

RESEARCH

Open Access



Evaluating the accuracy of Chat Generative Pre-trained Transformer version 4 (ChatGPT-4) responses to United States Food and Drug Administration (FDA) frequently asked questions about dental amalgam

Mehmet Buldur¹ and Berkant Sezer^{2*}

Abstract

Background The use of artificial intelligence in the field of health sciences is becoming widespread. It is known that patients benefit from artificial intelligence applications on various health issues, especially after the pandemic period. One of the most important issues in this regard is the accuracy of the information provided by artificial intelligence applications.

Objective The purpose of this study was to the frequently asked questions about dental amalgam, as determined by the United States Food and Drug Administration (FDA), which is one of these information resources, to Chat Generative Pre-trained Transformer version 4 (ChatGPT-4) and to compare the content of the answers given by the application with the answers of the FDA.

Methods The questions were directed to ChatGPT-4 on May 8th and May 16th, 2023, and the responses were recorded and compared at the word and meaning levels using ChatGPT. The answers from the FDA webpage were also recorded. The responses were compared for content similarity in "Main Idea", "Quality Analysis", "Common Ideas", and "Inconsistent Ideas" between ChatGPT-4's responses and FDA's responses.

Results ChatGPT-4 provided similar responses at one-week intervals. In comparison with FDA guidance, it provided answers with similar information content to frequently asked questions. However, although there were some similarities in the general aspects of the recommendation regarding amalgam removal in the question, the two texts are not the same, and they offered different perspectives on the replacement of fillings.

Conclusions The findings of this study indicate that ChatGPT-4, an artificial intelligence based application, encompasses current and accurate information regarding dental amalgam and its removal, providing it to individuals seeking access to such information. Nevertheless, we believe that numerous studies are required to assess the validity and reliability of ChatGPT-4 across diverse subjects.

*Correspondence:

Berkant Sezer

berkant.sezer@comu.edu.tr; dt.berkantsezer@gmail.com

Full list of author information is available at the end of the article



© The Author(s) 2024. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

Keywords Artificial intelligence, Public health, Dental amalgam, Health informatics, Mercury poisoning, ChatGPT, ChatGPT-4

Background

The potential for artificial intelligence (AI) to reduce the substantial time commitment required of dental professionals is noteworthy [1–3]. Moreover, AI holds promise in facilitating individuals to enhance their health at reduced expenses, affording personalized, anticipatory, and preventive dental care, and integrating healthcare services to cater to the needs of all individuals [2]. AI has the potential to enhance the quality of dental care by improving the precision and efficacy of diagnosis, generating superior treatment visuals, simulating results, and forecasting oral health and diseases [4, 5]. Incorporating AI systems as an additional tool for dentists can greatly enhance the provision of top-quality dental treatment. By utilizing AI, we can anticipate more accurate predictions of treatment outcomes and enhancements in the precision of diagnosis and treatment planning. Although deep learning mainly aids in dental diagnosis, AI purports to enhance accuracy and precision while also boosting dentists' productivity [6]. AI applications have demonstrated their usefulness in various areas of endodontics, such as root fractures, periapical lesions, dental and root caries, stem cell viability, root canal system anatomy, and other related fields [5, 7].

Artificial intelligence encompasses Large Language Models (LLMs), which replicate the capability of humans to comprehend written text. Language models are developed using deep learning algorithms that employ massive quantities of textual data collected from diverse sources such as books, articles, and websites. Through the analysis of patterns and connections in the input data, language models learn to anticipate the occurrence of specific words or phrases in a particular context. The latest addition to the category of LLMs is Chat Generative Pre-trained Transformer version 4 (ChatGPT-4) [8]. Utilizing deep learning techniques, ChatGPT-4 has been trained on copious amounts of data to generate responses to user inquiries that closely resemble human-like language. This model consists of 175 billion parameters [9]. Designed to function as a dialog agent for multiple purposes, ChatGPT-4 can provide responses on a vast range of topics [10]. Possibly introducing a new era of models that can enhance the representation of the amalgamation of clinical knowledge and conversation-based interaction, ChatGPT-4 may be the pioneer [11]. As a simulated physician or patient, ChatGPT-4's response-generating interface driven by narratives, has the potential to offer significant benefits [12]. Although demonstrating promise, these models have not yet been entirely successful in assessing clinical knowledge by means of

question-answering tasks [13]. Promoting awareness and education regarding the proper utilization and concealed risks of LLMs based on AI in the medical field is crucial, according to a different study [14].

ChatGPT-4, built upon the foundation of GPT-4, has captured the attention of millions of scientists with its remarkable, human-like conversational responses [15]. In contrast to previous applications, ChatGPT-4 has developed conversational capabilities by leveraging an extensive knowledge base. This enables informative communication aimed at enriching decision-making knowledge [16]. The user-friendly feature of ChatGPT-4 streamlines the diagnostic process, bringing about a notable shift in the current landscape. Additionally, its advancements are poised to shape the digital health landscape [17–19]. Used in a wide variety of applications, including virtual assistants, chatbots, and language translation, its ability to understand and respond to natural language makes it a tool for personalized health-related applications [20]. These applications, which have the ability to provide ideas from specific disease groups to public health, from diagnosis to treatment options, are increasingly finding wider use in the field of health [20, 21].

ChatGPT-4 exhibits a broad spectrum of applications within the dentistry domain. These encompass telemedicine services in dentistry, aiding dentists in their clinical decision-making, contributing to the education of dental students, and providing support in crafting scientific articles, evaluations, and patient information [22, 23]. Customized and refined applications that utilize extensive language models have the potential to enhance the quality of dental telemedicine services [16, 24]. In the imminent future, it is anticipated that numerous language model applications will adeptly collect patient data, analyze symptoms, and generate initial diagnoses, subject to subsequent review by a human dentist [17, 25]. Particularly in underserved areas with restricted access to dental care, expansive language models may offer increased utility in dental telemedicine services.

Despite the conclusions drawn from literature reviews indicating that the presence of mercury in dental amalgam could potentially lead to adverse health effects, the current scientific evidence does not provide compelling support for a causal relationship between dental amalgam and diseases [26]. Given the inherently toxic properties of mercury in all of its forms, it is conceivable that exposure to dental amalgam could give rise to both local and systemic deleterious effects [27]. For these reasons, individuals' demand and desire to obtain information

about amalgam fillings and their removal has increased in recent years [28].

Large language models have many potential applications, ranging from streamlined dental record-keeping to the clinical decision-making process [29]. However, LLMs can provide entirely incorrect answers, generate nonsensical content, and present misleading information, causing serious concerns in critical areas such as healthcare [18]. This requires the models to be trained with dentistry teaching materials, patient records, and other relevant domain information to capture pertinent patterns, terminology, and context, ensuring increased accuracy. As a result, the models develop a profound understanding of dentistry concepts and become capable of generating contextually relevant and informative responses [22]. In order to prevent the spread of false information on the subject and for the public to receive correct information on this subject, dentistry and public health institutions provide great benefits in informing by answering the questions frequently asked by patients. There is no research in the scientific literature evaluating the answers of ChatGPT-4 to the questions frequently asked by patients about dental amalgam. The aim of the study was to ask the frequently asked questions (FAQ) about dental amalgam determined by the United States Food and Drug Administration (FDA), which is one of these information resources, to ChatGPT-4 and to compare the content of the answers given by the application with the answers of the FDA. For comparison, the brochure on the FDA website was used due to its global presence as an organization that compiles and answers FAQ related to this topic, which patients can easily access through the website.

Methods

Methodological framework: study setup, search engine selection, bias avoidance, and data cleaning

A new email address was created for the purpose of working to avoid any negative bias effects of search algorithms. Access to the ChatGPT-4 application was made through a search engine query using Google (Google Inc., California, United States). The 'Continue with Google' tab was preferred at the application login. In order to mimic a typical user experience, the most widely used search engine worldwide was chosen. Prior to the question-and-answer (Q&A) session, all search history and cookies were cleared from the computer. The 'Clear Browsing Data' section under the 'Privacy and Security' tab in the 'Settings' menu of the search engine was utilized. By clicking the checkboxes for 'Cached Images and Files,' 'Cookies and Other Site Data,' and 'Browsing History,' the 'Time Range' was selected to align with the timeframe of the study, and the data was cleared.

Question selection process: source, exclusions, and inclusions

As part of the study, the first four questions out of six from the FAQ section of the FDA webpage were selected and analyzed [30]. The fifth question was excluded from the study as it inquired about material information. The sixth question was also not included in the analysis as it provided a brief explanation more in the form of guidance than information. The four questions analyzed in the study were as follows:

- Q.1. What Is Dental Amalgam?
- Q.2. Is Dental Amalgam Safe?
- Q.3. Who Should Be Concerned About Dental Amalgam?
- Q.4. Should Dental Amalgam Fillings Be Removed?

Access and testing protocol for ChatGPT-4

ChatGPT-4 relies on a general language model to understand user questions across various topics. Therefore, answers may change over time due to technological advancements, scientific discoveries, or other factors. Hence, a weekly difference assessment was conducted to monitor potential variations. The current version of the application (GPT-4) was accessed with internet connectivity from Çanakkale/Türkiye, on May 8, 2023, at 10:00 (<https://openai.com/gpt-4>) [31]. The selected four questions were asked in their original language, English, without modification, and the answers were reported in English as well. The same process was repeated one week later (May 16, 2023), at the same location and time, and the responses were once again documented. The questions were asked uninterruptedly and in the same order within a single text message. The specified four questions were asked one by one and consecutively within the same conversational context (Supplementary Material 1). At the same time, the answers on the FDA web page are also recorded and listed.

Answer comparison methodology: sources and response similarity analysis

The similarity of the responses given to the same questions with a one-week interval was also investigated by conducting a dialogue through ChatGPT-4 in the same manner. After the Q&A session held on May 8th, the responses provided by ChatGPT-4 were compared with the answers in the FDA's informational brochure. ChatGPT-4 and FDA response texts were individually subjected to similarity checks using the 'Doc-to-Doc Comparison' feature in a similarity detection program (iThenticate "<http://www.ithenticate.com>"), and their similarities were reported. Sentiment analysis is a field of natural language processing that aims to make inferences about emotions and thoughts from texts related to certain entities and topics, and research in this field aims

to investigate various aspects of one or more sentences in a text that can be associated with different emotions at different levels of detail [32]. It is important to have high accuracy in sentiment analysis in AI-based LLMs. Recent studies have reported that ChatGPT represents at least the right direction regarding specialized questions in the field of health and bio(medical) sciences [33–35]. Similarly, the increase in the effectiveness of topic modeling increases the validity of an AI-based application [36]. Since the first day ChatGPT was released, it has been reported that its effectiveness in topic modeling has increased with the updates to the program [36–38]. It has also been reported that ChatGPT has great potential in evaluating the realism of text summarization [39]. It has been shown that the text summarization proficiency of ChatGPT, which is another important parameter in AI-based applications, is quite good, and while text classification algorithms can distinguish between real and generated summaries, humans cannot distinguish between real summaries and those generated by ChatGPT [40]. Considering its reported success in all these matters, it was decided to use ChatGPT-4, the latest version of Chat-GPT, in the study.

ChatGPT-4 response evaluation: criteria and comparison methodology

The accuracy of the responses provided by ChatGPT-4 to the specified 4 questions, along with identifying similarities and differences, has been thoroughly examined within the scope of ChatGPT-4 capabilities based on the following 4 items, and reported by the researchers;

1. **Main Idea:** Analysis for identify the central theme or message of each answers. In the answer given to a curious question on the relevant subject, this issue was taken into account as the questioner first tried to learn the main idea of the answer given [41].
2. **Quality Analysis:** The accuracy and reliability of the information obtained by the questioner basically shows the quality of the information. Especially in health sciences, accurate and real information can only be achieved with quality answers [42–44]. Evaluating each answers using relevant criteria based on the questions provided below;

a. Relevance

- a.1. Does the answer address the core of the question and avoid straying off topic?
- a.2. Does the answer stay within the intended scope of the question?

b. Accuracy

- b.1. Is the information factually correct and verified where possible (or evidence-supported)?
 - b.1.1. Can the information be supported by relevant references?
- b.2. Are there any misleading or contradictory statements?

c. Completeness

- c.1. Does the answer provide enough information to fully answer the question?
- c.2. Are there any key points missing or left vague?

d. Clarity

- d.1. Is the language clear, concise, and easy to understand?
- d.2. Are there any jargon or complex terms that may hinder understanding?

3. **Summary of Common Ideas:** The common ideas between the answers prepared by human and scientific authorities, and the answers ensured by AI provide an idea about the validity and reliability of AI-based applications [45]. The aim is to identify overlapping themes, concepts, or information presented in both answers.
4. **Summary of Inconsistent Ideas:** As with common ideas, inconsistent ideas also provide insight into the validity and reliability of AI-based applications [46, 47]. Areas where the answers differ in their perspectives, information, or conclusions have been investigated.

Results

The ChatGPT-4 was used to obtain the following answer for the similarity of questions asked with a one-week interval and the similarity of the answers provided in terms of both meaning and word usage:

The two texts are quite similar in content and meaning, as both describe dental amalgam as a filling material that contains silver, tin, copper, and mercury. They both explain how the amalgam is prepared and used to fill cavities caused by tooth decay in both primary and permanent teeth. They also touch on the ongoing debate over the safety of dental amalgam due to the presence of mercury and mention the opinions of various organizations, such as the American Dental Association and the U.S. Food and Drug Administration. Overall, the two texts share similar information but present it in slightly different ways.

In the examination conducted through the word similarity program, the similarity percentages of the response texts were reported as 13%, 0%, 4%, and 9% according to the order of questions. Following this analysis, in the first Q&A session, the responses provided by ChatGPT-4 are outlined below, focusing on the main points and themes of the answers.

ChatGPT-4 responses

Q.1. What is dental amalgam?

The main focus of the answer to this question is generally on dental amalgam, a commonly used filling material in dentistry. The text discusses the composition of dental amalgam, which includes silver, tin, copper, and mercury, and how it is prepared for use in tooth cavities. The text emphasizes that dental amalgam is considered safe and effective by reputable organizations such as the American Dental Association (ADA), the FDA, and the World Health Organization (WHO). Additionally, it highlights the durability and long-lasting nature of dental amalgam as a filling material.

Q.2. Is dental amalgam safe?

The main focus of the text is on the reliability of dental amalgam. The text acknowledges that dental amalgam contains mercury, recognizing it as a toxic substance. However, it emphasizes that the mercury in dental amalgam is in a stable compound, not easily absorbed by the body, and the amount of mercury released by dental amalgam fillings is very low and within safe limits. The text highlights the consensus of reputable organizations, such as the ADA, the FDA, and the WHO, on the safety and effectiveness of dental amalgam in dental restorations. It underscores the importance of decisions regarding dental materials being made in consultation with a dentist based on the patient's specific needs and preferences.

Q.3. Who should be concerned about dental amalgam?

The text emphasizes specific groups that may be more sensitive to the potential negative effects of mercury exposure in dental amalgam. Individuals with kidney problems or allergies to metals in dental amalgam, pregnant women, and young children are mentioned as groups that could be more susceptible to potential harmful effects. The text notes that individuals with a large number of existing amalgam fillings may have a higher risk of mercury exposure over time. It highlights the importance of sharing concerns with a dentist, evaluating individual risks and benefits, and determining the best course of action based on specific needs and circumstances.

Q.4. Should dental amalgam fillings be removed?

The text emphasizes that the decision to remove dental amalgam fillings should be made individually, in consultation with a qualified dental professional. While acknowledging the presence of small amounts of mercury in these fillings, it suggests that the released mercury is generally considered safe. The text discourages removal solely for reducing mercury exposure, as the process itself may increase exposure. However, it acknowledges situations where removal may be recommended, such as for broken or decayed fillings or for individuals with known allergies or sensitivities. The key message is to consult with a qualified dental professional to make decisions based on individual dental and medical needs.

The comparative text analysis of the responses provided by the FDA to FAQs and the responses provided by ChatGPT-4 to the same questions was conducted by ChatGPT-4 under four main headings: 'Main Idea,' 'Quality Analysis,' 'Common Ideas,' and 'Inconsistent Ideas' (Supplementary Materials 2–5). These analyses are summarized in Tables 1, 2 and 3, and 4, respectively.

Discussion

Considering the effects of LLMs-based AI applications, which have been increasingly used in many fields, including the field of health, in recent years, the main purpose of this study was to investigate the accuracy and effectiveness of ChatGPT-4's answers to FAQ about dental amalgam by comparing them with the answers given by the FDA. After the evaluation, it was observed that although ChatGPT-4's answers to FAQ obtained from the FDA's website regarding dental amalgam, its effects, and removal were mostly different in form, they provided accurate and effective information in content. The importance of patients accessing accurate and reliable information about their health has once again emerged as face-to-face communication decreased during the pandemic period [48]. Considering the sociological effects of developing technologies, individuals' desire to access information on health-related issues through AI-based applications has also increased [49]. This has paved the way for the development of LLMs-based applications such as ChatGPT-4, beyond searches via the standard and traditional search engine. Considered in this context, taking into account the limitations and reservations, it is important for these applications to provide accurate and reliable information to users who want to access information, in terms of individual and social health in the field of health [7]. In this study, it has been shown that ChatGPT-4 provides accurate and reliable answers about dental amalgam, its effects, and removal.

Considering the increasing use in the field of health-care, ChatGPT-based scientific publications in this context in recent years have reported that the application is

Table 1 The main idea analysis of ChatGPT-4 answers and FDA guidance for the same questions (Generated by ChatGPT-4).

QUESTIONS	The Central Theme of Each Text of ChatGPT-4 and FDA	
	ChatGPT-4	FDA
Q.1 What Is Dental Amalgam?	The safety and effectiveness of dental amalgam as a filling material is highlighted, despite containing mercury, it is considered safe and effective by reputable organizations like the ADA, FDA, and WHO.	Provide a basic description of dental amalgam, outlining its composition and common name ("silver-filling") due to its appearance. The mixture of metals used in dental amalgam includes mercury, silver, copper, tin, and zinc, with a specific emphasis on the significant proportion of mercury in the mixture (approximately 50% by weight).
Q.2 Is Dental Amalgam Safe?	Dental amalgam's safety, especially its mercury content, is a subject of ongoing debate and research. Organizations like the ADA, FDA, and WHO generally consider it safe, but some individuals may be sensitive to its components. It's wise to consult a dentist to make informed decisions on filling materials.	Dental amalgam fillings can release small amounts of mercury vapor, especially during placement or removal. Ingesting small particles of amalgam is not known to pose a health risk, but inhaling mercury vapors may be risky for some. Studies have not conclusively linked amalgam to harmful health effects in the general population, though those with many fillings may have slightly higher mercury levels in their blood or urine.
Q.3 Who Should Be Concerned About Dental Amalgam?	Pregnant women, young children, those with specific health conditions, and individuals with many existing amalgam fillings are more vulnerable to mercury exposure. It's crucial to consult a dentist to determine the best approach based on individual needs.	Highlight the vulnerability of pregnant or nursing women, young children, those with allergies to mercury, and individuals with neurological or kidney issues to the adverse effects of mercury from dental amalgam. Emphasize caution in placing amalgam fillings in these groups. Mention the importance of research on mercury vapor exposure effects, especially on vulnerable populations, and signs of excessive exposure. Inform healthcare providers and patients about the risks and need for personalized decisions.
Q.4 Should Dental Amalgam Fillings Be Removed?	It's important to consult a dentist before removing dental amalgam fillings. While they release low levels of mercury, unnecessary removal solely to reduce exposure is cautioned against. Removal may be recommended for dental health reasons or due to allergies, but decisions should be based on individual needs and made in consultation with a professional.	Unless necessary for dental or medical reasons, it's not recommended to remove intact dental amalgam fillings. Highlight risks like tooth damage and temporary exposure to more mercury vapor. Even for higher-risk groups like pregnant women and children, removal should be considered only when medically necessary. Discuss with a dentist or physician, especially for those with health conditions or sensitivities.

Table 2 The summary of Quality Analysis ChatGPT-4 answers and FDA guidance for the same questions (Generated by ChatGPT-4).

		Answer 1		Answer 2		Answer 3		Answer 4	
		ChatGPT-4	FDA	ChatGPT-4	FDA	ChatGPT-4	FDA	ChatGPT-4	FDA
Relevance	a.1. Does the answer address the core of the question and avoid straying off topic?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	a.2. Does the answer stay within the intended scope of the question?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Accuracy	b.1. Is the information factually correct and verified where possible (or evidence-supported)?	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes
	b.1.1. Can the information be supported by relevant references?	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes
	b.2. Are there any misleading or contradictory statements?	No	No	No	No	No	No	No	No
Completeness	c.1. Does the answer provide enough information to fully answer the question?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	c.2. Are there any key points missing or left vague?	No	No	No	No	Yes	Yes	Yes	Yes
Clarity	d.1. Is the language clear, concise, and easy to understand?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	d.2. Are there any jargon or complex terms that may hinder understanding?	No	No	No	No	No	No	Yes	No

a potentially valuable tool for both healthcare providers and patients in facilitating the diagnosis, treatment and prevention of various diseases [7, 50–52]. Although it has been launched very recently, it is known to provide very high rates of accurate information on topics in medical and dental contexts [7, 52]. And as natural language processing technology continues to advance, ChatGPT's accuracy will likely increase, expanding its potential applications in healthcare [36]. In a study comparing the

answers of experts and ChatGPT on endodontic questions, it was reported that although ChatGPT was not yet sufficient for clinical decision-making, it could potentially be used with the development of the application [53]. In another study, ChatGPT's answers to 30 questions about tooth-supported fixed prostheses and removable dental prostheses were evaluated by different experts and it was reported that ChatGPT could not replace a dentist [54]. Vaira et al. [55] evaluated ChatGPT's capacity to

Table 3 The summary of common themes, concepts, or information in responses from ChatGPT-4 and FDA guidance to the same questions (Generated by ChatGPT-4).

Common Ideas of ChatGPT-4 and FDA	
QUESTIONS	
Q.1. What is Dental Amalgam?	<ol style="list-style-type: none"> 1. Both answers discuss dental amalgam, a common filling material used in dentistry. 2. They mention the composition of dental amalgam, which includes mercury, silver, tin, and copper. 3. Both describe the process of preparing and placing dental amalgam into tooth cavities. 4. They touch upon the debate surrounding the safety and effectiveness of dental amalgam. 5. ChatGPT-4 highlights endorsements from organizations like the ADA, FDA, and WHO. 6. FDA emphasizes the mercury content in dental amalgam.
Q.2. Is Dental Amalgam Safe?	<ol style="list-style-type: none"> 1. Both answers discuss safety concerns about dental amalgam, especially regarding mercury and its potential release in vapor form. 2. They mention assessments by organizations like the ADA, FDA, and WHO regarding dental amalgam safety. 3. Both acknowledge the possibility of allergic reactions or sensitivities to the metals in dental amalgam. 4. They mention the availability of alternative filling materials. 5. Both emphasize the importance of informed decision-making with a dentist based on individual needs and preferences.
Q.3. Who Should Be Concerned About Dental Amalgam?	<ol style="list-style-type: none"> 1. Both answers discuss the risks of mercury exposure from dental amalgam, especially for vulnerable groups like pregnant women, young children, nursing mothers, and individuals with specific health conditions. 2. They emphasize the importance of discussing concerns with a qualified dental professional. 3. Both acknowledge the limited data on the health effects of mercury exposure from dental amalgam in certain populations, highlighting the need for further research.
Q.4. Should Dental Amalgam Fillings Be Removed?	<ol style="list-style-type: none"> 1. Both emphasize the importance of individualized decision-making with a dental professional regarding the removal of dental amalgam fillings. 2. They highlight that intact amalgam fillings, especially those without decay beneath them, generally do not need to be removed solely to reduce mercury exposure. 3. Both caution against unnecessary removal, as it may lead to loss of healthy tooth structure and temporary increase in mercury vapor exposure. 4. They stress the need to consider medical necessity and specific health conditions when deciding on removal and replacement of amalgam fillings. 5. Overall, they advocate for a careful assessment of individual circumstances and dental health needs before deciding on dental amalgam removal.

Table 4 The summary of areas where the answers differ in their perspectives, information, or conclusions (Generated by ChatGPT-4).

		ChatGPT-4	FDA
Q.1 What Is Dental Amalgam?	a. Composition	ChatGPT-4 lists metals in dental amalgam without specifying proportions.	FDA specifies that dental amalgam is around 50% mercury by weight.
	b. Perspective on Safety	ChatGPT-4 reassures about the safety of dental amalgam, citing support from reputable organizations.	FDA indirectly raises concerns about mercury content, potentially implying health risks.
	c. Use of Terminology	ChatGPT-4 calls dental amalgam a “commonly used filling material.”	FDA mentions dental amalgam as “sometimes called a ‘silver-filling’ due to appearance,” offering an alternative term.
Q.2 Is Dental Amalgam Safe?	a. Perspective on Safety	ChatGPT-4 acknowledges mercury’s presence in dental amalgam, presenting it as a toxic substance but emphasizes its stable compound form and minimal release, backed by reputable organizations.	The FDA also recognizes mercury vapor release during dental procedures, particularly placement or removal of fillings, highlighting inhalation risks. It suggests no health risks from ingesting amalgam particles but potential harm from inhaling mercury vapor.
	b. Presentation of Information	ChatGPT-4 offers a comprehensive safety overview, discussing reassurances from organizations like ADA, FDA, WHO, and potential concerns about metal allergies.	The FDA focuses narrowly on mercury vapor release risks, avoiding broader safety discussions.
	c. Conclusion on Harmful Effects	ChatGPT-4 concludes no evidence of harmful effects from dental amalgam based on organizational reviews.	The FDA suggests inconclusive evidence on harmful effects, implying uncertainty rather than outright affirming safety.
Q.3 Who Should Be Concerned About Dental Amalgam?	a. Coverage of Susceptible Groups	ChatGPT-4 identifies pregnant women, young children, individuals with certain health conditions, and those with numerous amalgam fillings as susceptible to mercury exposure.	The FDA expands on susceptible groups, including pregnant women, nursing mothers, young children, individuals allergic to mercury, and those with neurological or kidney issues. It provides more detailed discussions on potential effects.
	b. Depth of Information	ChatGPT-4 provides a general overview without specific considerations or effects on infants and children.	The FDA offers in-depth information on each susceptible group, including considerations for pregnant women, nursing mothers, and children. It discusses potential effects on infants and children’s neurological systems due to mercury vapor exposure.
	c. Conclusions on Health Effects	ChatGPT-4 stresses the importance of discussing concerns with dental professionals but refrains from definitive conclusions on health effects.	The FDA acknowledges limited clinical data but suggests signs of excessive mercury exposure. It advocates a cautious approach for susceptible groups, emphasizing the need for further research and individual sensitivity consideration.
Q.4 Should Dental Amalgam Fillings Be Removed?	a. Conditions for Removal	ChatGPT-4 emphasizes that the decision to remove dental amalgam fillings should be made on an individual basis, considering factors such as dental and medical needs, as well as circumstances. It mentions situations such as broken, worn, or decayed fillings, or allergies/sensitivities to metals, where removal may be recommended.	FDA advises against the removal of intact amalgam fillings, particularly if the filling is in good condition and there is no decay beneath it. It suggests that removal may result in unnecessary loss of healthy tooth structure and temporary exposure to mercury vapor. It only recommends removal if it is considered medically necessary by a healthcare professional.
	b. Perspective on Risk	ChatGPT-4 acknowledges the potential risk of increased exposure to mercury vapor and particles during the removal process. However, it suggests that removal may be warranted in certain situations, such as allergies or sensitivities to metals.	FDA highlights the potential risks associated with removal, including loss of healthy tooth structure and temporary exposure to mercury vapor. It strongly advises against removal unless it is deemed medically necessary.
	c. Recommendations for Decision-Making	ChatGPT-4 emphasizes the importance of consulting with a qualified dental professional to make informed decisions about removal, based on individual circumstances and needs.	FDA underscores the need for discussions with a dentist or physician, particularly for individuals with health conditions such as allergies, sensitivities, or neurological/kidney disease.

answer questions and solve clinical scenarios related to head and neck surgery. The researchers noted that the results generally showed a good level of accuracy in the AI’s answers, and its ability to solve complex clinical scenarios was promising [55]. Hatia et al. [56], in their study in which they investigated the accuracy and completeness of ChatGPT’s answers on interseptive orthodontics

and clinical scenarios, concluded that the accuracy of the answers given by AI was highly accurate and complete, although not 100%. In another study where the answers given by the application to FAQ about orthodontics were evaluated, it was reported that the answers were not based on any scientific article and were reliable at a moderate level [57]. Mago et al. [58] reported 100% accuracy

in identifying radiographic landmarks with ChatGPT for questions regarding anatomical accidents, oral and maxillofacial pathologies, and radiographic features of oral and maxillofacial pathologies, respectively. Since not long has passed since its launch and it is a developing technology, there are not many studies in the scientific literature evaluating ChatGPT's answers to questions about dental issues and clinical scenarios. As stated, the results of the few available studies vary. While some researchers found ChatGPT's answers highly adequate, reliable, and accurate, others found them inadequate or moderately adequate.

Dental amalgam, which has been preferred in restorative dentistry for many years due to its longevity and durability, has been the subject of many studies in recent years due to the possible allergic and toxic effects of the mercury it contains. Due to the increase in individual and communal sensitivity to mercury in recent years, many patients apply to clinics to have their amalgam fillings removed [59]. Patients stated that headaches, fatigue, dizziness, memory and concentration impairment, anxiety, depression, irritability, and various musculoskeletal and gastrointestinal symptoms, which they often thought to be caused by dental amalgams, decreased when they removed their amalgam fillings [28, 59].

The gradual reduction of dental amalgam use as part of the Minamata Mercury Convention is one of the five key projects of the WHO's Oral Health Work Plan 2018–2020. The WHO aimed to provide assistance to low-income countries to accelerate their phasing out of amalgam use in the timeline from April 2018 to December 2023 [60]. Although the use of dental amalgam has decreased in many countries in order to support the use of these techniques and materials under appropriate conditions in the presence of better and more reliable restorative techniques and materials and to minimize the harmful environmental effects of mercury, the International Association of Dental Research (IADR) confirms the safety of dental amalgam for the general population without allergy to amalgam components or severe renal diseases.

The European Union Scientific Committee on Emerging and Newly Identified Health Risks (SCENIHR) and the US Agency for Toxic Substances and Disease Registry (ATSDR) reported that amalgam fillings should not be removed except for allergic reactions [61–63]. In addition, it has been shown that there is no significant decrease in blood mercury levels after many years in patients who have had amalgam fillings removed [63]. As one study stated in their qualitative data-based studies, it was concluded that it remains unclear how important it is to replace dental amalgam fillings in terms of improving health complaints [64]. In light of the developments in dental amalgam that have come to the fore in recent

years, it is important for patients to have accurate information about this subject, and the accuracy of ChatGPT-4's answers to questions about this subject is also important.

In addition to its positive aspects such as the convenience it offers in healthcare services, providing an alternative to patient-physician communication, and patients' easy access to information in difficult conditions such as pandemics, there are also some important difficulties in using ChatGPT-4 in healthcare services. Although the results of our study show that ChatGPT-4's responses on dental amalgam are consistent with the FDA's responses, the most significant potential issue with relying on AI technology for healthcare decision-making is the ethical implications. Patients may have difficulty trusting an AI-based technology, and there are concerns about bias in the data used to train the system. In addition, the findings obtained in these AI-based applications as a result of data obtained from patients raise concerns about the security and violation of patient data [20, 21]. The training data of ChatGPT-4 may include questions and answers from the examined FDA website, raising concerns that its responses might merely be rephrased versions of the website's content. Additionally, it's crucial to emphasize that ChatGPT lacks the ability for independent scientific reasoning; instead, it can only produce responses relying on recognized patterns and structures within the text it has been trained on [65, 66]. However, the authors have used a plagiarism detection program to confirm the originality of ChatGPT-4's responses when compared with the information on the website, alleviating this concern.

Considering the importance of oral and dental health, the high incidence of oral and dental diseases, the existence of different opinions on dental materials and methods, and the developments in technology, the efforts of dentists to provide oral and dental health for people are very important [67]. In this context, patients should be informed accurately. With the effects of the pandemic period, many of the patients are researching many subjects on web-based platforms instead of consulting clinics where the risk of transmission is high [68]. However, the validity and reliability of the information on web-based sites and applications are controversial. Studies on many different subjects have evaluated the accuracy of the information on different platforms, and it has been concluded that the Internet, which is like a vast ocean, will not replace the patient-physician relationship in the field of health [69]. However, the dramatic increase in the use of the Internet and AI in the developing world shows that people's applications to the Internet for health, as well as for many other subjects, are increasing [70–72]. For this reason, it is of great importance to present accurate and reliable information to individuals who need access to it.

In the Bloom classification, which is a classification of question types based on cognitive and hierarchical criteria based on medical education, there are question types on a scale ranging from lower-order to higher-order [73]. While lower-order cognitive skills are such as recall and comprehension, higher-order thinking skills include application, analysis, evaluation, and creation [74]. While LLMs-based applications can adequately respond to lower-order skills, they are not as successful in higher-order skills [56]. For this reason, the Bloom classification of the questions asked to AI applications in studies is also important and should also be considered as a limitation [74]. The fact that all four questions obtained from the FDA's website in our study were at a lower-order level may explain the accuracy of ChatGPT-4's answers to questions about dental amalgam and its effects. It is of great importance to evaluate the answers by asking higher-order level complex questions to these LLMs-based applications in future studies. As a matter of fact, the fact that the answers are not sufficient in some studies proves this. We believe that, in the context of dentistry, studies on issues that frequently worry patients and cause information pollution in society will contribute to the scientific literature and individual and social health.

Today, when the use of AI-based applications is increasing, it is known that ChatGPT-4 is used by physicians and patients to access accurate information. Since it is a fairly new application, it is a subject on which scientific research continues. In this context, aspects such as being the first study to evaluate the answers of AI on dental amalgam and being based on the questions on the website of the FDA, one of the institutions with the highest international recognition, validity and reliability regarding FAQ, constitute the strengths of our study. On the other hand, the fact that a limited number of questions were evaluated and therefore statistical analysis was not possible constitute the main limitations of our study. ChatGPT-4, like many other large language models, faces several significant challenges that impact its reliability and effectiveness. Firstly, its limited understanding of contextual meaning hinders its ability to conduct critical analysis and verify the accuracy of information. While adept at identifying patterns, it lacks the capacity for deep comprehension [75]. Secondly, the model inherits biases and errors present in its training data, which can manifest in its generated text, potentially leading to misleading or inaccurate results [76]. Lastly, the echo chamber effect poses a risk, as ChatGPT-4 may inadvertently perpetuate existing biases by recycling and reframing previously generated information without undergoing rigorous critical evaluation [77]. Moreover, since ChatGPT-4 has limited knowledge of scientific developments that occurred after January 2022, it can provide incomplete and misleading information in response to

current questions, which is another significant limitation. These challenges underscore the importance of ongoing research and development to address the limitations of large language models like ChatGPT-4.

Conclusions

The findings of this study show that the answers given by ChatGPT-4, which is seen as a potential information provider and is increasingly used in today's world where technology is constantly developing, about dental amalgam, its effects, and removal, are compatible in content with the answers given by the FDA to the same questions.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12903-024-04358-8>.

Supplementary Material 1
Supplementary Material 2
Supplementary Material 3
Supplementary Material 4
Supplementary Material 5

Acknowledgements

Not applicable.

Author contributions

M.B. and B.S. conceived the idea, M.B. collected the data, M.B. and B.S. wrote the main manuscript. All authors reviewed the manuscript.

Funding

No funding was obtained for this study.

Data availability

The datasets underpinning the outcomes of this study can be obtained from the corresponding author, B.S., upon a reasonable and substantiated request.

Declarations

Ethics approval and consent to participate

Not applicable. Ethics approval and consent to participate were not required for this study.

Consent for publication

Not applicable. Consent for publication was not required for this study.

Competing interests

The authors declare no competing interests.

Author details

¹Department of Restorative Dentistry, School of Dentistry, Çanakkale Onsekiz Mart University, Çanakkale, Türkiye

²Department of Pediatric Dentistry, School of Dentistry, Çanakkale Onsekiz Mart University, Çanakkale, Türkiye

Received: 19 October 2023 / Accepted: 9 May 2024

Published online: 24 May 2024

References

1. Abdullah R, Fakhieh B. Health Care employees' perceptions of the Use of Artificial Intelligence Applications: Survey Study. *J Med Internet Res*. 2020;22:e17620. <https://doi.org/10.2196/17620>.
2. Geis JR, Brady AP, Wu CC, Spencer J, Ranschaert E, Jaremko JL, Langer SG, Borondy Kitts A, Birch J, Shields WF, van den Hoven R, Kotter E, Wawira Gichoya J, Cook TS, Morgan MB, Tang A, Safdar NM, Kohli M. Ethics of Artificial Intelligence in Radiology: Summary of the joint European and north American Multisociety Statement. *Radiology*. 2019;293:436–40. <https://doi.org/10.1148/radiol.2019191586>.
3. Schwendicke F, Samek W, Krois J. Artificial Intelligence in Dentistry: chances and challenges. *J Dent Res*. 2020;99:769–74. <https://doi.org/10.1177/0022034520915714>.
4. Artificial Intelligence (AI). In *Dentistry 2023*; [cited 2023 July 5] <https://dental-reach.today/artificial-intelligence-ai-in-dentistry/>.
5. Agrawal P, Nikhade P. Artificial Intelligence in Dentistry: past, Present, and Future. *Cureus*. 2022;14:e27405. <https://doi.org/10.7759/cureus.27405>.
6. Lee JH, Kim DH, Jeong SN, Choi SH. Detection and diagnosis of dental caries using a deep learning-based convolutional neural network algorithm. *J Dent*. 2018;77:106–11. <https://doi.org/10.1016/j.jdent.2018.07.015>.
7. Umer F. Could AI offer practical solutions for dentistry in the future? *BDJ Team*. 2022;9:26–8. <https://doi.org/10.1038/s41407-022-0830-1>.
8. Brown TB, Mann B, Ryder N, Subbiah M, Kaplan J, Dhariwal P, Neelakantan A, Shyam P, Sastry G, Askell A, Agarwal S, Herbert-Voss A, Krueger G, Henighan T, Child R, Ramesh A, Ziegler DM, Wu J, Winter C, Hesse C, Chen M, Sigler E, Litwin M, Gray S, Chess B, Clark J, Berner C, McCandlish S, Radford A, Sutskever I, Amodei D. Language models are few-shot learners. *Adv Neural Inf Process*. 2020;33:1877–901. <https://doi.org/10.48550/arXiv.2005.14165>.
9. Scott K. Microsoft teams up with OpenAI to exclusively license GPT-3 language model. *The Official Microsoft Blog* 2020; [cited 2023 July 5] <https://blogs.microsoft.com/blog/2020/09/22/microsoft-teams-up-with-openai-to-exclusively-license-gpt-3-language-model/>.
10. Elkhatat AM. Evaluating the authenticity of ChatGPT responses: a study on text-matching capabilities. *Int J Educ Integr*. 2023;19:15. <https://doi.org/10.1007/s40979-023-00137-0>.
11. Suhag A, Kidd J, McGath M, Rajesh R, Gelfinbein J, Cacace N, Monteleone B, Chavez MR. ChatGPT: a pioneering approach to complex prenatal differential diagnosis. *Am J Obstet Gynecol MFM*. 2023;5:101029. <https://doi.org/10.1016/j.ajogmf.2023.101029>.
12. Gala D, Makaryus AN. The Utility of Language models in Cardiology: a narrative review of the benefits and concerns of ChatGPT-4. *Int J Environ Res Public Health*. 2023;20:6438. <https://doi.org/10.3390/ijerph20156438>.
13. Gutiérrez BJ, McNeal N, Washington C, Chen Y, Li L, Sun H, Su Y. Thinking about GPT-3 in-context learning for biomedical IE? Think again. *ACL Anthology*. 2022;4497–512. <https://doi.org/10.48550/arXiv.2203.08410>.
14. Cascella M, Montomoli J, Bellini V, Bignami E. Evaluating the feasibility of ChatGPT in Healthcare: an analysis of multiple clinical and research scenarios. *J Med Syst*. 2023;47:33. <https://doi.org/10.1007/s10916-023-01925-4>.
15. Zheng O, Abdel-Aty M, Wang D, Wang Z, Ding S. ChatGPT is on the horizon: could a large language model be all we need for Intelligent Transportation? Preprint at *arXiv:2303.05382*. 2023. <https://arxiv.org/abs/2303.05382>.
16. Kurian N, Cherian JM, Sudharson NA, Varghese KG, Wadhwa S. AI is now everywhere. *Br Dent J*. 2023;234:72. <https://doi.org/10.1038/s41415-023-5461-1>.
17. Alhaidry HM, Fatani B, Alrayes JO, Almana AM, Alfhaed NK. ChatGPT in Dentistry: a Comprehensive Review. *Cureus*. 2023;15:e38317. <https://doi.org/10.7759/cureus.38317>.
18. Eggmann F, Weiger R, Zitzmann NU, Blatz MB. Implications of large language models such as ChatGPT for dental medicine. *J Esthet Restor Dent*. 2023;35:1098–102. <https://doi.org/10.1111/jerd.13046>.
19. Fatani B. ChatGPT for Future Medical and Dental Research. *Cureus*. 2023;15:e37285. <https://doi.org/10.7759/cureus.37285>.
20. Arslan S. Exploring the potential of Chat GPT in personalized obesity treatment. *Ann Biomed Eng*. 2023;51:1887–8. <https://doi.org/10.1007/s10439-023-03227-9>.
21. Biswas SS. Role of Chat GPT in Public Health. *Ann Biomed Eng*. 2023;51:868–9. <https://doi.org/10.1007/s10439-023-03172-7>.
22. Huang H, Zheng O, Wang D, Yin J, Wang Z, Ding S, Yin H, Xu C, Yang R, Zheng Q, Shi B. ChatGPT for shaping the future of dentistry: the potential of multi-modal large language model. *Int J Oral Sci*. 2023;15:29. <https://doi.org/10.1038/s41368-023-00239-y>.
23. Ali K, Barhom N, Tamimi F, Duggal M. ChatGPT—A double-edged sword for healthcare education? Implications for assessments of dental students. *Eur J Dent Educ*. 2024;28:206–11. <https://doi.org/10.1111/eje.12937>.
24. Lahat A, Klang E. Can advanced technologies help address the global increase in demand for specialized medical care and improve telehealth services? *J Telemed Telecare*. 2023;1357633X231155520. <https://doi.org/10.1177/1357633X231155520>.
25. Babayigit O, Tastan Eroglu Z, Ozkan Sen D, Ucan Yarkac F. Potential use of ChatGPT for Patient Information in Periodontology: a descriptive pilot study. *Cureus*. 2023;15:e48518. <https://doi.org/10.7759/cureus.48518>.
26. Bates MN, Fawcett J, Garrett N, Cutress T, Kjellstrom T. Health effects of dental amalgam exposure: a retrospective cohort study. *Int J Epidemiol*. 2004;33:894–902. <https://doi.org/10.1093/ije/dyh164>.
27. Issa Y, Brunton PA, Glenny AM, Duxbury AJ. Healing of oral lichenoid lesions after replacing amalgam restorations: a systematic review. *Oral Surg Oral Med Oral Pathol Oral Radiol Endod*. 2004;98:553–65. <https://doi.org/10.1016/j.tripleo.2003.12.027>.
28. Björkman L, Musial F, Alraek T, Werner EL, Weidenhammer W, Hamre HJ. Removal of dental amalgam restorations in patients with health complaints attributed to amalgam: a prospective cohort study. *J Oral Rehabil*. 2020;47:1422–34. <https://doi.org/10.1111/joor.13080>.
29. Shen Y, Heacock L, Elias J, Hentel KD, Reig B, Shih G, Moy L. ChatGPT and other large Language models are double-edged swords. *Radiology*. 2023;307:e230163. <https://doi.org/10.1148/radiol.230163>.
30. Information for Patients About Dental Amalgam Fillings. [cited 2023 July 5] <https://www.fda.gov/medical-devices/dental-amalgam-fillings/information-patients-about-dental-amalgam-fillings>.
31. ChatGPT App. 2023; [cited 2023 July 5] <https://openai.com/blog/chatgpt>.
32. Feldman R. Techniques and applications for sentiment analysis. *Commun ACM* Apr. 2013;56:82–9.
33. Lossio-Ventura JA, Weger R, Lee AY, Guinee EP, Chung J, Atlas L, Linos E, Pereira. Against Widely Used Sentiment Analysis Tools: Sentiment Analysis of COVID-19 Survey Data. *JMIR Ment Health*. 2024;11:e50150. FA Comparison of ChatGPT and Fine-Tuned Open Pre-Trained Transformers (OPT).
34. Li J, Dada A, Puladi B, Kleesiek J, Egger J. ChatGPT in healthcare: a taxonomy and systematic review. *Comput Methods Programs Biomed*. 2024;245:108013. <https://doi.org/10.1016/j.cmpb.2024.108013>.
35. Kücüç D, Arici N. Deep learning-based sentiment and stance analysis of Tweets about Vaccination. *Int J Semant Web Inf Syst*. 2023;19:1–18. <https://doi.org/10.4018/IJWSWS.333865>.
36. Rijcken E, Scheepers F, Zervanov K, Spruit M, Mosteiro P, Kaymak U. Towards Interpreting Topic Models with ChatGPT. In: Paper presented at The 20th World Congress of the International Fuzzy Systems Association, Daegu, Republic of Korea. 2023. <https://research.tue.nl/en/publications/towards-interpreting-topic-models-with-chatgpt>. Accessed 18 Apr 2024.
37. Praveen SV, Vijaya S. Examining otolaryngologists' attitudes towards large language models (LLMs) such as ChatGPT: a comprehensive deep learning analysis. *Eur Arch Otorhinolaryngol*. 2024;281:1061–3. <https://doi.org/10.1007/s00405-023-08325-x>.
38. Fütterer T, Fischer C, Alekseeva A, Chen X, Tate T, Warschauer M, Gerjets P. ChatGPT in education: global reactions to AI innovations. *Sci Rep*. 2023;13:15310. <https://doi.org/10.1038/s41598-023-42227-6>.
39. Luo Z, Xie Q, Ananiadou S. ChatGPT as a factual inconsistency evaluator for Abstractive text summarization. *ArXiv Abs*. 2023;15621. <https://doi.org/10.48550/arXiv.2303.15621>.
40. Mayank S, Wade V. Comparing Abstractive Summaries generated by ChatGPT to Real Summaries through Blinded reviewers and text classification algorithms. *ArXiv Abs*. 2023;17650. <https://doi.org/10.48550/arXiv.2303.17650>.
41. Kocoń J, Cichecki I, Kaszyca O, Kochanek M, Szydło D, Baran J, Bielaniec J, Gruza M, Janz A, Kanclerz K, Kocoń A, Koptyra B, Mieleśczenko-Kowszewicz W, Milkowski P, Oleksy M, Piasecki M, Radlński L, Wojtasik K, Woźniak S, Kazienko P. ChatGPT: Jack of all trades, master of none. *Inf Fusion*. 2023;99:101861. <https://doi.org/10.1016/j.inffus.2023.101861>.
42. Oh S, Yi YJ, Worrall A. Quality of health answers in social Q&A. *Proc Am Soc Info Sci Tech*. 2012;49:1–6. <https://doi.org/10.1002/meet.14504901075>.
43. Johnson D, Goodman R, Patrinely J, Stone C, Zimmerman E, Donald R, Chang S, Berkowitz S, Finn A, Jahangir E, Scoville E, Reese T, Friedman D, Bastarache J, van der Heijden Y, Wright J, Carter N, Alexander M, Choe J, Chastain C, Zic J, Horst S, Turker I, Agarwal R, Osmundson E, Idrees K, Kieman C, Padmanabhan C, Bailey C, Schlegel C, Chambliss L, Gibson M, Osterman T, Wheelless L. Assessing the accuracy and reliability of AI-Generated medical

- responses: an evaluation of the Chat-GPT model. Res sq [Preprint]. 2023 Feb 28;rs.3.rs-2566942. <https://doi.org/10.21203/rs.3.rs-2566942/v1>.
44. Wei Q, Yao Z, Cui Y, Wei B, Jin Z, Xu X. Evaluation of ChatGPT-generated medical responses: a systematic review and meta-analysis. *J Biomed Inf*. 2024;151:104620. <https://doi.org/10.1016/j.jbi.2024.104620>.
 45. Hulman A, Døllerup OL, Mortensen JF, Fenech ME, Norman K, Støvring H, Hansen TK. ChatGPT- versus human-generated answers to frequently asked questions about diabetes: a turing test-inspired survey among employees of a Danish diabetes center. *PLoS ONE*. 2023;18:e0290773. <https://doi.org/10.1371/journal.pone.0290773>.
 46. Gregorcic B, Pendrill AM. ChatGPT and the frustrated Socrates. *Phys Educ*. 2023;58:035021. <https://doi.org/10.1088/1361-6552/acc299>.
 47. Amaro I, Della Greca A, Francese R, Tortora G, Tucci C. AI unreliable answers: A case study on ChatGPT. In: International Conference on Human-Computer Interaction. Switzerland: Springer Nature; 2023. pp. 23–40.
 48. Rao A, Pang M, Kim J, Kamineneni M, Lie W, Prasad AK, Landman A, Dreyer K, Succi MD. Assessing the utility of ChatGPT throughout the entire clinical workflow: development and usability study. *J Med Internet Res*. 2023;25:e48659. <https://doi.org/10.2196/48659>.
 49. Walker HL, Ghani S, Kuemmerli C, Nebiker CA, Müller BP, Raptis DA, Staubli SM. Reliability of Medical Information provided by ChatGPT: Assessment Against Clinical Guidelines and Patient Information Quality Instrument. *J Med Internet Res*. 2023;25:e47479. <https://doi.org/10.2196/47479>.
 50. Gunawan J. Exploring the future of nursing: insights from the ChatGPT model. *Belitung Nurs J*. 2023;9:1–5. <https://doi.org/10.33546/bnj.2551>.
 51. Mijwil M, Mohammad A, Ahmed HA. ChatGPT: exploring the role of Cybersecurity in the Protection of Medical Information. *Mesopotamian J Cybersecur*. 2023;18–21. <https://doi.org/10.58496/MJCS/2023/004>.
 52. Bagde H, Dhopte A, Alam MK, Basri R. A systematic review and meta-analysis on ChatGPT and its utilization in medical and dental research. *Heliyon*. 2023;9:e23050. <https://doi.org/10.1016/j.heliyon.2023.e23050>.
 53. Suárez A, Díaz-Flores García V, Algar J, Gómez Sánchez M, Llorente de Pedro M, Freire Y. Unveiling the ChatGPT phenomenon: evaluating the consistency and accuracy of endodontic question answers. *Int Endod J*. 2024;57:108–13. <https://doi.org/10.1111/iej.13985>.
 54. Freire Y, Santamaría Laorden A, Orejas Pérez J, Gómez Sánchez M, Díaz-Flores García V, Suárez A. ChatGPT performance in prosthodontics: Assessment of accuracy and repeatability in answer generation. *J Prosthet Dent*. 2024;131. <https://doi.org/10.1016/j.prosdent.2024.01.018>. :659.e1-659.e6.
 55. Vaira LA, Lechien JR, Abbate V, Allevi F, Audino G, Beltraminini GA, Bergonzani M, Bolzoni A, Comitteri U, Crimi S, Gabriele G, Lonardi F, Maglificio F, Petrocelli M, Pucci R, Saponaro G, Tel A, Vellone V, Chiesa-Estomba CM, Boscolo-Rizzo P, Salzano G, De Riu G. Accuracy of ChatGPT-Generated Information on Head and Neck and Oromaxillofacial Surgery: A Multicenter Collaborative Analysis. *Otolaryngol Head Neck Surg*. 2023 Aug 18. Epub ahead of print. <https://doi.org/10.1002/ohn.489>.
 56. Hatia A, Doldo T, Parrini S, Chisci E, Cipriani L, Montagna L, Lagana G, Guenza G, Agosta E, Vinjulli F, Hoxha M, D'Amelio C, Favaretto N, Chisci G. Accuracy and completeness of ChatGPT-Generated information on interceptive orthodontics: a Multicenter Collaborative Study. *J Clin Med*. 2024;13:735. <https://doi.org/10.3390/jcm13030735>.
 57. Kılınc DD, Mansız D. Examination of the reliability and readability of Chatbot Generative Pretrained Transformer's (ChatGPT) responses to questions about orthodontics and the evolution of these responses in an updated version. *Am J Orthod Dentofacial Orthop*. 2024;S0889-5406(24)00007–6. Epub ahead of print. <https://doi.org/10.1016/j.ajodo.2023.11.012>.
 58. Mago J, Sharma M. The potential usefulness of ChatGPT in oral and maxillofacial Radiology. *Cureus*. 2023;15:e42133. <https://doi.org/10.7759/cureus.42133>.
 59. Kristoffersen AE, Alræk T, Stub T, Hamre HJ, Björkman L, Musial F. Health complaints attributed to Dental Amalgam: a retrospective survey exploring Perceived Health changes related to amalgam removal. *Open Dent J*. 2016;10:739–51. <https://doi.org/10.2174/1874210601610010739>.
 60. Broadbent JM, Murray CM, Schwass DR, Brosnan M, Brunton PA, Lyons KS, Thomson WM. The Dental Amalgam Phasedown in New Zealand: a 20-year Trend. *Oper Dent*. 2020;45:255–64. <https://doi.org/10.2341/19-024-C>.
 61. Scientific Committee on Emerging and Newly Identified Health Risks. 2015; [cited 2023 July 5] The safety of dental amalgam and alternative dental restoration materials for patients and users. Brussels (Belgium): European Commission. https://ec.europa.eu/health/scientific_committees/emerging/docs/scenihr_o_046.pdf.
 62. Agency for Toxic Substance and Disease Registry, Public Health Service. 1999; [cited 2023 July 5] Toxicological profile for mercury. Atlanta (GA): US Department of Health and Human Services. <https://www.atsdr.cdc.gov/toxprofiles/tp46.pdf>.
 63. National Center for Toxicological Research, US Food and Drug Administration. 2023; [cited 2023 July 5] White paper: FDA update/review of potential adverse health risks associated with exposure to mercury in dental amalgam. Jefferson (AR): US Department of Health and Human Services. <https://www.fda.gov/medical-devices/dental-amalgam-fillings/white-paper-fda-updatereview-potential-adverse-health-risks-associated-exposure-mercury-dental>.
 64. Sjørusen TT, Binder PE, Lygre GB, Helland V, Dalen K, Björkman L. Patients' experiences of changes in health complaints before, during, and after removal of dental amalgam. *Int J Qual Stud Health Well-being*. 2015;10:28157. <https://doi.org/10.3402/qhw.v10.28157>.
 65. Alkaissi H, McFarlane SI. Artificial Hallucinations in ChatGPT: implications in Scientific writing. *Cureus*. 2023;15:e35179. <https://doi.org/10.7759/cureus.35179>.
 66. Ollivier M, Pareek A, Dahmen J, Kayaalp ME, Winkler PW, Hirschmann MT, Karlsson J. A deeper dive into ChatGPT: history, use and future perspectives for orthopaedic research. *Knee Surg Sports Traumatol Arthrosc*. 2023;31:1190–2. <https://doi.org/10.1007/s00167-023-07372-5>.
 67. Sezer B, Giritlioğlu B, Siddikoğlu D, Lussi A, Kargül B. Relationship between erosive tooth wear and possible etiological factors among dental students. *Clin Oral Investig*. 2022;26:4229–38. <https://doi.org/10.1007/s00784-022-04425-w>.
 68. Kumar G, Rehman F, Al-Muzian L, Farsi D, Hiremath S. Global Scenario of Teledentistry during COVID-19 pandemic: an insight. *Int J Clin Pediatr Dent*. 2021;14:426–9. <https://doi.org/10.5005/jp-journals-10005-1952>.
 69. Hesse BW, Nelson DE, Kreps GL, Croyle RT, Arora NK, Rimer BK, Viswanath K. Trust and sources of health information: the impact of the internet and its implications for health care providers: findings from the first Health Information National trends Survey. *Arch Intern Med*. 2005;165:2618–24. <https://doi.org/10.1001/archinte.165.22.2618>.
 70. Hanna K, Sambrook P, Armfield JM, Brennan DS. Internet use, online information seeking and knowledge among third molar patients attending public dental services. *Aust Dent J*. 2017;62:323–30. <https://doi.org/10.1111/adj.12509>.
 71. Cheng K, Li Z, He Y, Guo Q, Lu Y, Gu S, Wu H. Potential use of Artificial Intelligence in Infectious Disease: take ChatGPT as an Example. *Ann Biomed Eng*. 2023;51:1130–5. <https://doi.org/10.1007/s10439-023-03203-3>.
 72. Buldur M, Sezer B. Can Artificial Intelligence effectively respond to frequently asked questions about fluoride usage and effects? A qualitative study on ChatGPT. *Fluoride – Q*. 2023;56:201–16.
 73. Krathwohl DR. A revision of Bloom's taxonomy: an overview. *Theory into Pract*. 2010;41:212–8.
 74. Herrmann-Werner A, Festl-Wietek T, Holderried F, Herschbach L, Griewatz J, Masters K, Zipfel S, Mahling M. Assessing ChatGPT's mastery of Bloom's taxonomy using psychosomatic medicine exam questions: mixed-methods study. *J Med Internet Res*. 2024;26:e52113. <https://doi.org/10.2196/52113>.
 75. Mitrovic S, Andreoletti D, Ayoub O. ChatGPT or Human? Detect and explain. Explaining decisions of machine learning model for detecting short ChatGPT-generated text. *ArXiv Abs*. 2023;13852. <https://doi.org/10.48550/arXiv.2301.13852>.
 76. Ferrara E. Should ChatGPT be biased? Challenges and risks of Bias in large Language models. *ArXiv Abs*. 2023;03738. <https://doi.org/10.48550/arXiv.2304.03738>.
 77. Sharma N, Liao QV, Xiao Z. Generative Echo Chamber? Effects of LLM-Powered Search systems on Diverse Information seeking. *ArXiv Abs*. 2024;05880. <https://doi.org/10.48550/arXiv.2402.05880>.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.