

Revista Brasileira de Odontologia Legal – RBOL

ISSN 2359-3466

<http://www.portalabol.com.br/rbol>



Odontologia legal

USO DE APLICATIVO DE INTELIGÊNCIA ARTIFICIAL COMO FONTE DE INFORMAÇÃO PARA RESOLUÇÃO DE PROVAS DE ODONTOLOGIA.

Use of an artificial intelligence application as a source of information for solving Dentistry exams.

Hellen Cristina de Paiva MACHADO¹, Fernanda Ferreira de Albuquerque JASSÉ², Diandra Costa ARANTES².

1. Graduanda em Odontologia, Universidade Federal do Pará, Brasil.

2. Docente da Faculdade de Odontologia, Universidade Federal do Pará, Brasil.

Informações sobre o manuscrito:

Recebido: 29 de janeiro de 2024

Aceito: 22 de abril de 2024

Autor(a) para contato:

Diandra Costa Arantes.

Universidade Federal do Pará, Faculdade de Odontologia. Rua Augusto Correa, 1. Bairro: Guamá. CEP: 66075-110. Belém-Pará-Brasil.

E-mail: diandracosta@ufpa.br.

RESUMO

Introdução: O ChatGPT® é uma ferramenta pública desenvolvida pela OpenAI que utiliza a tecnologia do modelo de linguagem GPT. Este chatbot é capaz de atender a variadas solicitações de texto. **Objetivo:** avaliar se o ChatGPT® é capaz de ser a única fonte de informação para resolução de provas de Odontologia. **Material e métodos:** consiste em um estudo transversal quantitativo analítico. Para a coleta de dados, foi elaborada uma prova fictícia constituída por questões do ENADE e de outros concursos públicos. Os participantes responderam a prova em dois momentos: T1, sem o ChatGPT® e, após 15 dias (T2), utilizando-o. A amostra foi de 30 discentes de graduação em Odontologia, divididos igualmente entre 3 grupos: 1º ao 4º semestre, 5º ao 6º semestre e 7º ao 10º semestre. Para análise de dados foram aplicadas análises estatísticas descritiva e inferencial, por meio do software SPSS, com os testes de Wilcoxon e de McNemar. **Resultados:** revelaram uma eficácia notável do ChatGPT® na resolução de questões discursivas, com 83,3% de taxa de acerto, enquanto os discentes deram mais respostas incorretas ou incompletas. Porém, foram observadas limitações da base de dados do ChatGPT® quanto às questões objetivas. É crucial ressaltar que, apesar de resultados promissores, a aplicação do Chat levanta questões éticas e pedagógicas. Assim, a introdução do ChatGPT® na educação preocupa quanto à validade e equidade nas avaliações, destacando a importância de encontrar equilíbrio entre a inovação tecnológica e a preservação da integridade acadêmica.

PALAVRAS-CHAVE

Odontologia legal; Educação em odontologia, Inteligência artificial.

INTRODUÇÃO

A Inteligência Artificial (IA) é um campo de estudo interdisciplinar que envolve o uso de modelos computacionais de algoritmos baseados em comportamento, com mínima intervenção

humana, e é direcionado por lógica, estatística, teoria da decisão, neurociência, cibernética e engenharia de computação. Esse campo abrange um conjunto sofisticado de algoritmos e técnicas, capacitando as máquinas a aprender e

executar tarefas que anteriormente eram exclusivas da inteligência humana.^{1-3.}

O uso da tecnologia de IA está criando um impacto global e está transformando rapidamente todas as esferas da vida moderna, incluindo indústria, mídia social, saúde, tecnologia espacial, bem como uma ampla gama de funções no âmbito governamental. Neste sentido, a OpenAI é um laboratório de pesquisa fundado em 2015, que fez progressos rápidos no desenvolvimento de tecnologias de IA e lançou vários produtos de aprendizado de máquina para o público em geral.^{4-6.}

O ChatGPT® é uma ferramenta pública desenvolvida pela OpenAI que utiliza a tecnologia do modelo de linguagem GPT (“Generative Pre-trained Transformer” - Transformador Pré-Treinado Generativo em tradução livre). Este *chatbot* altamente sofisticado é capaz de atender a uma ampla variedade de solicitações baseadas em texto, incluindo responder perguntas simples e realizar tarefas mais avançadas com precisão e rapidez. O ChatGPT® é capaz de alcançar esse resultado por meio da alavancagem de suas extensas bases de dados e de seu design eficiente, que permitem a compreensão e a interpretação precisa das solicitações dos usuários, seguidas pela geração de respostas adequadas em uma linguagem praticamente natural. Esta ferramenta começou a ganhar popularidade no Brasil em meados de junho de 2020, com o lançamento do modelo GPT-3. Desde então, vem sendo utilizada em diversos contextos, inclusive dentro do cenário

acadêmico, gerando conflitos e esbarrando em questões éticas no processo.⁶

O uso de IA na Odontologia avançou dramaticamente à medida que a ciência e a tecnologia progrediram. Essa tecnologia tem sido usada principalmente para diagnosticar condições dentárias patológicas, planejar tratamentos e prever o prognóstico de casos, com destaque para as áreas de radiologia e de ortodontia, principalmente para avaliação de métodos de diagnóstico digital.^{7,8} Entretanto, no contexto clínico, o uso desta ferramenta encontra alguns entraves como a incerteza de segurança da privacidade dos dados dos pacientes e o papel fundamental da avaliação humana para tomada de decisões clínicas.^{9,10.}

Na esfera educacional, a implementação da IA desponta como uma tendência promissora, carregando consigo um potencial significativo para transformar os métodos de ensino e de aprendizagem. Um dos benefícios mais notáveis dessa integração é a capacidade de personalizar o ensino, adaptando-se de maneira precisa às necessidades individuais dos discentes. Por meio do emprego de algoritmos de “*machine learning*” e de análise de dados, abre-se espaço para o desenvolvimento de sistemas educacionais inteligentes, capazes de ajustar conteúdos e metodologias conforme as características, o ritmo e o desempenho de cada estudante. Essa personalização não apenas otimiza a eficácia do processo de aprendizagem, mas também o torna mais envolvente, contribuindo para a realização dos objetivos educacionais propostos.¹¹

A aplicação da IA na área acadêmica tem sido um tópico bastante discutido na área da educação. A utilização de APIs ("Application Programming Interface" - Interface de Programação de Aplicativos, em tradução livre) de chat e GPT no ensino superior pode trazer inúmeros benefícios, como o aumento do engajamento dos discentes, da colaboração e da acessibilidade. A tecnologia também é notável por sua capacidade de realizar uma ampla gama de tarefas linguísticas, incluindo tradução, resposta a perguntas e geração de texto, com pouco ou nenhum treinamento específico para a tarefa. Entretanto, essas ferramentas também apresentam desafios e preocupações, sobretudo no que diz respeito à integridade acadêmica e à possibilidade de plágio.^{12,13.}

Sistemas de redação de artigos ou de trabalhos acadêmicos baseados em IA são projetados para gerar textos com base em um conjunto de parâmetros ou sugestões. Isso significa que os discentes podem potencialmente usar esses sistemas para trapacear em suas tarefas, enviando materiais que não são de sua autoria. Diante disso, torna-se urgente e necessário debater o uso desta ferramenta no ambiente acadêmico, analisando seus aspectos e seus impactos positivos e negativos, debatendo, também, propostas alternativas ao seu uso.¹⁴

Este trabalho teve por objetivo avaliar a possibilidade de uso do aplicativo de IA ChatGPT® no processo de resolução de uma prova de conhecimentos específicos da Odontologia por parte de

discentes da graduação da Universidade Federal do Pará (UFPA).

METODOLOGIA

Desenho de estudo e aspectos éticos

Esta pesquisa consiste em um estudo transversal analítico quantitativo, aprovado pelo Comitê de Ética em Pesquisa do Instituto de Ciências da Saúde da UFPA (parecer número 6.098.014 – CAAE: 68595823.0.0000.0018) e reportado segundo as recomendações da iniciativa *Strengthening the Reporting of Observational Studies in Epidemiology* (Declaração STROBE). Previamente à coleta de dados, todos os participantes assinaram o Termo de Consentimento Livre e Esclarecido.

Instrumento de coleta de dados

A coleta de dados foi realizada por meio de uma prova fictícia de Odontologia. Para a elaboração deste instrumento de coleta de dados, foram selecionadas, pelos pesquisadores, questões de Odontologia disponíveis em provas do Exame Nacional de Desempenho dos Estudantes (ENADE) e em provas de concursos públicos. Foram selecionadas questões objetivas, questões de verdadeiro (V) ou falso (F), questões de análise de itens e questões discursivas correspondentes a diferentes especialidades da Odontologia. As questões de concurso e as questões do ENADE selecionadas possuíam gabaritos e espelhos oficiais disponíveis nos sites oficiais das bancas promotoras dessas avaliações. Ao final, a prova fictícia elaborada pelos pesquisadores contava com duas questões de V ou F com cinco

itens a serem julgados em cada, totalizando 10 itens; duas questões de análise de itens; 10 questões objetivas e duas questões discursivas.

Amostra, critérios de inclusão e de exclusão

Foram convidados a participar da pesquisa, em amostragem por conveniência, 30 discentes de cursos de graduação em Odontologia da Faculdade de Odontologia da Universidade Federal do Pará (FO-UFPA) de todos os 10 semestres divididos em três grupos (n = 10): Grupo 1) ciclo básico (discentes do 1º ao 4º semestre); Grupo 2) ciclo laboratorial (discentes do 5º ao 6º semestre); e 3) ciclo clínico (discentes do 7º ao 10º semestre). Foram incluídos indivíduos de ambos os gêneros, com idade mínima de 18 anos de idade e com matrícula ativa na FO-UFPA. Foram excluídos da pesquisa discentes que estivessem afastados das atividades acadêmicas, discentes que não possuíssem familiaridade com equipamentos eletrônicos ou com internet e discentes que apresentassem deficiências intelectuais.

Estudo piloto

O tamanho amostral foi determinado por estudo piloto realizado com três discentes. Nesta etapa, foi solicitado aos voluntários participantes que respondessem à prova uma única vez utilizando o ChatGPT® como fonte de informação e que escrevessem à mão os comandos utilizados em cada questão. Observou-se que, desta maneira, a metodologia utilizada acabou dificultando o

processo e aumentando o tempo de resolução consideravelmente. Então, posteriormente, a metodologia foi alterada para a coleta e armazenamento dos comandos utilizados de forma virtual. O instrumento de coleta de dados também foi alterado, excluindo-se a área destinada à transcrição dos comandos.

Coleta de dados

A coleta de dados da pesquisa foi dividida em dois momentos. No primeiro momento (T1), os discentes foram submetidos ao instrumento de coleta de dados para responderem às questões simulando uma prova convencional, ou seja, sem nenhum tipo de auxílio de ferramenta. Após 15 dias (T2), os discentes foram novamente submetidos ao mesmo instrumento de coleta de dados para responderem às questões utilizando somente o aplicativo ChatGPT®.

Análise estatística

Os dados obtidos foram analisados por meio de análise estatística descritiva e de análise estatística inferencial, por meio do software SPSS (SPSS Statistics, Version 23.0; IBM Corp, Armonk, NY), considerando $\alpha = 0,05$. A hipótese nula da pesquisa é que não há diferença no número de acertos das questões com ou sem o uso do ChatGPT®.

Inicialmente, foram mensuradas as médias de acerto de cada formato de questão e foram mensuradas as frequências absoluta e relativa das variáveis de caracterização do perfil dos discentes, conforme os grupos. A fim de analisar a normalidade das variáveis

quantitativas, foi aplicado teste de Shapiro-Wilk. A fim de avaliar diferenças, entre os tempos T1 (sem uso do ChatGPT®) e T2 (com uso do ChatGPT®), com relação ao número de acertos de questões de V ou F, de questões de análise de itens e de questões objetivas, foi aplicado Teste de Wilcoxon. Para avaliar diferenças de respostas de questões discursivas nos dois tempos (T1 e T2), as variáveis foram transformadas em dicotômicas (opções de resposta correta ou incorreta, em que as incompletas foram consideradas incorretas) e a análise foi executada por meio do Teste de McNemar.

RESULTADOS

Participaram da pesquisa 30 discentes divididos igualmente entre os grupos (n = 10). Os dados detalhados de caracterização do perfil dos discentes estão descritos na Tabela 1. Quase todos os participantes (93,30%) afirmaram ter muita familiaridade com a internet. Entretanto, somente três dos 30 discentes afirmaram ter muita familiaridade com o uso de IA. Apesar disso, quase metade dos participantes afirmou fazer uso de outras ferramentas que utilizam a tecnologia de IA incorporada, como o *Google Assistant* e a Siri, presentes em aparelhos celulares Android e da empresa Apple, respectivamente.

O grupo amostral foi composto em 60% pelo gênero feminino e em 40% pelo gênero masculino. A maioria dos discentes tinha de 18 a 24 anos. Todos os participantes do Grupo 1 afirmaram já ter utilizado o ChatGPT® para atividades acadêmicas, enquanto somente três

participantes do Grupo 2 e quatro participantes do Grupo 3 afirmaram o mesmo. Dos 30 participantes, 28 afirmaram participar de atividades acadêmicas complementares: o estágio foi o mais prevalente (63,33%), seguido pela pesquisa e pela monitoria (43,33% cada). Outras atividades acadêmicas foram relatadas, sendo as atividades de extensão e de participação em ligas acadêmicas as mais citadas.

A Tabela 2 aponta a média de acertos obtidos pelos participantes sem o uso do ChatGPT® (T1) e com o uso do aplicativo (T2). Os dados obtidos mostraram que houve diferença estatisticamente significativa do número de acertos das questões de V ou F e das duas questões discursivas entre os dois tempos da pesquisa. Já a Tabela 3 expõe a média de acertos por grupo, mostrando que houve uma queda significativa no Grupo 2 em relação às questões de V ou F entre T1 e T2: no T1 o grupo apresentou média sete de acertos e no T2 caiu para quatro.

A Tabela 4 apresenta os números totais de acertos, de erros e de respostas incompletas em relação às questões discursivas com (T2) e sem (T1) a utilização do ChatGPT®. Para a questão discursiva 1, 63,2% os discentes responderam corretamente sem a utilização do ChatGPT®. Utilizando a ferramenta, este percentual foi para 93,3%. Em relação à questão discursiva 2, a taxa de acerto ficou em 33,3% sem a utilização do ChatGPT®, enquanto com o uso da IA ficou em 76,6%. Ainda em relação à questão discursiva 2, 53,2% dos discentes responderam de forma incompleta em T1,

enquanto esse percentual ficou em 23,2% em T2.

Tabela 1 - Caracterização do perfil dos participantes da pesquisa.

	DISCENTES					
	Grupo 1 1º ao 4º semestre		Grupo 2 5º ao 6º semestre		Grupo 3 7º ao 10º semestre	
	n	%	n	%	n	%
Gênero						
Feminino	4	13,3	6	20,0	8	26,7
Masculino	6	20,0	4	13,3	2	6,7
Idade						
18 a 24 anos	10	33,3	8	26,7	8	26,7
Mais de 25 anos	-	-	2	6,7	2	6,7
Familiaridade com internet						
Muito familiar	10	33,3	9	30,0	9	30,0
Razoavelmente familiar	-	-	1	3,3	1	3,3
Familiaridade com uso de IA						
Muito familiar	2	6,7	1	3,3	-	-
Razoavelmente familiar	8	26,7	2	6,7	6	20,0
Pouco familiar	-	-	4	13,3	2	6,7
Nada familiar	-	-	3	10,0	2	6,7
Uso do ChatGPT® em atividades acadêmicas						
Sim	10	33,3	3	10,0	4	13,3
Não	-	-	7	23,3	6	20,0
Uso de outras ferramentas com IA						
Sim	7	23,3	3	10,0	4	13,3
Não	3	10,0	7	23,3	6	20,0
Participação em atividades acadêmicas						
Sim	9	30,0	9	30,0	10	33,3
Não	1	3,3	1	3,3	-	-
Faz estágio						
Sim	9	30,0	5	16,7	5	16,7
Não	1	3,3	5	16,7	5	16,7
Faz monitoria						

Sim	3	10,0	5	16,7	5	16,7
Não	7	23,3	5	16,7	5	16,7
Faz pesquisa						
Sim	3	10,0	3	10,0	7	23,3
Não	7	23,3	7	23,3	3	10,0
Participação em outras atividades acadêmicas						
Sim	3	10,0	5	16,7	1	3,3
Não	7	23,3	5	16,7	9	30,0

n: número absoluto; %: frequência relativa; -: valor numérico igual a zero.

Tabela 2 - Média de acertos por tipo de questão sem a utilização do ChatGPT® (T1) e utilizando o ChatGPT® (T2).

Tipo de questões	Média de acertos T1	Média de acertos T2	Valor de p
Questões de V ou F	5,6	4,3	0,032*
Questões de itens	1,0	1,0	0,311*
Questões de objetivas	5,0	5,0	0,738*
Questão discursiva 1	6,5	9,5	0,022#
Questão discursiva 2	3,5	7,6	0,004#

*Teste de Wilcoxon; #Teste de McNemar.

Tabela 3 - Média de acertos das questões de V ou F, de análise de itens e objetivas, segundo os grupos.

Acertos	DISCENTES					
	Grupo 1		Grupo 2		Grupo 3	
	1º ao 4º semestre		5º ao 6º semestre		7º ao 10º semestre	
	Média (Desvio padrão)	Mediana (intervalo interquartil)	Média (Desvio padrão)	Mediana (intervalo interquartil)	Média (Desvio padrão)	Mediana (intervalo interquartil)
Questões de V ou F - T1	5,10 (1,85)	5,00 (4,00)	6,60 (1,95)	6,50 (3,00)	4,70 (2,31)	4,50 (3,00)
Questões de itens - T1	1,00 (0,81)	1,00 (2,00)	1,30 (0,48)	1,00 (1,00)	1,40 (0,51)	1,00 (1,00)
Questões objetivas - T1	4,00 (2,10)	4,00 (3,00)	4,70 (1,88)	4,50 (3,00)	6,10 (1,59)	6,50 (3,00)
Questões de V ou F - T2	4,50 (1,08)	5,00 (1,00)	4,00 (1,56)	3,50 (2,00)	4,40 (1,17)	4,50 (2,00)
Questões de itens - T2	0,90 (0,73)	1,00 (1,00)	1,40 (0,69)	1,50 (1,00)	0,90 (0,87)	1,00 (2,00)
Questões objetivas - T2	5,10 (1,10)	5,00 (2,00)	5,20 (1,31)	5,00 (2,00)	5,10 (1,91)	5,00 (4,00)

T1 = sem uso do ChatGPT®; T2 = com uso do ChatGPT®.

Tabela 4 - Dados relacionados à correção das questões discursivas com (T2) e sem (T1) a utilização do ChatGPT®, segundo o grupo.

	DISCENTES					
	Grupo 1		Grupo 2		Grupo 3	
	1º ao 4º semestre		5º ao 6º semestre		7º ao 10º semestre	
	n	%	n	%	n	%
Discursiva 1 - T1						
Correta	7	23,3	7	23,3	5	16,6
Incompleta	3	10,0	2	6,6	3	10,0
Incorreta	-	-	1	3,3	2	6,6
Discursiva 2 - T1						
Correta	4	13,3	3	10,0	3	10,0
Incompleta	5	16,6	6	20,0	5	16,6
Incorreta	1	3,3	1	3,3	2	6,6
Discursiva 1 - T2						
Correta	9	30,0	10	33,3	9	30,0
Incompleta	1	3,3	-	-	1	3,3
Incorreta	-	-	-	-	-	-
Discursiva 2 - T2						
Correta	8	26,6	9	30,0	6	20,0
Incompleta	2	6,6	1	3,3	4	13,3
Incorreta	-	-	-	-	-	-

n: número absoluto; %: frequência relativa; -: valor numérico igual a zero.

DISCUSSÃO

Considerando que o avanço das tecnologias é iminente e que sua utilização no âmbito acadêmico é quase que inevitável, estudar seu uso na educação é imprescindível para avaliar seu impacto em curto, médio e longo prazo. Neste estudo, buscou-se testar o uso do ChatGPT® em resolução de provas, especificamente questões de Odontologia, em modalidades distintas: questões de V ou F, questões de avaliação de itens, questões objetivas e questões discursivas. O foco da pesquisa foi analisar potenciais diferenças de

respostas obtidas pela IA em relação à resolução de humanos sem auxílio de nenhuma ferramenta digital.

Vinte e oito dos 30 dos participantes da pesquisa afirmaram ter familiaridade com internet e quase a metade declarou utilizar ou já ter utilizado o ChatGPT® em atividades acadêmicas, destacando-se o grupo de acadêmicos do primeiro ao quarto semestre, visto que todos afirmaram utilizar a ferramenta em atividades relacionadas à graduação, o que comprova o avanço deste tipo de tecnologia dentro do ambiente acadêmico,

desde o início do curso, tornando urgente a discussão acerca do tema.

Nesta pesquisa, notou-se que houve diferença entre os grupos das questões de V ou F: utilizando o chat, o número médio de acertos dos grupos caiu de 5,6 para 4,3. Isto aponta e exemplifica a limitação técnico-científica da base de dados do ChatGPT®, que em muitas vezes não conseguia distinguir quais itens eram falsos e apontava que todos os itens eram verdadeiros, erroneamente.

A mesma situação ocorreu com as questões de avaliação de itens, em que o ChatGPT®, com 10 participantes, deu respostas que não condiziam com nenhuma alternativa disponível, o que foi considerado como incorreto. Já os discentes se viam obrigados a selecionarem uma das opções e não apontavam resposta fora das alternativas.

Ao utilizar o ChatGPT® na resposta de questões objetivas, observou-se uma discreta semelhança entre os momentos de resposta, quando utilizaram e quando não utilizaram a ferramenta. Esta semelhança pode ser atribuída, em grande parte, à natureza direta e técnica das questões, que exigiam conhecimento específico em assuntos variados relacionados à prática clínica. Além disso, os discentes participantes do estudo podem não ter tido contato prévio com o conteúdo abordado nas questões, já que alguns discentes eram de semestres iniciais do curso, o que pode ter contribuído para a similaridade nos resultados medianos entre os grupos. Vale ressaltar que, embora o ChatGPT® seja uma ferramenta poderosa em muitos aspectos, sua limitação no conhecimento

técnico-científico pode ter influenciado na consistência das respostas, especialmente em um contexto acadêmico que demanda profundo entendimento de temas especializados, como é a Odontologia.

Segundo Brown *et al.* (2020)¹⁵, o ChatGPT® demonstra versatilidade ao realizar uma ampla gama de tarefas relacionadas ao processamento de linguagem natural. Essas atividades abrangem desde tradução automática, resumo automático de texto, geração de conteúdo até preenchimento de lacunas em textos, entre outras funcionalidades. Adicionalmente, o modelo pode ser personalizado para desempenhar funções específicas, tais como *chatbots* para atendimento ao cliente, assistentes virtuais e análise de sentimentos em redes sociais. A adaptabilidade do modelo permite sua aplicação em diversas áreas, proporcionando soluções eficazes e ajustadas às necessidades particulares de cada contexto.

Neste contexto, o grande ponto de destaque positivo para o ChatGPT® nesta pesquisa foi na resolução das questões discursivas, visto que, com o ChatGPT®, os discentes obtiveram um percentual de 83,3% de acerto, pois a ferramenta atende a todos os comandos solicitados na formulação de textos, enquanto os discentes, muitas vezes, acabam dando respostas incorretas ou incompletas, o que destaca a eficácia do ChatGPT® na compreensão dos questionamentos e na produção de resposta a perguntas mais elaboradas. Entretanto, é crucial ressaltar que, apesar dos resultados promissores em questões discursivas, a aplicação do

ChatGPT® em avaliações acadêmicas suscita questões éticas, legais e pedagógicas importantes.

Outro ponto a ser avaliado neste sentido é que as questões de V ou F, de análise de itens e objetivas tratavam de conhecimentos técnicos da clínica odontológica, como cirurgia, prescrição medicamentosa e anestesiologia. Já as questões discursivas, retiradas de provas do Exame Nacional de Desempenho de Estudantes (ENADE), abordaram questões gerais de Saúde Pública. Isto mostra que a base de dados do ChatGPT® pode ser mais assertiva em questões de Saúde Coletiva, mas pode ser limitada em outras disciplinas da Odontologia.

A introdução de tecnologias como o ChatGPT® na educação, embora proporcione benefícios em termos de automação e eficiência, também gera inquietações sobre a validade e equidade das avaliações. O desafio consiste em alcançar um equilíbrio entre a incorporação de ferramentas avançadas e a preservação da integridade acadêmica. Algumas limitações durante a execução deste estudo estiveram relacionadas à captação de voluntários para manter a quantidade equivalente entre os grupos. Muitos discentes foram convidados a participar, porém vários alegaram não terem familiaridade com a utilização do ChatGPT® com a finalidade proposta na pesquisa e dispensaram sua participação. Outra dificuldade encontrada foi durante a execução da metodologia, pois muitas dúvidas surgiram quanto à utilização da IA na formulação das respostas para as questões propostas, mesmo por aqueles

que alegaram ter familiaridade com a plataforma.

Considerando os aspectos observados, a pesquisa proporcionou uma visão abrangente sobre a receptividade dos estudantes em relação ao ChatGPT®, destacando a familiaridade crescente com essa tecnologia no ambiente acadêmico. A constatação de que quase metade dos participantes já utilizou o ChatGPT® em atividades acadêmicas, especialmente aqueles no início da graduação, sinaliza a rápida adoção dessa ferramenta desde seu lançamento. Contudo, a variação nos resultados entre os grupos evidencia a complexidade de incorporar uma tecnologia que, embora versátil, ainda apresenta limitações específicas em contextos educacionais especializados, como a Odontologia. Provavelmente, os dados obtidos neste estudo podem ser extrapolados para outros discentes de graduação em Odontologia do Brasil, por compartilharem semelhanças entre os novos projetos pedagógicos dos cursos de Odontologia e pela base de dados do ChatGPT® ser a mesma.

Nesse sentido, a pesquisa proporciona *insights* valiosos para educadores, instituições de ensino e desenvolvedores de tecnologia, destacando a necessidade de avaliações cuidadosas, adaptações contínuas e discussões éticas para uma integração eficaz e responsável do ChatGPT® no ambiente acadêmico.

CONCLUSÃO

Esta pesquisa demonstrou a eficácia do ChatGPT® para resolução de

questões discursivas. Nelas, a ferramenta é capaz de dar uma resposta extensa e detalhada acerca do comando que lhe é solicitado. Nas questões de V ou F, apesar do desempenho mediano nas respostas corretas, utilizando o ChatGPT® os voluntários conseguiram ter um ponto acima na média de acertos. Contudo, o número médio de acertos entre os momentos em que os discentes utilizaram

e não utilizaram a ferramenta em questões objetivas e de avaliação de itens foram iguais (média de 5 acertos de 10 questões e média de 1 acerto de 2 questões, respectivamente). Desta forma, conclui-se que há limitações quanto ao uso da ferramenta no âmbito acadêmico, provavelmente por conta da base de dados insuficiente do ChatGPT® até o momento de realização da pesquisa.

ABSTRACT

Introduction: ChatGPT® is a public tool developed by OpenAI that employs the language model technology of GPT. This chatbot is capable of addressing various text-based requests. Objective: To assess whether ChatGPT® can be the sole source of information for resolving Dentistry exams. Materials and Methods: This is an analytical quantitative cross-sectional study. For data collection, a fictitious exam was created, consisting of questions from the National Student Performance Exam (ENADE) and other public competitions. Participants answered the exam at two different times: T1, without ChatGPT®, and, after 15 days (T2), using it. The sample included 30 undergraduate Dentistry students, equally divided into three groups: 1st to 4th semester, 5th to 6th semester, and 7th to 10th semester. Descriptive and inferential statistical analyses were applied using SPSS software, including the Wilcoxon and McNemar tests. Results: They revealed a notable effectiveness of ChatGPT® in resolving essay questions, with an 83.3% accuracy rate, while students provided more incorrect or incomplete answers. However, limitations of the ChatGPT® database were observed regarding objective questions. It is crucial to emphasize that, despite promising results, the application of Chat raises ethical and pedagogical questions. Therefore, the introduction of ChatGPT® in education raises concerns about the validity and fairness of assessments, underscoring the importance of finding a balance between technological innovation and the preservation of academic integrity.

KEYWORDS

Forensic dentistry; Dental education; Artificial intelligence.

REFERÊNCIAS

1. Howard J. Artificial intelligence: Implications for the Future of Work. *American Journal of Industrial Medicine* [Internet]. 2019 Aug 22;62(11):917–26. <https://doi.org/10.1002/ajim.23037>
2. Silva JAS da, Mairink CHP. Inteligência artificial: LIBERTAS: Revista de Ciências Sociais Aplicadas [Internet]. 2019 Dec 13;9(2):64–85.
3. Garcia AC. Ética e Inteligência Artificial. *Computação Brasil* [Internet]. 2020 Nov 16;(43):14–22. <https://doi.org/10.5753/CompBR.2020.43.1791>
4. King MR. The Future of AI in Medicine: A Perspective from a Chatbot. *Annals of Biomedical Engineering*. 2022 Dec 26; <https://doi.org/10.1007/s10439-022-03121-w>
5. OpenAI. OpenAI [Internet]. OpenAI. 2023.
6. Brockman G, Cheung V, Pettersson L, Schneider J, Schulman J, Tang J, et al. OpenAI Gym [Internet]. <https://doi.org/10.48550/arXiv.1606.01540>
7. Karobari MI, Adil AH, Basheer SN, Murugesan S, Savadamoorthi KS, Mustafa M, Abdulwahed A, Almokhatieb AA. Evaluation of the Diagnostic and Prognostic Accuracy of Artificial Intelligence in Endodontic Dentistry: A Comprehensive Review of Literature. *Comput Math Methods Med*. 2023 Jan 31;2023:7049360. <https://doi.org/10.1155/2023/7049360>
8. Ahmed N, Abbasi MS, Zuberi F, Qamar W, Halim MSB, Maqsood A, Alam MK. Artificial Intelligence Techniques: Analysis, Application, and Outcome in Dentistry-A Systematic Review. *Biomed Res Int*. 2021 Jun 22;2021:9751564. <https://doi.org/10.1155/2021/9751564>

9. Bonny T, Al Nassan W, Obaideen K, Al Mallahi MN, Mohammad Y, El-Damamhoury HM. Contemporary Role and Applications of Artificial Intelligence in Dentistry. *F1000Res*. 2023 Sep 20;12:1179. <https://doi.org/10.1155/2023/7049360>
10. Thurzo A, Urbanová W, Novák B, Czako L, Siebert T, Stano P, Mareková S, Fountoulaki G, Kosnáčová H, Varga I. Where Is the Artificial Intelligence Applied in Dentistry? Systematic Review and Literature Analysis. *Healthcare (Basel)*. 2022 Jul 8;10(7):1269. <https://doi.org/10.3390%2Fhealthcare10071269>
11. Lund B, Ting W. Chatting about ChatGPT: How May AI and GPT Impact Academia and Libraries? [Internet]. *papers.ssrn.com*. Rochester, NY; 2023. <https://dx.doi.org/10.2139/ssrn.4333415>
12. Webber C, Flores D, Fracasso D. Inteligência Artificial na Escola: Rumo às Novas Experiências Computacionais. [Internet]. 2021. 9:27. <https://dx.doi.org/10.2139/ssrn.4333415>
13. Cotton DRE, Cotton PA, Shipway JR. Chatting and cheating: Ensuring academic integrity in the era of ChatGPT. *Innovations in Education and Teaching International*. 2023 Mar 13;1–12. <https://doi.org/10.1080/14703297.2023.2190148>
14. Dehouche N. Plagiarism in the age of massive Generative Pre-trained Transformers (GPT-3). *Ethics in Science and Environmental Politics*. 2021 Mar 25;21:17–23. <https://doi.org/10.3354/esep00195>
15. Brown TB, Mann B, Ryder N, Subbiah M, Kaplan J, Dhariwal P, et al. Language Models are Few-Shot Learners. *arxiv.org* [Internet]. 2020 May 28; <https://doi.org/10.48550/arXiv.2005.14165>