver, suas implicações potenciais em todos os setores, incluindo a medicina, são cada vez mais evidentes. Este artigo de opinião visa apresentar a perspectiva dos autores sobre a integração de modelos de IA, como o <i>ChaGPT</i> , na medicina clínica e na educação médica. A capacidade sem precedentes do <i>ChaGPT</i> em gerar respostas semelhantes às humanas, refinadas através do Aprendizado por Reforço com Feedback Humano, poderia remodelar significativamente as metodologias pedagógicas dentro da educação médica. Através de uma revisão abrangente e das experiências pessoais dos autores, este artigo de opinião explora os prós, contras e considerações éticas do uso do <i>ChaGPT</i> na medicina clínica, destacando suas implicações para a educação médica. Essa exploração é crucial em uma era transformadora, onde a IA pode potencialmente ampliar as capacidades humanas no processo de criação e disseminação do conhecimento, revolucionando a educação médica e a prática clínica. A importância de manter a integridade acadêmica e os padrões profissionais é enfatizada, assim como a relevância de estabelecer diretrizes claras para o uso responsável e ético das tecnologias de IA na medicina clínica e na educação médica.
29	Teaching AI Ethics in Medical Education: A Scoping Review of Current Literature and Practices.	Weidener e Fischer (2023)	Representar e sintetizar a literatura sobre o ensino de ética em IA como parte da educação médica.	Weidener e Fischer (2023), abordam as crescentes preocupações éticas relacionadas ao aumento do uso da Inteligência Artificial (IA) na medicina, incluindo questões como autonomia do paciente, preconceito e transparência. Em resposta a essas preocupações, estudos recentes indicam a necessidade de integrar o ensino de ética em IA nos currículos médicos. Sua revisão de	Introduction: Com o aumento do uso da Inteligência Artificial (IA) na medicina, surgiram preocupações éticas, como autonomia do paciente, viés e transparência. Estudos recentes sugerem a necessidade de ensinar ética em IA como parte dos currículos médicos. Esta revisão abrangente teve como objetivo representar e sintetizar a literatura sobre o ensino de ética em IA como parte da educação médica. Métodos: As diretrizes PRISMA-SCR e a metodologia JBI guiaram uma busca na literatura em quatro bases de dados (PubMed, Embase, Scopus e Web of Science) nos últimos 22 anos (2000-2022). Para incluir aplicações de chat baseadas em IA, como o <i>ChaGPT</i> , a busca foi atualizada para incluir publicações

	Ensino de Ética em Inteligência Artificial na Educação Médica: Uma Revisão Abrangente da Literatura Atual e Práticas.			escopo buscou representar e sintetizar a literatura sobre o ensino dessa ética como parte da educação médica. Inicialmente, 1.384 publicações foram identificadas, e, após triagem, 87 foram avaliadas na íntegra, resultando na inclusão de 10 publicações para análise. Atualizações bibliográficas adicionaram duas publicações relevantes de 2023 ao conjunto de dados. Todas as 12 publicações revisadas recomendaram o ensino da ética da IA nos currículos médicos, destacando as potenciais implicações dessa tecnologia na medicina. Os desafios éticos antecipados, especialmente o preconceito, foram identificados como a base essencial para o ensino, junto com os princípios fundamentais de ética médica. Recomenda-se a abordagem de ensino baseada em casos, envolvendo exemplos do mundo real em seminários interativos e pequenos grupos, para proporcionar uma educação mais eficaz nesse domínio."	até o final de junho de 2023. Originalmente, foram identificadas 1384 publicações, das quais 87 foram avaliadas na íntegra. Após essa avaliação, 10 publicações foram incluídas para análise adicional. Na busca atualizada, duas publicações adicionais relevantes de 2023 foram identificadas e incluídas na análise. Todas as 12 publicações recomendaram o ensino de ética em IA nos currículos médicos devido às potenciais implicações da IA na medicina. Desafios éticos antecipados, como viés, foram identificados como base recomendada para o conteúdo de ensino, além dos princípios básicos de ética médica. O ensino baseado em casos, utilizando exemplos do mundo real em seminários interativos e grupos pequenos, foi recomendado como modalidade de ensino. Conclusão: Esta revisão abrangente revela uma escassez de literatura sobre o ensino de ética em IA na educação médica, com grande parte da literatura disponível sendo recente e teórica. Esses achados destacam a importância de mais estudos empíricos e definições fundamentais de ética em IA para orientar o desenvolvimento de conteúdos e modalidades de ensino. Reconhecendo o impacto significativo da IA na medicina, são necessárias mais pesquisas sobre o ensino de ética em IA na educação médica para preparar adequadamente os discentes de medicina para os desafios éticos futuros.
30	New Artificial Intelligence <i>ChaGPT</i> Performs Poorly on the 2022 Self-assessment Study Program for Urology. Novo <i>ChaGPT</i> de Inteligência Artificial Apresenta Desempenho Insatisfatório no Programa de Estudo de Autoavaliação para Urologia de 2022.	Huynh, Bonebrake e Schultis (2023)	Avaliar o uso do <i>ChaGPT</i> no Programa de Estudo de Autoavaliação da American Urological Association como complemento educacional para estagiários de urologia e médicos praticantes. Selecionando 150 questões do exame de 2022, o estudo envolveu a codificação das respostas do <i>ChaGPT</i> como corretas, incorretas ou indeterminadas, com a regeneração permitida até duas vezes para respostas indeterminadas. O <i>ChaGPT</i> obteve uma taxa de acerto de 26,7% em questões abertas e 28,2% em questões de múltipla escolha. Embora a regeneração tenha reduzido as respostas indeterminadas, a proporção de respostas corretas não aumentou. O estudo também destacou que o <i>ChaGPT</i> forneceu justificativas consistentes para respostas	Huynh, Bonebrake e Schultis (2023), investigaram o uso do <i>ChaGPT</i> no Programa de Estudo de Autoavaliação da American Urological Association como complemento educacional para estagiários de urologia e médicos praticantes. Selecionando 150 questões do exame de 2022, o estudo envolveu a codificação das respostas do <i>ChaGPT</i> como corretas, incorretas ou indeterminadas, com a regeneração permitida até duas vezes para respostas indeterminadas. O <i>ChaGPT</i> obteve uma taxa de acerto de 26,7% em questões abertas e 28,2% em questões de múltipla escolha. Embora a regeneração tenha reduzido as respostas indeterminadas, a proporção de respostas corretas não aumentou. O estudo também destacou que o <i>ChaGPT</i> forneceu justificativas consistentes para respostas	Os grandes modelos de linguagem têm demonstrado capacidades impressionantes, mas sua aplicação na medicina ainda não está clara. Nós buscamos avaliar o uso do <i>ChaGPT</i> no Programa de Estudo de Autoavaliação da Associação Americana de Urologia como um complemento educacional para residentes de urologia e médicos em prática. MÉTODOS: Cento e cinquenta questões do exame do Programa de Estudo de Autoavaliação de 2022 foram selecionadas, e aquelas que continham recursos visuais (n=15) foram removidas. Os itens restantes foram codificados como abertos ou de múltipla escolha. A saída do <i>ChaGPT</i> foi codificada como correta, incorreta ou indeterminada; se indeterminada, as respostas foram regeneradas até 2 vezes. Concordância, qualidade e precisão foram determinadas por 3 pesquisadores independentes e revisadas por 2 médicos avaliadores. Uma nova sessão foi iniciada para cada entrada para evitar aprendizado cruzado. RESULTADOS: O <i>ChaGPT</i> foi correto em 36/135 (26,7%) perguntas abertas e 38/135 (28,2%) perguntas de múltipla escolha. Respostas indeterminadas foram geradas em 40 (29,6%) e 4 (3,0%), respectivamente. Das respostas corretas, 24/36 (66,7%) e 36/38 (94,7%) foram na saída inicial, 8

				incorrectas e manteve a concordância entre respostas corretas e incorretas para ambas as categorias de questões.	(22,2%) e 1 (2,6%) na segunda saída, e 4 (11,1%) e 1 (2,6%) na saída final, respectivamente. Embora a regeneração tenha reduzido as respostas indeterminadas, a proporção de respostas corretas não aumentou. Para perguntas abertas e de múltipla escolha, o ChaGPT forneceu justificativas consistentes para respostas incorretas e manteve concordância entre respostas corretas e incorretas. CONCLUSÕES: O ChaGPT mostrou promessa anteriormente em exames de licenciamento médico; no entanto, sua aplicação no Programa de Estudo de Autoavaliação de 2022 não foi demonstrada. O desempenho melhorou com perguntas de múltipla escolha em relação às perguntas abertas. Mais importante ainda foram as justificativas persistentes para respostas incorretas - deixado sem controle, a utilização do ChaGPT na medicina pode facilitar a desinformação médica.
31	Artificial intelligence: How will <i>ChaGPT</i> and other AI applications change our everyday medical practice? Inteligência artificial: Como o <i>ChaGPT</i> e outras aplicações de IA irão mudar nossa prática médica cotidiana?	Sonntagbauer, Haar e Kluge (2023)	Explicar a funcionalidade básica dos grandes modelos de linguagem (LLM), apresentar suas aplicações na medicina, e discutir os possíveis perigos associados ao uso da IA.	No trabalho de Sonntagbauer, Haar e Kluge (2023), é oferecida uma explicação abrangente sobre a funcionalidade básica dos grandes modelos de linguagem (LLM), seguida por uma análise das possíveis aplicações do <i>ChaGPT</i> na medicina. O artigo abrange também uma perspectiva e discussão sobre os potenciais perigos associados às aplicações de Inteligência Artificial (IA) na área médica. O aumento notável no uso de aplicações de IA em trabalhos científicos, especialmente na redação científica, é destacado, com uma consideração especial para a possível ampla aplicação do LLM na documentação médica. O texto enfatiza a funcionalidade técnica dessas aplicações, explorando a utilização do <i>ChaGPT</i> como sistema de apoio ao diagnóstico. No entanto, alerta-se para o risco de disseminação e consolidação de imprecisões e preconceitos por meio da implementação do LLM. A discussão ressalta a necessidade pendente de regulamentação diante dessa nova tecnologia, destacando questões importantes relacionadas à ética e à qualidade das informações geradas por esses modelos.	Com a disponibilização gratuita do robô de chat "ChaGPT" pela OpenAI em novembro de 2022, a aplicação da Inteligência Artificial (IA) tornou-se acessível a todos. Este desenvolvimento representa uma inovação significativa, especialmente no uso de Modelos de Linguagem Avançados (LLM), como o ChaGPT. Questão de Estudo: Este artigo explora o funcionamento básico dos Modelos de Linguagem Avançados e examina as opções de aplicação do ChaGPT na medicina. Segue-se uma discussão sobre as possíveis ameaças das aplicações de IA. Material e Métodos: Resolução de problemas utilizando o ChaGPT com exemplos concretos. Análise e discussão da literatura científica relevante. Resultados: Observa-se um aumento significativo no uso de aplicações de IA na pesquisa científica, especialmente na redação científica. Existe a possibilidade de ampla aplicação de Modelos de Linguagem Avançados na documentação médica, como sistemas de apoio ao diagnóstico. Contudo, há o risco de disseminação de imprecisões e vieses através do uso desses modelos. A regulamentação desta nova tecnologia ainda está em desenvolvimento. Conclusão: Aplicações de IA como o ChaGPT têm o potencial de transformar profundamente o cotidiano médico. É crucial um debate aprofundado sobre esta tecnologia e uma avaliação cuidadosa de suas oportunidades e desafios. Palavras-chave: Big Data, Sistemas de Apoio ao Diagnóstico, Aprendizado de Máquina, Digitalização na Medicina, Modelos de Linguagem Avançados
32	From human writing to artificial intelligence generated text:	Dergaa, Chamari e Zmijewski (2023)	Explorar os benefícios e ameaças potenciais do <i>ChaGPT</i> e	O estudo de Dergaa, Chamari e Zmijewski (2023), aborda a análise das implicações éticas e potenciais impactos do <i>ChaGPT</i> e outras	Este manuscrito explora os potenciais benefícios e riscos do uso de modelos avançados de inteligência artificial, como o Chat Generative Pre-trained Transformer (ChaGPT), e outras tecnologias de processamento de linguagem

	<p>examining the prospects and potential threats of <i>ChaGPT</i> in academic writing.</p> <p>Da escrita humana ao texto gerado por inteligência artificial: examinando as perspectivas e potenciais ameaças do <i>ChaGPT</i> na escrita acadêmica.</p>	<p>outras tecnologias de PNL em escritos acadêmicos e publicações de pesquisa; (ii) destaca as considerações éticas envolvidas na utilização destas ferramentas e (iii) considera o impacto que podem ter na autenticidade e credibilidade do trabalho acadêmico.</p>	<p>tecnologias de Processamento de Linguagem Natural (PNL) na escrita acadêmica e na pesquisa. A revisão, baseada em uma abordagem quase qualitativa, examina a literatura de artigos acadêmicos indexados no Scopus, usando palavras-chave específicas. A análise revela o potencial das tecnologias de PNL, incluindo o <i>ChaGPT</i>, para melhorar a eficiência e a qualidade da redação acadêmica. No entanto, o estudo destaca preocupações significativas relacionadas à autenticidade e credibilidade do trabalho acadêmico, enfatizando a importância de debates éticos abrangentes. Recomenda-se cautela aos acadêmicos no uso dessas ferramentas, com ênfase na transparência e na preservação dos princípios éticos e acadêmicos, reconhecendo o papel crucial da inteligência humana e do pensamento crítico no processo de pesquisa.</p>	<p>natural (NLP) na escrita acadêmica e na publicação de pesquisas. O objetivo é destacar considerações éticas e avaliar seu impacto na autenticidade e credibilidade do trabalho acadêmico. O estudo realizou uma revisão da literatura de artigos revisados por pares indexados no Scopus, focando em palavras-chave como "ChaGPT", "texto gerado por IA", "escrita acadêmica" e "processamento de linguagem natural". Utilizando uma abordagem quasi-qualitativa, dados relevantes foram avaliados criticamente para responder às questões de pesquisa. Os resultados indicam que o ChaGPT e tecnologias de NLP similares podem aumentar a eficiência acadêmica ao realizar diversas tarefas linguísticas e gerar respostas semelhantes às humanas. No entanto, seu uso levanta preocupações quanto à autenticidade e credibilidade das produções acadêmicas. O estudo enfatiza a necessidade de discussões abrangentes sobre os benefícios potenciais, ameaças e limitações dessas ferramentas. Sublinha a importância de manter padrões éticos e acadêmicos, defendendo a integração da inteligência humana e do pensamento crítico no processo de pesquisa. Em conclusão, o estudo recomenda o uso cauteloso de ferramentas de IA na academia, promovendo transparência em sua aplicação. Destaca o papel essencial do intelecto humano e da análise crítica na preservação da integridade do trabalho acadêmico diante dos avanços tecnológicos em NLP e IA.</p>	
33	<p><i>ChaGPT</i> Utility in Healthcare Education, Research, and Practice: Systematic Review on the Promising Perspectives and Valid Concerns.</p> <p>Utilidade do <i>ChaGPT</i> na Educação em Saúde, Pesquisa e Prática: Revisão Sistemática sobre Perspectivas Promissoras e Preocupações Válidas.</p>	<p><u>Sallam</u> (2023)</p>	<p>Investigar a utilidade do <i>ChaGPT</i> na educação, pesquisa e prática em saúde e destacar suas potenciais limitações.</p>	<p>A revisão sistemática conduzida por Sallam (2023), explorou a utilidade do <i>ChaGPT</i> na educação, pesquisa e prática em saúde, além de destacar suas potenciais limitações. Dos 60 registros elegíveis, 85% citaram benefícios, incluindo melhor redação científica, aumento da equidade na pesquisa, eficiência na análise de conjuntos de dados e geração de código, economia de tempo na pesquisa em saúde, otimização do fluxo de trabalho na prática de cuidados de saúde e melhor aprendizagem personalizada na educação em saúde. No entanto, 96,7% dos registros expressaram preocupações, abrangendo questões éticas, de direitos autorais, transparência e legais, risco de preconceito, plágio, falta de originalidade, conteúdo impreciso, conhecimento limitado, citações incorretas, segurança cibernética e risco de infodemia. Embora o <i>ChaGPT</i> possa induzir mudanças paradigmáticas nas áreas mencionadas, a revisão enfatiza a necessidade de uma abordagem cautelosa na adoção desse <i>Chatbot</i> de IA, recomendando a formulação de</p>	<p>ChaGPT é um modelo de linguagem grande (LLM) conversacional baseado em inteligência artificial (IA). As aplicações potenciais de <i>LLMs</i> na educação, pesquisa e prática em saúde podem ser promissoras se as preocupações válidas associadas forem examinadas e abordadas proativamente. Esta revisão sistemática teve como objetivo investigar a utilidade de ChaGPT na educação, pesquisa e prática em saúde, além de destacar suas potenciais limitações. Utilizando as diretrizes PRISMA, foi realizada uma busca sistemática para recuperar registros em inglês no PubMed/MEDLINE e Google Scholar (pesquisas publicadas ou pré-impressões) que examinaram ChaGPT no contexto da educação, pesquisa ou prática em saúde. Um total de 60 registros foram incluídos. Os benefícios de ChaGPT foram mencionados em 51/60 (85,0%) registros e incluíram: (1) melhoria na escrita científica e no aumento da equidade e versatilidade da pesquisa; (2) utilidade na pesquisa em saúde (análise eficiente de conjuntos de dados, geração de código, revisões de literatura, economia de tempo para focar no desenho experimental, descoberta e desenvolvimento de medicamentos); (3) benefícios na prática em saúde (otimização do fluxo de trabalho, economia de custos, documentação, medicina personalizada e melhoria da literacia em saúde); e (4) benefícios na educação em saúde, incluindo melhoria na aprendizagem personalizada e foco em pensamento crítico e aprendizado baseado em problemas. As preocupações quanto ao uso de ChaGPT foram relatadas em 58/60 (96,7%) registros, incluindo questões éticas, de direitos autorais, transparência e legais, risco de viés, plágio, falta de originalidade, conteúdo impreciso com risco de criação de informações falsas, conhecimento limitado, citações incorretas, problemas de cibersegurança e risco de infodemia. As aplicações promissoras de</p>

				um código de ética para orientar seu uso responsável na saúde e na academia.	ChaGPT podem induzir mudanças paradigmáticas na educação, pesquisa e prática em saúde. No entanto, a adoção deste <i>Chatbot</i> de IA deve ser realizada com extrema cautela, considerando suas potenciais limitações. Atualmente, ChaGPT não se qualifica para ser listado como autor em artigos científicos, a menos que as diretrizes do ICMJE/COPE sejam revisadas ou modificadas. É urgente uma iniciativa envolvendo todos os stakeholders na educação, pesquisa e prática em saúde para estabelecer um código de ética que oriente o uso responsável de ChaGPT e outros <i>LLMs</i> na saúde e na academia.
34	The Now and Future of <i>ChaGPT</i> and GPT in Psychiatry O Presente e o Futuro do <i>ChaGPT</i> e GPT em Psiquiatria.	Cheng et al. (2023)	Fornecer insights sobre o estado atual das aplicações <i>ChaGPT</i> no campo da psiquiatria e imaginar um futuro potencial dos cuidados digitais de saúde mental <i>por meio</i> da integração e dos avanços na tecnologia GPT.	A pesquisa conduzida por Cheng, Chang e Chang et al (2023), destaca o potencial transformador das aplicações do <i>ChaGPT</i> no cenário da psiquiatria, delineando perspectivas promissoras para o futuro dos cuidados digitais de saúde mental. A habilidade do <i>ChaGPT</i> de desempenhar funções além do seu papel inicial como <i>Chatbot</i> genérico em cuidados de saúde mental sugere possibilidades significativas. O treinamento avançado em linguagem emerge como uma ferramenta valiosa, capaz de contribuir em áreas como diagnóstico assistido, psicoterapia clínica e detecção precoce de sinais de alerta. No entanto, é crucial salientar que a eficácia dessas aplicações depende da implementação cuidadosa de padrões éticos robustos, visando assegurar a privacidade e a confidencialidade dos pacientes. A colaboração estreita entre profissionais de saúde mental, pesquisadores e desenvolvedores é imperativa para uma integração eficiente da tecnologia GPT na evolução dos cuidados de saúde mental.	ChaGPT tem suscitado extensas discussões na comunidade de saúde desde seu lançamento em novembro de 2022. No entanto, as potenciais aplicações no campo da psiquiatria têm recebido atenção limitada. O aprendizado profundo tem demonstrado benefícios para a psiquiatria, e o GPT é um poderoso modelo de linguagem baseado em aprendizado profundo com imenso potencial para este campo. Apesar da conveniência do ChaGPT, este avançado <i>Chatbot</i> atualmente possui aplicações práticas limitadas na psiquiatria. Ele pode ser utilizado para apoiar psiquiatras em tarefas rotineiras como completar prontuários médicos, facilitar comunicações entre clínicos e com pacientes, aprimorar textos acadêmicos e apresentações, e programar e realizar análises para pesquisa. O treinamento e a aplicação atual do ChaGPT requerem o uso de prompts apropriados para maximizar os resultados adequados e minimizar imprecisões prejudiciais e erros fantasmagóricos. Além disso, avanços futuros do GPT que incorporam empatia, reconhecimento de emoções, avaliação de personalidade e detecção de sinais de alerta de saúde mental são essenciais para sua integração eficaz no cuidado psiquiátrico. No futuro próximo, o desenvolvimento de um sistema de psicoterapia totalmente automatizado, treinado para comunicação especializada (como verbatim de psicoterapia), é concebível com base na tecnologia GPT fundamental. Este sistema ideal deverá integrar entradas práticas do "mundo real" e interfaces amigáveis de usuário e paciente via algoritmos clinicamente validados, módulos de compreensão/geração de voz e algoritmos de discriminação de emoções baseados em expressões faciais e entradas fisiológicas de dispositivos vestíveis. Além dos desafios tecnológicos, acreditamos ser crucial estabelecer padrões éticos geralmente aceitos para a aplicação de ferramentas relacionadas ao ChaGPT em todos os ambientes de saúde mental, incluindo telemedicina e ambientes acadêmicos/treinamento.
35	Diagnostic and Management Applications of <i>ChaGPT</i> in Structured Otolaryngology Clinical Scenarios	Qu, Qureshi e Petersen (2023)	Avaliar as aplicações clínicas e limitações do transformador pré-treinado gerativo de chat (<i>ChaGPT</i>) em otorrinolaringologia.	A investigação realizada por Qu, Qureshi e Petersen (2023) evidenciam a eficácia do <i>ChaGPT</i> na otorrinolaringologia. Com uma participação significativa de onze médicos assistentes (taxa de resposta de 61%), o <i>ChaGPT</i> demonstrou habilidade em lidar com casos de dificuldade variada, conforme indicado pela classificação das vinhetas.	Resumo: Objetivo: Avaliar as aplicações clínicas e limitações do chat generative pretrained transformer (<i>ChaGPT</i>) em otorrinolaringologia. Desenho do Estudo: Estudo transversal. Ambiente: Centro acadêmico terciário. Métodos: O <i>ChaGPT</i> 4.0 foi consultado para diagnósticos e planos de manejo de 20 vinhetas clínicas escritas por médicos em otorrinolaringologia. Médicos supervisores foram então solicitados a avaliar a dificuldade das vinhetas clínicas e concordância com os diagnósticos diferenciais e planos de manejo das respostas do <i>ChaGPT</i> em uma escala Likert de 5 pontos. Estatísticas resumidas foram calculadas. Regressão

	Aplicações Diagnósticas e de Gerenciamento do <i>ChaGPT</i> em Cenários Clínicos Estruturados de Otorrinolaringologia.			Houve alta concordância com os diagnósticos diferenciais e planos de tratamento propostos pelo <i>ChaGPT</i> , evidenciando sua utilidade percebida pelos profissionais. A consistência na performance do <i>ChaGPT</i> , independentemente da dificuldade das vinhetas, sugere sua robustez. A correlação entre escores mais baixos no diagnóstico e tratamento destaca a importância de analisar a confiabilidade em ambas as fases do processo clínico. Esses achados contribuem para a compreensão das aplicações clínicas do <i>ChaGPT</i> em otorrinolaringologia, fornecendo insights valiosos para o seu desenvolvimento contínuo na prática médica.	ordinal univariada foi realizada entre a dificuldade da vinheta e a qualidade dos diagnósticos e planos de manejo. Resultados: Onze médicos supervisores completaram a pesquisa (taxa de resposta de 61%). Globalmente, as vinhetas foram classificadas como muito fáceis até dificuldade neutra (intervalo de escore mediano: 1,00-4,00; mediana geral 2,00). Houve alta concordância com o diagnóstico diferencial fornecido pelo <i>ChaGPT</i> (intervalo de escore mediano: 3,00-5,00; mediana geral: 5,00). Também houve alta concordância com os planos de tratamento (intervalo de escore mediano: 3,00-5,00; mediana geral: 5,00). Não houve associação entre a dificuldade da vinheta e a concordância com o diagnóstico diferencial ou tratamento. Menores escores de diagnóstico tiveram maiores chances de ter menores escores de tratamento. Conclusão: Modelos geradores de inteligência artificial como o <i>ChaGPT</i> estão sendo rapidamente adotados na medicina. O desempenho com cenários de otorrinolaringologia curados, de fácil a moderada dificuldade, indica alta concordância com médicos para diagnóstico e manejo. No entanto, uma diminuição na qualidade do diagnóstico está associada a uma diminuição na qualidade do manejo. Mais pesquisas são necessárias sobre a capacidade do <i>ChaGPT</i> em lidar com informações clínicas não estruturadas.
36	Using artificial intelligence to create diverse and inclusive medical case vignettes for education Utilizando a inteligência artificial para criar vinhetas de casos médicos diversas e inclusivas para fins educacionais.	Bakkum et al. (2023)	Criar uma quantidade significativa de casos autênticos e diversos que possam ser reutilizados e personalizados como REA em contextos educacionais locais. Além disso, esforça-se por melhorar a acessibilidade da criação de casos utilizando IA para todos os educadores médicos, especialmente aqueles que ensinam farmacologia clínica e terapêutica que estão ligados através do EurOP 2 E.	Bakkum, Hartjes, Piët et al (2023), destacam em sua abordagem a meta central de criar uma quantidade substancial de casos autênticos e diversificados, visando sua reutilização e personalização como Recursos Educacionais Autênticos (REA) em contextos educacionais locais. Além disso, a iniciativa visa aprimorar a acessibilidade da criação de casos utilizando Inteligência Artificial (IA) para todos os educadores médicos, especialmente aqueles envolvidos no ensino de farmacologia clínica e terapêutica, conectados através do EurOP 2 E. A geração consistente de casos para diversas tarefas, compreendendo conjuntos de 30 casos por vez, é realizada utilizando os prompts mencionados. O processo, que inclui verificações e formatações obrigatórias, é concluído em aproximadamente 60 minutos por conjunto, destacando a eficiência e praticidade dessa abordagem para a criação de material educacional relevante e personalizado.	Objetivos: As vinhetas de casos médicos desempenham um papel crucial na educação médica, porém frequentemente falham em representar autenticamente pacientes diversos. Além disso, essas vinhetas tendem a simplificar em excesso a complexa relação entre as características dos pacientes e as condições médicas, levando a perspectivas tendenciosas e potencialmente prejudiciais entre os discentes. A inclusão de aspectos da diversidade dos pacientes, como etnia, em casos escritos é desafiadora. Além disso, a criação desses casos impõe uma carga significativa aos professores em termos de trabalho e tempo. Nossa objetivo é explorar o potencial de casos clínicos assistidos por inteligência artificial (IA) para acelerar a criação de casos e melhorar a diversidade, incluindo fotografias geradas por IA para uma representação mais realista dos pacientes. Métodos: Neste estudo, utilizamos o <i>ChaGPT</i> (OpenAI, GPT 3.5) para desenvolver vinhetas de casos médicos diversas e inclusivas. Avaliamos diferentes abordagens e identificamos um conjunto de oito prompts consecutivos que podem ser facilmente personalizados para atender a contextos locais e tarefas específicas. Para melhorar a representação visual, utilizamos o Adobe Firefly beta para geração de imagens. Resultados: Usando os prompts descritos, consistentemente geramos casos para diversas tarefas, produzindo conjuntos de 30 casos por vez. Garantimos a inclusão de verificações obrigatórias e formatação, completando o processo em aproximadamente 60 minutos por conjunto. Conclusões: Nossa método acelerou significativamente a criação de casos e melhorou a diversidade, embora priorizar a máxima diversidade tenha comprometido a representatividade em certa medida. Embora os prompts otimizados sejam facilmente reutilizáveis, o processo em si

					exige habilidades em informática que nem todos os educadores possuem. Para resolver isso, pretendemos compartilhar todos os pacientes criados como recursos educacionais abertos, capacitando os educadores a criar casos de forma independente.
37	Accuracy of <i>ChaGPT</i> -Generated Information on Head and Neck and Oromaxillofacial Surgery: A Multicenter Collaborative Analysis Precisão das Informações Geradas pelo <i>ChaGPT</i> em Cirurgia de Cabeça e Pescoço e Cirurgia Oro-Maxilo-Facial: Uma Análise Colaborativa Multicêntrica.	Vaira et al. (2023)	Investigar a precisão do Transformador Pré-treinado Generativo Baseado em Chat (<i>ChaGPT</i>) em responder perguntas e resolver cenários clínicos de cirurgia de cabeça e pescoço.	Vaira et al (2023), investigaram a precisão do <i>ChaGPT</i> em responder perguntas e resolver cenários clínicos de cirurgia de cabeça e pescoço. As pontuações médias para precisão e completude em questões abertas foram 6 (IQR: 5-6) e 3 (IQR: 2-3), respectivamente. O modelo obteve 87,2% de respostas totalmente ou quase totalmente corretas e cobriu todos os aspectos da questão em 73% dos casos. Para questões fechadas, a IA acertou em 84,7%, com 11 respostas incorretas. Nos cenários clínicos, o <i>ChaGPT</i> forneceu diagnósticos precisos em 81,7% dos casos, enquanto os procedimentos propostos foram considerados completos em 56,7%. A qualidade das referências bibliográficas foi avaliada como ruim em geral, com 46,4% das fontes ausentes. Esses resultados destacam as capacidades e limitações do <i>ChaGPT</i> na abordagem de questões clínicas em cirurgia de cabeça e pescoço.	Resumo: Objetivo: Investigar a precisão do Chat-Based Generative Pre-trained Transformer (<i>ChaGPT</i>) em responder perguntas e resolver cenários clínicos de cirurgia de cabeça e pescoço. Desenho do Estudo: Estudo observacional e avaliativo. Ambiente: Dezoito cirurgiões de quatorze unidades italianas de cirurgia de cabeça e pescoço. Métodos: Um total de 144 perguntas clínicas abrangendo diferentes subespecialidades da cirurgia de cabeça e pescoço e 15 cenários clínicos abrangentes foram desenvolvidos. As perguntas e cenários foram inseridos no <i>ChaGPT4</i> , e as respostas resultantes foram avaliadas pelos pesquisadores usando escalas Likert de precisão (faixa 1-6), completude (faixa 1-3) e qualidade das referências. Resultados: A pontuação mediana geral das perguntas abertas foi de 6 (intervalo interquartil [IQR]: 5-6) para precisão e 3 (IQR: 2-3) para completude. No geral, os revisores classificaram a resposta como totalmente ou quase totalmente correta em 87,2% dos casos e como abrangente e cobrindo todos os aspectos da pergunta em 73% dos casos. O modelo de inteligência artificial (IA) obteve uma resposta correta em 84,7% das perguntas de resposta fechada (11 respostas incorretas). Quanto aos cenários clínicos, o <i>ChaGPT</i> forneceu um diagnóstico totalmente ou quase totalmente correto em 81,7% dos casos. O procedimento diagnóstico ou terapêutico proposto foi considerado completo em 56,7% dos casos. A qualidade geral das referências bibliográficas foi baixa, e as fontes eram inexistentes em 46,4% dos casos. Conclusão: Os resultados geralmente demonstram um bom nível de precisão nas respostas da IA. A capacidade do AI em resolver cenários clínicos complexos é promissora, mas ainda não é considerada um suporte confiável para o processo de tomada de decisão de especialistas em cirurgia de cabeça e pescoço.
38	Brain versus bot: Distinguishing letters of recommendation authored by humans compared with artificial intelligence Cérebro versus robô: Distinguindo cartas de recomendação escritas por humanos em	Preiksaitis et al. (2023)	Avaliar a viabilidade do uso de ferramentas de inteligência artificial generativa, como o <i>ChatGPT</i> , na redação de cartas de recomendação (LORs) em medicina acadêmica.	Preiksaitis et al (2023), investigaram a viabilidade do uso de ferramentas de inteligência artificial, como o <i>ChaGPT</i> , na redação de cartas de recomendação (LORs) na medicina acadêmica. O estudo envolveu 32 participantes avaliando LORs de autoria humana e de IA. A precisão média em distinguir entre as duas foi de 59,4%, sem impacto significativo da certeza do revisor ou do tempo de deliberação. LORs suspeitas de serem de autoria humana foram avaliadas mais	Resumo: Objetivos: As cartas de recomendação (LORs) são essenciais na medicina acadêmica, influenciando decisões importantes sobre avanços na carreira, porém exigem tempo e trabalho significativos para serem preparadas. O uso de ferramentas de inteligência artificial (IA) generativa, como <i>ChaGPT</i> , está ganhando popularidade para diversas tarefas de escrita acadêmica e oferece uma solução inovadora para aliviar o ônus da redação de cartas. Ainda não foi determinado se <i>ChaGPT</i> poderia auxiliar na elaboração de LORs, especialmente em contextos de alta importância como a promoção de professores. Para determinar a viabilidade desse processo e se há diferença significativa entre cartas escritas por IA e humanos, realizamos um estudo para verificar se os médicos

	comparação com inteligência artificial.			<p>favoravelmente em qualidade e persuasão. Além disso, observou-se uma diferença na linguagem de gênero, com LORs humanas tendo mais palavras associadas a mulheres, enquanto LORs de IA apresentavam mais palavras associadas a homens. Estes resultados contribuem para a compreensão da aplicação da IA na redação acadêmica, considerando aspectos de percepção e linguagem.</p>	<p>acadêmicos conseguem distinguir entre elas. Métodos: Foi conduzido um estudo quasi-experimental com um desenho de único cego. Médicos acadêmicos com experiência em revisão de LORs foram apresentados com LORs para promoção a professor associado, escritas por humanos ou por IA. Os participantes revisaram as LORs e identificaram a autoria. Análise estatística foi realizada para determinar a precisão na distinção entre LORs escritas por humanos e por IA. Além disso, a qualidade percebida e a persuasividade das LORs foram comparadas com base na autoria suspeita e real. Resultados: Um total de 32 participantes completaram a revisão das cartas. A precisão média na distinção entre LORs escritas por humanos versus por IA foi de 59,4%. A certeza do revisor e o tempo gasto na deliberação não impactaram significativamente a precisão. As LORs suspeitas de serem escritas por humanos foram avaliadas mais favoravelmente em termos de qualidade e persuasividade. Foi observada uma diferença na linguagem com viés de gênero em nossas cartas: as cartas escritas por humanos continham significativamente mais palavras associadas a mulheres, enquanto a maioria das cartas escritas por IA tendeu a usar mais palavras associadas a homens. Conclusões: Os participantes não conseguiram diferenciar de forma confiável entre LORs escritas por humanos e por IA para promoção. A IA pode ser capaz de gerar LORs e aliviar o ônus da redação de cartas para acadêmicos. Novas estratégias, políticas e diretrizes são necessárias para equilibrar os benefícios da IA enquanto se preserva a integridade e a equidade nas decisões de promoção acadêmica.</p>
39	Scientometric analysis on the use of <i>ChaGPT</i> , artificial intelligence, or intelligent conversational agent in the role of medical training Análise Cientométrica sobre o Uso do <i>ChaGPT</i> , Inteligência Artificial ou Agente Conversacional Inteligente no Papel do Treinamento Médico.	Mayta-Tovalino et al. (2023)	Examinar a utilização de agentes conversacionais inteligentes, <i>ChaGPT</i> e bots de inteligência artificial na educação médica.	<p>Segundo Mayta-Tovalino et al. (2023), o estudo bibliométrico exploratório revela que a utilização de agentes conversacionais inteligentes, como o <i>ChaGPT</i> e bots de IA, tem sido amplamente investigada na área da educação médica. Timothy W. Bickmore é destacado como o autor com o maior número de publicações e citações recebidas, indicando sua influência significativa nesse campo. A subcategoria com maior produção acadêmica é a Informática da Saúde, enquanto Geriatria e Gerontologia apresentam baixa representatividade. Colaborações dentro do mesmo país são comuns, com instituições como Swiss Federal Institute of Technology Zurich e Imperial College London se destacando por seu impacto nas respectivas áreas. No entanto, apesar do Brasil ter um número considerável de publicações, o país</p>	<p>Nos tempos recentes, houve um aumento perceptível no uso de inteligência artificial, incluindo <i>ChaGPT</i> e outros tipos, no campo da educação em ciências da saúde. Neste contexto, um estudo bibliométrico exploratório foi realizado para examinar a utilização de agentes conversacionais inteligentes, <i>ChaGPT</i>, e bots de inteligência artificial na educação médica. Métodos: Foi empregada uma análise bibliométrica retrospectiva, observacional e transversal para avaliar as publicações científicas listadas no Scopus. O estudo foi conduzido em 11 de março de 2023, com a busca de informações no Scopus. Um total de 220 documentos relevantes foram identificados na base de dados Scopus no período de 2017 a 2022. Foi utilizado o software SciVal da Elsevier. Posteriormente, tabelas e gráficos estatísticos foram preparados para apresentação no software Bibliometrix. Resultados: Entre os autores, Timothy W. Bickmore, dos Estados Unidos, possui o maior número de publicações (10) e citações recebidas (172), com um índice h de 45, sugerindo uma influência significativa no campo de estudo. A subcategoria com o maior número de publicações acadêmicas é Informática em Saúde, com 133 publicações, enquanto Geriatria e Gerontologia tem apenas 3. A maioria das publicações analisadas (44,2%) originou-se de colaborações dentro do mesmo país. Notavelmente, o Instituto Federal Suíço de Tecnologia de Zurique e o Imperial College London destacaram-se com 12 publicações cada, recebendo mais</p>

				possui um impacto de citações ponderado por campo relativamente baixo e recebeu o menor número de citações. Uma análise de agrupamento mostrou que os termos "agentes conversacionais", "Chatbots", "agente conversacional" e "Chatbot" estão estreitamente relacionados, indicando a interconexão desses conceitos.	de 200 citações, indicando seu impacto significativo em seus respectivos campos. Apesar de ter o maior número de publicações acadêmicas (15), o Brasil teve um impacto de citação ponderado pelo campo relativamente baixo (0,64) e recebeu o menor número de citações (81). Foi realizada uma análise de agrupamento em uma amostra de 10 conceitos usando 2 dimensões. Os resultados indicaram que todos os termos faziam parte do mesmo cluster. Notavelmente, os termos 'agentes conversacionais', 'Chatbots', 'agente conversacional' e 'Chatbot' estavam intimamente relacionados. Conclusões: Verificou-se que o americano Timothy W. Bickmore liderou os 10 principais pesquisadores, e que Informática em Saúde foi a área temática mais predominante. No entanto, Brasil e Alemanha foram os países líderes em termos de produção de pesquisa, principalmente publicada em periódicos de alto impacto (Q1).
40	Harnessing the benefits of <i>ChaGPT</i> for radiography education: A discussion paper Explorando os Benefícios do <i>ChaGPT</i> para a Educação em Radiografia: Um Artigo de Discussão.	Amedu e Ohene-Botwe (2024)	Explora a possibilidade de integração do <i>ChaGPT</i> no ensino de radiografia, um modelo de linguagem de IA conversacional de última geração, no ensino de radiografia.	Segundo Amedu e Ohene-Botwe (2024), a integração do <i>ChaGPT</i> no ensino de radiografia pode trazer oportunidades de aprendizagem interativas, aprendizagem individualizada, plataformas de revisão e apoio aos educadores na criação de cenários, desenvolvimento de avaliações, colaboração em grupo e atividades profissionais e de pesquisa.	A educação em radiografia é fundamental para treinar radiologistas habilidosos em imagens diagnósticas e aplicações terapêuticas. Com os avanços tecnológicos, há um interesse crescente em ferramentas educacionais inovadoras para aprimorar os métodos de ensino tradicionais. Este artigo de discussão explora a possibilidade de integração do <i>ChaGPT</i> , um modelo de linguagem de inteligência artificial conversacional de ponta, na educação em radiografia. Principais descobertas: Relatamos que o <i>ChaGPT</i> oferece oportunidades de aprendizado interativo que podem facilitar a aprendizagem. Ele também proporciona aprendizado autodirigido, plataformas de revisão e apoia educadores na criação de cenários, desenvolvimento de avaliações, colaboração em grupo e atividades profissionais e de pesquisa. Apesar desses benefícios, é importante considerar cuidadosamente questões relacionadas à integridade acadêmica e privacidade, juntamente com as oportunidades e desafios apresentados por esta nova tecnologia na educação em radiografia. Conclusão: Este artigo destaca algumas das perspectivas e limitações das potenciais aplicações do <i>ChaGPT</i> na educação em radiografia, enfatizando os benefícios tanto para discentes quanto para educadores. No entanto, sua implementação deve ser pensada de forma cuidadosa e ética, levando em consideração seus pontos fortes e limitações. Implicações para a prática: A integração do <i>ChaGPT</i> na educação em radiografia tem o potencial de melhorar a educação em radiografia ao aumentar a literacia digital e os resultados dos graduandos, ao mesmo tempo em que simplifica o processo de preparação para os educadores. No entanto, a implementação ética é vital para resultados ótimos.
41	<i>ChaGPT</i> 3.5 fails to write appropriate multiple choice practice exam questions <i>ChaGPT</i> 3.5 falha ao escrever questões	Ngo et al. (2024)	Investigar se o <i>ChaGPT</i> poderia ser usado para criar questões de múltipla escolha juntas com explicações de escolhas corretas e incorretas para aulas de imunologia de	De acordo com Ngo et al (2024), foi investigado se o <i>ChaGPT</i> poderia ser aplicado na criação de questões de múltipla escolha acompanhadas de explicações para escolhas corretas e incorretas, a fim de auxiliar nas aulas de imunologia tanto em nível de graduação quanto de faculdade de medicina. Os resultados indicaram que, no geral, o <i>ChaGPT</i> apresentou dificuldades em produzir	A inteligência artificial (IA) pode ter um impacto profundo no ensino tradicional em ambientes acadêmicos. Várias preocupações têm sido levantadas, especialmente relacionadas ao uso do <i>ChaGPT</i> para criar ensaios de novo. No entanto, programas de IA como o <i>ChaGPT</i> podem aprimorar técnicas de ensino. Neste artigo, utilizamos o <i>ChaGPT</i> 3.5 para criar 60 questões de múltipla escolha. Textos escritos pelos autores foram carregados e o <i>ChaGPT</i> foi solicitado a criar questões de múltipla escolha com uma explicação para a resposta correta e explicações para as respostas incorretas. Infelizmente, o <i>ChaGPT</i> gerou apenas

	apropriadas para exames práticos de múltipla escolha.		graduação e faculdade de medicina.	perguntas e respostas adequadas para um curso de imunologia ministrado por professores especializados no Departamento de Patologia e Medicina Laboratorial. Contudo, evidenciou-se que o <i>ChaGPT</i> conseguiu criar corretamente as hastes das perguntas, identificou a resposta correta e forneceu explicações sucintas, porém detalhadas, para a resposta correta. Além disso, explicações adequadas também foram oferecidas para cada uma das respostas incorretas.	questões e respostas corretas com explicações em 32% das questões (19 de 60). Em muitos casos, o <i>ChaGPT</i> não conseguiu fornecer uma explicação para as respostas incorretas. Um adicional de 25% das questões teve respostas que estavam erradas ou enganosas. Uma nota de 32% seria considerada insuficiente na maioria dos cursos. Apesar desses problemas, instrutores ainda podem achar o <i>ChaGPT</i> útil para criar exames práticos com explicações—com a ressalva de que pode ser necessária uma edição extensiva. Palavras-chave: Imunologia, Inteligência artificial, Educação, Patologia.
42	Role of <i>Chatbots</i> on gastroenterology: Let's chat about the future O Papel dos <i>Chatbots</i> em Gastroenterologia: Vamos Conversar sobre o Futuro.	Li (2023)	Sintetizar relatórios da literatura atual sobre as funções dos <i>Chatbots</i> e avaliar suas aplicações potenciais em diversas áreas, como diagnóstico clínico e planos de tratamento, triagem de pacientes, pesquisa científica e redação acadêmica.	De acordo com Li (2023), os <i>Chatbots</i> têm o potencial de melhorar a eficiência e os resultados dos pacientes em diversas áreas, como diagnóstico clínico e planos de tratamento, triagem de pacientes, pesquisa científica e redação acadêmica. Além disso, eles podem permitir que equipes médicas e discentes tenham acesso rápido a informações e recursos médicos relevantes, auxiliando no seu aprendizado e desenvolvimento, o que é essencial para o crescimento e avanço da gastroenterologia.	Os avanços recentes nos modelos de <i>Chatbot</i> , utilizando arquitetura de transformer e treinamento extensivo com dados, estão prontos para revolucionar a saúde, incluindo a gastroenterologia. Este artigo sintetiza a literatura atual sobre os papéis dos <i>Chatbots</i> e avalia suas potenciais aplicações no diagnóstico clínico, planejamento de tratamento, triagem de pacientes, pesquisa científica e redação acadêmica dentro da gastroenterologia. Argumentamos que os <i>Chatbots</i> têm a capacidade de melhorar a eficiência e os resultados dos pacientes nessas áreas. No entanto, desafios significativos como preocupações com privacidade e considerações emocionais devem ser abordados. No geral, embora promissores, a integração de <i>Chatbots</i> na gastroenterologia exige mais pesquisas para entender plenamente suas capacidades e implicações.
43	<i>ChaGPT</i> in medicine: an overview of its applications, advantages, limitations, future prospects, and ethical considerations <i>ChaGPT</i> na medicina: uma visão geral de suas aplicações, vantagens, limitações, perspectivas futuras e considerações éticas.	Dave, Athuluri e Sing (2023)	Analizar as vantagens, limitações, considerações éticas, perspectivas futuras e aplicações práticas do <i>ChaGPT</i> e da inteligência artificial (IA) nos domínios da saúde e da medicina.	De acordo com a pesquisa realizada por Dave, Athuluri e Sing (2023), é possível afirmar que o <i>ChaGPT</i> representa um modelo avançado de linguagem que utiliza técnicas de aprendizagem profunda para gerar respostas similares às humanas em situações de linguagem natural. Nesse contexto, as áreas médica e de saúde mostram-se como potenciais domínios de aplicação para o <i>ChaGPT</i> . Contudo, é importante destacar que tais aplicações enfrentam limitações e considerações éticas, os quais são abordados detalhadamente juntamente com as perspectivas futuras no campo da saúde e medicina.	Este artigo apresenta uma análise das vantagens, limitações, considerações éticas, perspectivas futuras e aplicações práticas de <i>ChaGPT</i> e inteligência artificial (IA) nos domínios da saúde e medicina. <i>ChaGPT</i> é um modelo avançado de linguagem que utiliza técnicas de aprendizado profundo para produzir respostas semelhantes às humanas a entradas de linguagem natural. Ele faz parte da família de modelos GPT (generative pre-training transformer) desenvolvidos pela OpenAI e atualmente é um dos maiores modelos de linguagem disponíveis publicamente. <i>ChaGPT</i> é capaz de capturar nuances e complexidades da linguagem humana, permitindo gerar respostas apropriadas e contextualmente relevantes em uma ampla gama de estímulos. As aplicações potenciais de <i>ChaGPT</i> no campo médico vão desde a identificação de tópicos de pesquisa até o auxílio de profissionais no diagnóstico clínico e laboratorial. Além disso, pode ser usado para ajudar discentes de medicina, médicos, enfermeiros e todos os membros da fraternidade da saúde a se manterem atualizados sobre novidades e desenvolvimentos em suas respectivas áreas. O desenvolvimento de assistentes virtuais para auxiliar pacientes na gestão de sua saúde é outra aplicação importante de <i>ChaGPT</i> na medicina. Apesar de suas potenciais aplicações, o uso de <i>ChaGPT</i> e outras ferramentas de IA na escrita médica também apresenta preocupações éticas e legais. Estas incluem possíveis violações de leis de direitos autorais, complicações

					médico-legais e a necessidade de transparência no conteúdo gerado por IA. Em conclusão, ChaGPT tem diversas aplicações potenciais nos campos médico e da saúde. No entanto, essas aplicações vêm acompanhadas de várias limitações e considerações éticas, que são apresentadas detalhadamente juntamente com as perspectivas futuras na medicina e saúde.
44	Assessing the Capability of <i>ChaGPT</i> in Answering First- and Second-Order Knowledge Questions on Microbiology as per CompetencyBased Medical Education Curriculum Avaliando a Capacidade do <i>ChaGPT</i> em Responder a Perguntas de Conhecimento de Primeira e Segunda Ordem sobre Microbiologia de Acordo com o Currículo de Educação Médica Baseada em Competências.	Das et al. (2023)	Analizar a capacidade do <i>ChaGPT</i> em responder questões de primeira e segunda ordem sobre o tema microbiologia.	De acordo com Das et al. (2023), foi analisada a capacidade do <i>ChaGPT</i> em responder questões de primeira e segunda ordem sobre microbiologia. Os resultados mostraram que a pontuação média das respostas foi de $4,04 \pm 0,37$, indicando um desempenho consistente. No entanto, as respostas às questões de conhecimento de segunda ordem tiveram uma pontuação ligeiramente menor, sendo de $3,99 \pm 0,43$. Isso sugere uma certa inconsistência no desempenho do <i>ChaGPT</i> em diferentes tópicos da microbiologia.	Contexto e objetivo: ChaGPT é um modelo de linguagem de inteligência artificial (IA) treinado para processar e responder a perguntas em uma ampla gama de tópicos. Ele também é capaz de resolver problemas em tópicos educacionais médicos. No entanto, a capacidade do ChaGPT de responder com precisão a perguntas de conhecimento de primeiro e segundo nível no campo da microbiologia ainda não foi explorada até o momento. Portanto, neste estudo, nosso objetivo foi analisar a capacidade do ChaGPT em responder a perguntas de primeiro e segundo nível sobre o tema da microbiologia. Materiais e métodos: Com base no currículo de educação médica baseada em competências (CBME) do tema da microbiologia, preparamos um conjunto de perguntas de conhecimento de primeiro e segundo nível. Para os oito módulos no currículo CBME de microbiologia, preparamos seis perguntas de conhecimento de primeiro e seis de segundo nível de acordo com o currículo CBME recomendado pela Comissão Médica Nacional, totalizando $(8 \times 12) = 96$ perguntas. As perguntas foram verificadas quanto à validade de conteúdo por três microbiologistas especialistas. Essas perguntas foram usadas para interagir com o ChaGPT por um único usuário e as respostas foram registradas para análise posterior. As respostas foram avaliadas por três microbiologistas em uma escala de pontuação de 0-5. A média das três pontuações foi considerada como a pontuação final para análise. Como os dados não apresentaram distribuição normal, foi utilizado um teste estatístico não paramétrico. As pontuações gerais foram testadas por um teste de mediana de amostra única com valores hipotéticos de 4 e 5. As pontuações das respostas a perguntas de conhecimento de primeiro e segundo nível foram comparadas pelo teste U de Mann-Whitney. As respostas por módulo foram testadas pelo teste de Kruskall-Wallis seguido pelo teste post hoc para comparações em pares. Resultados: A pontuação geral de 96 respostas foi de $4,04 \pm 0,37$ (mediana: 4,17, Q1-Q3: 3,88-4,33), com a média das respostas às perguntas de conhecimento de primeiro nível sendo de $4,07 \pm 0,32$ (mediana: 4,17, Q1-Q3: 4,00-4,33) e das respostas às perguntas de conhecimento de segundo nível sendo de $3,99 \pm 0,43$ (mediana: 4, Q1-Q3: 3,67-4,33) ($p=0,4$ no teste U de Mann-Whitney). A pontuação foi significativamente abaixo da pontuação de 5 (teste de mediana de amostra única $p<0,0001$) e similar a 4 (teste de mediana de amostra única $p=0,09$). No geral, houve uma variação nas medianas das pontuações obtidas em oito categorias de tópicos em microbiologia, indicando desempenho inconsistente em diferentes tópicos. Conclusão: Os resultados do estudo indicam que o ChaGPT é capaz de responder a perguntas de conhecimento de primeiro e segundo nível relacionadas ao tema da microbiologia. O modelo alcançou uma precisão de aproximadamente 80% e não houve diferença na capacidade do modelo de

					responder a perguntas de primeiro e segundo nível de conhecimento. Os achados deste estudo sugerem que o ChaGPT tem o potencial de ser uma ferramenta eficaz para a resposta automatizada de perguntas no campo da microbiologia. No entanto, melhorias contínuas no treinamento e desenvolvimento de modelos de linguagem são necessárias para aprimorar seu desempenho e torná-los adequados para uso acadêmico.
45	Are <i>ChaGPT</i> 's knowledge and interpretation ability comparable to those of medical students in Korea for taking a parasitology examination?: a descriptive study A capacidade de conhecimento e interpretação do <i>ChaGPT</i> é comparável à dos discentes de medicina na Coreia para realizar um exame de parasitologia? Um estudo descritivo.	Huh, (2023)	Comparar o conhecimento e a capacidade de interpretação do <i>ChaGPT</i> , um modelo de linguagem de inteligência artificial geral, com os de discentes de medicina na Coreia.	De acordo com o estudo realizado por Huh (2023), o desempenho inferior do <i>ChaGPT</i> em relação aos discentes de medicina na Coreia, durante um exame de parasitologia, não foi influenciado pelo nível de conhecimento dos itens, porém, houve uma relação significativa entre explicações aceitáveis e respostas corretas. Assim, conclui-se que o conhecimento e a capacidade de interpretação do <i>ChaGPT</i> para este exame ainda não são comparáveis aos dos discentes de medicina na Coreia.	Este estudo teve como objetivo comparar o conhecimento e a capacidade de interpretação do <i>ChaGPT</i> , um modelo de linguagem de inteligência artificial geral, com os de discentes de medicina na Coreia, administrando um exame de parasitologia tanto ao <i>ChaGPT</i> quanto aos discentes de medicina. Métodos: O exame consistiu em 79 itens e foi administrado ao <i>ChaGPT</i> em 1º de janeiro de 2023. Os resultados do exame foram analisados em termos do desempenho geral do <i>ChaGPT</i> , da taxa de respostas corretas por nível de conhecimento dos itens e da aceitabilidade das explicações fornecidas pelo <i>ChaGPT</i> para os itens. Resultados: O desempenho do <i>ChaGPT</i> foi inferior ao dos discentes de medicina, e a taxa de respostas corretas do <i>ChaGPT</i> não estava relacionada ao nível de conhecimento dos itens. No entanto, houve uma relação entre explicações aceitáveis e respostas corretas. Conclusão: Em conclusão, a capacidade de conhecimento e interpretação do <i>ChaGPT</i> neste exame de parasitologia ainda não é comparável à dos discentes de medicina na Coreia.
46	Practical Applications of <i>ChaGPT</i> in Undergraduate Medical Education Aplicações Práticas do <i>ChaGPT</i> na Educação Médica de Graduação.	Tsang, (2023)	Delinear como o <i>ChaGPT</i> pode ser aplicado para apoiar a educação médica de graduação durante os anos pré-clínicos e clínicos, e destacar possíveis preocupações quanto ao seu uso, o que exige a criação de políticas formais e treinamento por parte das escolas médicas.	Conforme destacado por Tsang (2023), a aplicação do <i>ChaGPT</i> na educação médica de graduação durante os anos pré-clínicos e clínicos pode ser uma forma de apoiar os discentes, porém é importante considerar as potenciais preocupações associadas a seu uso, o que demanda a implementação de políticas formais e treinamentos por parte das escolas médicas. Pesquisas anteriores têm demonstrado que os discentes de medicina se beneficiam do uso de ferramentas e aplicativos médicos móveis no local de atendimento, promovendo o aumento do conhecimento, a confiabilidade e a eficiência (Tsang, 2023, p.	RESUMO: <i>ChaGPT</i> é um <i>Chatbot</i> desenvolvido pela OpenAI que tem recebido atenção significativa por alcançar ou estar próximo do padrão de aprovação no Exame de Licenciamento Médico dos Estados Unidos (USMLE). Atualmente, pesquisadores e usuários estão explorando as amplas possíveis aplicações de <i>ChaGPT</i> na academia, negócios, programação e além. Tentamos delinejar como <i>ChaGPT</i> pode ser aplicado para apoiar a educação médica de graduação durante os anos pré-clínicos e clínicos, e destacar possíveis preocupações quanto ao seu uso, o que necessita a criação de políticas formais e treinamentos por parte das escolas médicas. PALAVRAS-CHAVE: Educação médica, UGME, <i>ChaGPT</i> , <i>Chatbot</i> , IA.

				X). Embora atualmente não exista um aplicativo móvel dedicado para o <i>ChaGPT</i> , o acesso através de navegadores da web em dispositivos móveis permite que os usuários utilizem a ferramenta de acordo com suas necessidades.	
47	Comparing <i>ChaGPT</i> and GPT-4 performance in USMLE soft skill assessments Comparando o Desempenho do <i>ChaGPT</i> e GPT-4 em Avaliações de Habilidades Sociais no USMLE.	Brin et al. (2023)	Avaliar <i>ChaGPT</i> e GPT-4 em questões USMLE envolvendo habilidades de comunicação, ética, empatia e profissionalismo.	De acordo com o estudo conduzido por Brin et al. (2023), o objetivo foi avaliar a eficácia do <i>ChaGPT</i> e GPT-4 em questões USMLE relacionadas a habilidades de comunicação, ética, empatia e profissionalismo. Para essa avaliação, foram utilizadas 80 questões de múltipla escolha, selecionadas para refletir os requisitos dos exames USMLE, além de 59 perguntas adicionais selecionadas do banco de perguntas AMBOSS. Essas perguntas foram escolhidas para incluir casos éticos semelhantes aos encontrados nos exemplos de questões do teste USMLE. O AMBOSS também fornece estatísticas de desempenho de usuários anteriores, permitindo uma análise comparativa do desempenho dos <i>LLMs</i> em relação aos discentes de medicina e médicos.	Métodos: Grandes modelos de linguagem. Neste estudo, avaliamos o desempenho dos modelos de IA <i>ChaGPT</i> (GPT-3.5-turbo) e GPT-4, desenvolvidos pela OpenAI (https://openai.com). Como modelos de linguagem, eles são projetados para processar e gerar texto que imita conversas humanas. Eles alcançam isso prevendo qual a próxima palavra (token) que deve vir após uma entrada, permitindo-lhes gerar respostas coerentes e contextualmente relevantes. <i>ChaGPT</i> , uma versão anterior, tem sido amplamente utilizada em uma variedade de aplicações devido à sua capacidade de gerar respostas coerentes e contextualmente relevantes. GPT-4, uma versão mais recente, é semelhante em design, mas muito maior em escala, o que significa que foi treinado com mais dados e pode entender e gerar texto com ainda maior precisão. Datasets de questões médicas: Para avaliar <i>ChaGPT</i> e GPT-4, utilizamos um conjunto de 80 questões de múltipla escolha projetadas para espelhar os requisitos dos exames USMLE. Este conjunto de questões foi compilado a partir de duas fontes respeitáveis. A primeira é um conjunto de questões de teste amostrais para Step1, Step2CK e Step3, lançado entre junho de 2022 e março de 2023, disponível no site oficial do USMLE (https://www.usmle.org/prepare-your-exam). Fizemos uma triagem de todas as questões de teste de exemplo e selecionamos 21 questões que não requerem conhecimento médico científico, mas exigem habilidades de comunicação e interpessoais, profissionalismo, questões legais e éticas, competência cultural, comportamento organizacional e liderança. A segunda fonte é a AMBOSS, um banco de questões amplamente reconhecido para médicos e discentes de medicina (https://www.amboss.com/us), da qual selecionamos mais 59 questões. As questões escolhidas incluem tipos de questões Step1, Step2CK e Step3, lidando com cenários éticos, comparáveis às questões escolhidas dos testes de exemplo do USMLE. A AMBOSS fornece estatísticas de desempenho de seus usuários anteriores, permitindo uma análise comparativa do desempenho dos <i>LLMs</i> em relação ao de discentes de medicina e médicos. Tanto <i>ChaGPT</i> quanto GPT-4 foram encarregados de responder a todas as 80 questões incluídas neste estudo, e suas respostas e desempenho foram subsequentemente comparados. Formatação de prompts: Em conformidade com as práticas metodológicas padrão para avaliar modelos de linguagem de IA, formatamos uma estrutura de prompt que incluía o texto da pergunta seguido das respostas de múltipla escolha separadas por uma nova linha. Após a resposta do modelo, perguntamos "Você tem certeza?", o que

				nos permitiu avaliar a consistência e estabilidade do modelo e provocar uma potencial reavaliação de sua resposta inicial. Se um modelo altera sua resposta, isso pode indicar que ele possui algum nível de "incerteza" sobre sua resposta inicial. Acompanhar com que frequência e em que circunstâncias o modelo altera sua resposta pode fornecer insights valiosos sobre a capacidade do modelo de auto-revisão, que é um aspecto importante dos sistemas de aprendizado e tomada de decisão. Exemplos de prompts são mostrados nos itens supplementares 1–4. Resultados: Neste estudo, um total de 80 questões de habilidades sociais do USMLE foram incluídas, e os diferentes assuntos são detalhados na Tabela 1 e na Figura 1. A precisão do ChaGPT para questões de teste amostrais do USMLE e da AMBOSS foi de 66,6% e 61%, respectivamente, com uma precisão geral de 62,5%. O GPT-4 demonstrou um desempenho superior, com uma precisão de 100% e 86,4% para questões de teste amostrais do USMLE e da AMBOSS, respectivamente, e uma precisão geral de 90%. Os resultados também são mostrados na Tabela 1 e na Figura 1.	
48	<p><i>ChaGPT</i> Knowledge Evaluation in Basic and Clinical Medical Sciences: Multiple Choice Question Examination-Based Performance</p> <p>Avaliação do Conhecimento do <i>ChaGPT</i> em Ciências Médicas Básicas e Clínicas: Desempenho Baseado em Exame de Questões de Múltipla Escolha.</p>	<p>Meo et al. (2023)</p>	<p>Investigar o nível de conhecimento do <i>ChaGPT</i> na educação médica, tanto nas ciências médicas básicas quanto nas clínicas, no desempenho baseado em exames de questões de múltipla escolha (MCQs) e seu impacto no sistema de exames médicos.</p>	<p>De acordo com Meo et al. (2023), o estudo demonstrou que o <i>ChaGPT</i> apresentou um nível satisfatório de conhecimento na área da educação médica, tanto nas ciências médicas básicas quanto nas clínicas. O desempenho baseado em exames de questões de múltipla escolha foi avaliado, revelando que o <i>ChaGPT</i> foi capaz de fornecer respostas em MCQs individuais nessas áreas. O estudo utilizou uma questão de teste padronizada abrangente e de alto risco, tornando-o um excelente substrato para avaliar o <i>ChaGPT</i>. As pontuações obtidas nas MCQs aplicadas indicaram que o <i>ChaGPT</i> apresentou um desempenho mais preciso em questões de ciências médicas básicas em comparação com questões de ciências clínicas.</p>	<p>Resumo: O <i>Chatbot</i> Generative Pre-Trained Transformer (<i>ChaGPT</i>) tem recebido grande atenção do público, acadêmicos e comunidades científicas. Ele responde com respostas apropriadas e articuladas, bem como explicações em várias disciplinas. Para o uso do <i>ChaGPT</i> na educação, pesquisa e saúde, diferentes perspectivas existem com algum nível de ambiguidade em relação à sua aceitabilidade e usos ideais. No entanto, a literatura carece profundamente de estabelecer uma ligação para avaliar os níveis intelectuais do <i>ChaGPT</i> nas ciências médicas. Portanto, o presente estudo teve como objetivo investigar o nível de conhecimento do <i>ChaGPT</i> na educação médica, tanto nas ciências médicas básicas quanto clínicas, baseado no desempenho em exames de questões de múltipla escolha (MCQs) e seu impacto no sistema de exames médicos. Neste estudo, inicialmente, um banco de questões por disciplina foi estabelecido com uma variedade de questões de múltipla escolha (MCQs) de diversos livros didáticos médicos e pools de exames universitários. Os membros da equipe de pesquisa revisaram cuidadosamente o conteúdo das MCQs e garantiram que fossem relevantes para o conteúdo da disciplina. Cada questão era baseada em cenários com quatro subitens e tinha uma única resposta correta. Foram selecionadas aleatoriamente 100 MCQs em diversas disciplinas, incluindo ciências médicas básicas (50 MCQs) e ciências médicas clínicas (50 MCQs), a partir do banco de MCQs. As MCQs foram inseridas manualmente uma por uma, e uma nova sessão de <i>ChaGPT</i> foi iniciada para cada entrada para evitar viés de retenção de memória. Foi dado a tarefa ao <i>ChaGPT</i> de avaliar a resposta e o nível de conhecimento. A primeira resposta obtida foi considerada como a resposta final. Com base em uma chave de resposta pré-determinada, a pontuação foi feita em uma escala de 0 a 1, sendo zero representando resposta incorreta e um representando resposta correta. Os resultados revelaram que, das 100 MCQs em diversas disciplinas das ciências</p>

					médicas básicas e clínicas, ChaGPT tentou todas as MCQs e obteve 37/50 (74%) pontos nas ciências médicas básicas e 35/50 (70%) pontos nas ciências médicas clínicas, com uma pontuação geral de 72/100 (72%) em ambas as ciências médicas básicas e clínicas. Conclui-se que ChaGPT obteve uma pontuação satisfatória tanto nas disciplinas médicas básicas quanto clínicas e demonstrou um grau de entendimento e explicação. Os achados deste estudo sugerem que ChaGPT pode ser capaz de auxiliar discentes e professores de medicina em ambientes de educação médica, uma vez que possui potencial como inovação no contexto das ciências e educação médicas.
49	<i>ChaGPT in Medical Education and Research: A Boon or a Bane?</i> <i>ChaGPT na Educação Médica e Pesquisa: Uma Bênção ou uma Maldição?</i>	Jeyaraman et al. (2023)	Enfoca-se o <i>ChaGPT</i> na educação e pesquisa médica é uma bênção ou uma maldição. <i>ChaGPT</i> pode ser uma bênção para a educação e pesquisa médica. Tem o potencial de fornecer acesso instantâneo a uma vasta quantidade de informações médicas e auxiliar na análise de dados complexos, o <i>ChaGPT</i> apresenta desafios e limitações, como a falta de precisão e confiabilidade das informações fornecidas.	Segundo Jeyaraman et al. (2023), a utilização do <i>ChaGPT</i> na educação e pesquisa médica pode ser considerada tanto uma bênção quanto uma maldição. Enquanto ele possui o potencial de fornecer acesso instantâneo a uma vasta quantidade de informações médicas e auxiliar na análise de dados complexos, o <i>ChaGPT</i> apresenta desafios e limitações, como a falta de precisão e confiabilidade das informações fornecidas.	Resumo: A literatura de ficção científica e os filmes não se concentram mais apenas na inteligência artificial. Ao contrário de todos os outros aspectos da vida, a educação médica e o cuidado clínico ao paciente têm progredido lentamente. Recentemente, uma grande quantidade de texto da internet foi utilizada para construir e treinar <i>Chatbots</i> , especialmente o <i>ChaGPT</i> . O modelo de linguagem <i>ChaGPT</i> , criado pela OpenAI, emergiu como um recurso útil para pesquisa e educação médica. Ele se mostrou uma ferramenta útil para pesquisadores, discentes e profissionais médicos devido à sua capacidade de produzir respostas humanas para questões médicas desafiadoras. No entanto, o uso do <i>ChaGPT</i> também apresenta significativas desvantagens. A possibilidade de disseminação de informações errôneas ou tendenciosas, que poderiam ter efeitos negativos no cuidado ao paciente, é uma das principais preocupações. Além disso, a superdependência da tecnologia na educação médica também pode levar a uma diminuição das habilidades de pensamento crítico e tomada de decisão clínica. No geral, o <i>ChaGPT</i> tem o potencial de ser uma bênção para a educação e pesquisa médicas, mas seu uso deve ser acompanhado de cautela e avaliação crítica.
50	Assessment of <i>ChaGPT</i> 's performance on neurology written board examination questions Avaliação do desempenho do <i>ChaGPT</i> em questões do exame escrito de neurologia.	Chen et al. (2023)	Quantificar a capacidade do <i>ChaGPT</i> de processar informações médicas altamente especializadas.	De acordo com Chen et al. (2023), o <i>ChaGPT</i> demonstrou uma capacidade promissora em processar informações médicas altamente especializadas. O estudo revelou que o sistema obteve uma taxa de acerto de 65,8% na primeira tentativa e 75,3% após três tentativas, ao responder corretamente a 335 e 383 perguntas, respectivamente. Além disso, observou-se que o desempenho do <i>ChaGPT</i> superou o 25º percentil na primeira tentativa e o 50º percentil após três tentativas, em comparação com as respostas dos discentes do BoardVitals. Os resultados também indicaram que determinados temas, como dor, epilepsia e convulsões, e genética, foram abordados com maior precisão pelo <i>ChaGPT</i> , enquanto estudos de imagem/diagnóstico, cuidados	Resumo: Contexto e objetivos: O <i>ChaGPT</i> tem mostrado promessa na área da saúde. Para avaliar a utilidade dessa nova ferramenta na educação em saúde, nós avaliamos o desempenho do <i>ChaGPT</i> na resposta a questões do exame de neurologia. Métodos: Questões de exames no estilo de neurologia foram acessadas do BoardVitals, um banco comercial de questões de neurologia. O <i>ChaGPT</i> recebeu o prompt completo da questão e múltiplas opções de resposta. Foram fornecidas três tentativas para que o <i>ChaGPT</i> selecionasse a resposta correta, sendo que tentativas subsequentes foram dadas para chegar à resposta correta. Um total de 560 questões (14 blocos de 40 questões) foram utilizadas, embora quaisquer questões baseadas em imagens tenham sido descartadas devido à incapacidade do <i>ChaGPT</i> de processar entradas visuais. As respostas da inteligência artificial (IA) foram então comparadas com os dados de usuários humanos fornecidos pelo banco de questões para medir seu desempenho. Resultados: De um total de 509 questões elegíveis distribuídas em 14 blocos de questões, o <i>ChaGPT</i> respondeu corretamente a 335 questões (65,8%) na primeira tentativa/iteração e a 383 questões (75,3%) ao longo de três tentativas/iterações,

			<p>intensivos e nervos cranianos apresentaram um desempenho inferior.</p> <p>pontuando aproximadamente nos percentis 26 e 50, respectivamente. As disciplinas com melhor desempenho foram dor (100%), epilepsia e convulsões (85%) e genética (82%), enquanto as disciplinas com pior desempenho foram estudos de imagem/diagnóstico (27%), cuidados críticos (41%) e nervos cranianos (48%). Discussão: Este estudo encontrou que o ChaGPT teve desempenho semelhante aos seus equivalentes humanos. A precisão da IA aumentou com múltiplas tentativas e o desempenho ficou dentro da faixa esperada para aprendizes de neurologia residentes. Este estudo demonstra o potencial do ChaGPT no processamento de informações médicas especializadas. Estudos futuros melhor definirão até que ponto a IA será capaz de se integrar à tomada de decisão médica.</p>
--	--	--	--