The Saudi Dental Journal xxx (xxxx) xxx

Contents lists available at ScienceDirect

The Saudi Dental Journal

journal homepage: www.ksu.edu.sa
www.sciencedirect.com

Original Article

Assessing the quality of AI information from ChatGPT regarding oral surgery, preventive dentistry, and oral cancer: An exploration study

Arwa A. Alsayed [a,*], Mariam B. Aldajani [b], Marwan H. Aljohani [c], Hamdan Alamri [d], Maram A. Alwadi [e], Bodor Z. Alshammari [f], Falah R. Alshammari [g,*]

[a] Sijam Dental Centre, Riyadh, Saudi Arabia
[b] Department of Paediatric Dentistry, Faculty of Dentistry, King Abdulaziz University, Jeddah, Saudi Arabia
[c] Oral and Maxillofacial Diagnostic Sciences, College of Dentistry, Taibah University, Madinah city, Saudi Arabia
[d] Department of Preventive Dentistry, College of Dentistry, Majmaah University, Al Majmaah, Saudi Arabia
[e] Department of Dental Health, College of Applied Medical Sciences, King Saud University, Riyadh, Saudi Arabia
[f] Ministry of Health, Qassim, Saudi Arabia
[g] College of Dentistry, University of Ha'il, Ha'il city, Saudi Arabia

ARTICLE INFO

Keywords:
ChatGPT
Oral surgery
Preventive dentistry
Oral cancer

ABSTRACT

Aim: Evaluation of the quality of dental information produced by the ChatGPT artificial intelligence language model within the context of oral surgery, preventive dentistry, and oral cancer.
Methodology: This study adopted quantitative methods approach. The experts prepared 50 questions (including dimensions of, risk factors, preventive measures, diagnostic methods, and treatment options) that would be presented to ChatGPT, and its responses were rated for their accuracy, completeness, relevance, clarity or comprehensibility, and possible risks using a standardized rubric. To carry out the assessment of the responses by ChatGPT, a standardized scoring rubric was used. Evaluation process included feedback concerning the strengths, weaknesses, and potential areas of improvement in the responses provided by ChatGPT.
Results: While achieving the highest score for preventive dentistry at 4.3/5 and being able to communicate the complex information coherently, the tool showed lower accuracy for oral surgery and oral cancer, scoring 3.9/5 and 3.6/5, respectively, with several gaps for post-operative instructions, personalized risk assessments, and specialized diagnostic methods. Potential risks, such as a lack of individualized advice, were shown in 53% of the oral cancer and in 40% of the oral surgery. While showing promise in some domains, ChatGPT had important limitations in specialized areas that require nuanced expertise.
Conclusion: The findings point to the need for professional supervision while using AI-generated information and ongoing evaluation as capabilities evolve, for the assurance of responsible implementation in the best interest of patient care.

1. Introduction

Oral surgery involves a variety of procedures crucial for the maintenance and restoration of health in the oral cavity (Aghiorghiesei et al. 2022). According to Alauddin et al. (2021), these procedures range from simple, everyday cases of tooth extraction and implant placement to highly complex reconstructive surgery of the jaws and other associated structures (Alauddin, Baharuddin, and Ghazali 2021). One of the most common oral surgical procedures is a tooth extraction. This may be for a variety of reasons: gross tooth decay, severe periodontal diseases, impacted third molars, and orthodontic treatment (Lindahl and Ventä 2023). Proper extraction techniques and postoperative care are mandatory for the reduction of complications and the achievement of optimal healing.

Other more complex oral surgeries include orthognathic surgery, repositioning of the jawbones for skeletal deformity correction and improved bite alignment. Ngọc et al. (2016) showed that orthognathic surgeries are frequently performed with orthodontists for maximum function and aesthetic results (Ngọc et al. 2016). Other complex operations, as shown by Neville et al. (2023), include the surgical correction

* Corresponding authors.
E-mail addresses: difp_as@yahoo.com (A.A. Alsayed), mariam.aldajani@gmail.com (M.B. Aldajani), al_marwan_y@hotmail.com (M.H. Aljohani), hamdanamri1@gmail.com (H. Alamri), Malwadi@ksu.edu.sa (M.A. Alwadi), bodor.z.alshammari@gmail.com (B.Z. Alshammari), falah_R_M_SH@hotmail.com (F.R. Alshammari).

https://doi.org/10.1016/j.sdentj.2024.09.009
Received 13 June 2024; Received in revised form 7 September 2024; Accepted 8 September 2024
Available online 12 September 2024

of cleft lip and palate, the management of TMJ disorders, and the removal of pathological lesions or tumors within the mouth (Neville et al. 2023). Accurate and current information on these procedures is very essential for patient education, informed consent, and successful treatment outcomes. Patients need to be adequately informed of the risks, benefits, and possible complications of every procedure so they can make sound decisions concerning their care.

Preventive dentistry aims at optimal oral health maintenance to prevent the development or progression of dental diseases (Birch et al. 2015). In this way, complex and costly treatments are reduced, and general well-being is promoted. Birch et al. (2015) identified that regular dental check-ups and professional cleanings are cornerstones of preventive dentistry (Birch et al. 2015). Dental professionals can use these appointments to identify and address any potential issues at their early stages, such as cavities, gum disease, or oral lesions. In many cases, early treatment is able to stop further progression of the problem and often more extensive treatment. Fluoride treatment, according to Ten Cate (2013), is the main essence of preventive dentistry (Ten Cate 2013). This is because fluoride is a strengthening element to the enamel of the tooth, making it more resistant to acid and decay. Professional fluoride applications, in association with the application of fluoridated toothpaste and fluoridated water supplies, can significantly reduce the risk of caries in children and those people with a high caries risk (Carey 2014).

Another important area of preventive dentistry is dietary advice. It has been realized that the reduction of sugar and acid intake will be able to provide a healthy environment in the mouth, which in turn will not only reduce the formation of caries on teeth but also help with erosion (Jevtić et al. 2015). Dental professionals can give tailored advice on how to ensure a balanced diet that will help in optimal oral health.

Preventive care includes more than just routine care and includes the application of dental sealants as a protective barrier to cavities in the grooves and fissures of the teeth. Other key constituents of preventive dentistry are education in oral hygiene and promotion of good brushing and flossing habits, which allow a person to maintain their own oral hygiene (Alshammari et al. 2022; Wells 2019).

Oral cancer goes beyond routine care to encompass the application of dental sealants, which accord some kind of shield from cavities against the grooves and fissures of teeth. Also, oral hygiene education and encouragement of good brushing and flossing habits will be required inputs of preventive dentistry to empower the patient to take active participation in oral health maintenance. Oral cancer, though much rarer than some other kinds of cancer, may have devastating consequences if not diagnosed and treated early (Wells 2019). In addition, a narrative review claimed that early diagnosis and early intervention drastically improve survival rates and reduce the long-term impact of the disease (Abati et al., 2020). On the other hand, early diagnosis and early intervention drastically improve survival rates and reduce the long-term impact of the disease (Abati et al. 2020). The risk factors for oral cancer include the use of tobacco and alcohol, which make up a high proportion of the cases (Tenore et al. 2020). However, Aghiorghiesei et al. (2022) established that other risk factors, such as human papillomavirus infection, chronic inflammation, and genetic predispositions, increase the chances of a person suffering from oral cancer (Aghiorghiesei et al. 2022).

Early signs and symptoms of oral cancer include persistent sores, lumps, or thickened areas in the mouth or throat, difficulty in swallowing, or unexplained bleeding. According to Hadziabdic et al. (2017), dental professionals are on the front line in conducting regular oral cancer screening during routine check-ups for the purpose of identifying patients at an early stage and referring them for further diagnosis and management (Hadziabdic, Sulejmanagic, and Kurtovic-Kozaric 2017). Diagnostic tools in oral cancer include visual and tactile examinations; specialized procedures also include brush biopsies and molecular testing. Other advanced imaging techniques include CT scans and MRI, which may also help to assess the extent of the disease and assist in planning the process of treatment (Carreras-Torras and Gay-Escoda 2015).

Akbulut and Altan (2018) identified the fact that treatment for oral cancer generally employs an interdisciplinary approach that will generally include surgery, radiation therapy, chemotherapy, and targeted therapies (Akbulut et al. 2018). The precise treatment option is decided by the case, the stage and location of the cancer, the general health of the patient, and personal choices. Supportive care and rehabilitation services are equally essential to the management of the physical and psychological impacts of oral cancer (Akbulut et al. 2018). Akbulut and Altan, (2018) further indicated that speech therapy, nutritional counseling, and mental health support can be given to a patient to deal with difficulties arising from this disease and help ensure a quality life (Akbulut et al. 2018).

AI is revolutionizing oral surgery, preventative dentistry, and cancer detection. AI technologies, such as machine learning algorithms and deep learning models, are improving diagnosis accuracy, treatment planning, and patient outcomes. In dentistry, AI can help diagnose oral disorders early, such as mouth cancer, by analyzing imaging data, patient records, and clinical symptoms. It can also enhance patient care by providing risk assessment, treatment suggestions, and personalized care. Large datasets have trained ChatGPT, an AI-powered tool, to provide human-like responses to user inputs, thereby enhancing computational linguistics and communication skills. It provides advice, coaching, and simplified explanations for education, healthcare, and patient care (Gilson et al. 2023; Balel, 2023). These applications are expected to enhance patient empowerment, service efficiency, safety, sustainability, care accessibility, quality, and engagement in healthcare and dentistry. D patient outcomes. In dentistry, AI can help diagnose oral disorders early, such as mouth cancer, by analyzing imaging data, patient records, and clinical symptoms. It can also enhance patient care by providing risk assessment, treatment suggestions, and personalized care. Large datasets have trained ChatGPT, an AI-powered tool, to provide human-like responses to user inputs, thereby enhancing computational linguistics and communication skills. It provides advice, coaching, and simplified explanations for education, healthcare, and patient care. These applications are expected to enhance patient empowerment, service efficiency, safety, sustainability, care accessibility, quality, and engagement in healthcare and dentistry (Eggmann et al. 2023; Strunga et al. 2023).

In recent years, artificial intelligence (AI) has rapidly advanced, and language models like ChatGPT have demonstrated remarkable capabilities in understanding and generating human-like text (Orrù et al. 2023). As these AI systems become more accessible, potential uses in the health sciences have gained significant attention (Sarker 2022). However, the quality and accuracy of information provided by AI systems, particularly in specialized areas like dentistry, need further testing for patient safety and optimum care. As AI and algorithms continue to evolve, it's crucial to continually assess and improve the quality of the information they provide. This research project will focus on assessing the quality of AI-generated information from ChatGPT in three critical areas: oral surgery, preventive dentistry, and oral cancer. In so doing, it will analyze how information given through the AI mode of operation will be accurate, complete, and relevant in clarity and risk or limitations to the client or patient. The study seeks to determine whether AI systems like ChatGPT can be effectively used as dental information for both professionals and patients.

2. Methodology

The current study adopted quantitative methods approach. In assessing the quality of information generated through AI from ChatGPT, a total of 50 questions in the three areas of interest—oral surgery, preventive dentistry, and oral cancer—were developed by six consultants in each field (AR, FR, MJ, HA, MD and MW) and tested via a recent graduated dentist BZ. The development of questions was guided by an extensive literature review and cross-checked against relevant clinical practice guidelines. The questions were meant to encompass as

much information as possible regarding the topics in each area, including procedures used, risk factors, preventive measures, diagnostic methods, and treatment options (Table 1).

For oral surgery, the questions ranged from tooth extractions to dental implants, orthognathic surgeries, and temporomandibular joint (TMJ) treatments. Questions address the preoperative concerns, surgical techniques, postoperative care, and the complications of surgery. In the domain of preventive dentistry, questions related to oral hygiene practices were asked, as well as fluoride treatments, dietary recommendations, risk assessment for dental caries and periodontal disease, and application of preventive measures such as dental sealants. For oral cancer, questions included risk factors: tobacco, alcohol, HPV, genetic predispositions; signs and symptoms; diagnostic methods, including visual examinations, brush biopsies, and imaging techniques; treatment modalities, including surgery, radiation, chemotherapy, targeted therapies; and supportive care services.

The set of 50 questions was prepared to represent topics in each area in a balanced way, considering the complexity and specificity of a given question and exploring its performance. These questions were subsequently put to ChatGPT, and its responses were recorded verbatim, without giving it further prompting or context beyond that provided within the question. This was a way of re-creating a real-life scenario where users would use ChatGPT without providing extensive background information.

To carry out the assessment of ChatGPT's responses, a standardized scoring rubric was developed (Table 2). The rubric was designed to ensure homogeneity and objectivity in the process of estimation and reduce the possibilities of bias or subjective perceptions. To improve reliability in the evaluation process, a minimum of three panel members scored the responses for any question. In the case of inconsistencies regarding the score given by different panel members, a consensus discussion was done to revisit the response, clarify misunderstandings, and come to a final agreed-upon score.

Further, the evaluation process included qualitative analysis through which the panel members provided detailed feedback concerning the strengths, weaknesses, and potential areas of improvement in the responses provided by ChatGPT. Thematic analysis was used for the qualitative aspect of this study. Thematic analysis is a method of analyzing qualitative data that involves identifying, analyzing, and interpreting patterns or themes within the data (Finlay, 2021). This qualitative data also completed the quantitative scores, bringing valuable insight into the specific challenges or limitations that the AI system was likely to face. All through the research, strict protocols to keep all information confidential and ethically handle any information regarding patients that might incidentally be included in the ChatGPT responses were adhered to. The version of ChatGPT that was used in this was OpenAI's GPT-4 between the 1st of February 2024 and the 1st of May 2024.

Accuracy was measured by assessing how well responses aligned with established dental guidelines and best practices, ranging from completely accurate (5) to completely inaccurate (0). Completeness evaluated the thoroughness of responses in covering all relevant aspects of the topic, with scores ranging from comprehensive coverage (5) to no relevant information provided (0). Relevance measured the degree to which responses pertained to the question or topic, with higher scores indicating higher pertinence and lower scores indicating irrelevance. Clarity and Comprehensibility assessed the clarity and understandability of responses for a general audience, with higher scores indicating clear and well-organized responses and lower scores indicating incomprehensibility. Potential Risks/Limitations identified any potential risks or limitations in the responses, such as lack of personalized advice or potential for misinterpretation, with a score of 0 indicating no risks identified and 1 indicating potential risks present.

To calculate the average scores and percentages for each criterion and topic area, individual scores were collected and totaled, then divided by the total number of responses evaluated. The percentage

Table 1
Questions Used to Input in ChatGPT.

Area of interest	The Question
Oral Surgery (15 questions)	1 What are the main indications for a tooth extraction? 2 Describe the surgical procedure for a simple tooth extraction. 3 What pre-operative instructions should be given to a patient before tooth extraction? 4 What are some potential complications that can occur after a tooth extraction? 5 How should a patient care for the extraction site in the days following the procedure? 6 What is the process for getting a dental implant? 7 What are the risks and potential complications of dental implant surgery? 8 How does orthodontic treatment sometimes require oral surgery procedures? 9Explain the orthognathic surgery process for correcting jaw deformities. 10 What are the most common complications following orthognathic surgery? 11 How is the temporomandibular joint (TMJ) surgically treated for disorders? 12 Describe some surgical techniques for cleft lip and palate repair. 13 What oral pathological lesions may require surgical removal? 14 How is tumor surgery in the oral cavity performed? 15 What types of reconstructive surgery may be needed after oral cancer treatment?
Preventive Dentistry (15 questions)	1 Why are regular dental cleanings and check-ups important for prevention? 2 How does fluoride help prevent cavities and what are the main sources? 3 What dietary changes can help reduce the risk of tooth decay and gum disease? 4 Describe proper brushing and flossing techniques for good oral hygiene. 5 At what ages should dental sealants be applied and how do they prevent cavities? 6 How do you assess a patient's risk level for developing dental caries? 7 What factors put a person at higher risk for periodontal (gum) disease? 8 How can pregnant women take extra preventive steps for optimal oral health? 9 What are the main causes of tooth erosion and how can it be prevented? 10 Describe the ideal mouth rinsing routine as part of a preventive regimen. 11 How can fluoride varnishes be incorporated into a prevention plan? 12 What specific oral hygiene practices should diabetic patients follow? 13 How can preventive dental visits help catch oral cancer early? 14 What lifestyle counseling can dentists provide to help prevent oral diseases? 15 How can dental professionals properly assess and treat dental erosion?
Oral Cancer (20 questions)	1 What are the main risk factors for developing oral cancer? 2 How does HPV infection increase the risk of oral cancer? 3 Describe the role genetics and family history play in oral cancer risk. 4 What are the most common early signs and symptoms of oral cancer? 5 How can oral cancer be detected through regular screening exams? 6 Explain the visual inspection and palpation techniques for oral cancer screening. 7 What is the brush biopsy procedure and when is it used for oral cancer diagnosis?

(*continued on next page*)

Table 1 (*continued*)

Area of interest	The Question
	8 Describe the process of taking a tissue biopsy for oral cancer diagnosis. 9 How are imaging tests like CT, MRI and PET used for oral cancer diagnosis and staging? 10 What is the role of the dentist in early detection and referral of suspected oral cancers? 11 How is the stage of oral cancer determined and what factors are considered? 12 What is the typical treatment approach for early stage oral cancers? 13 Explain the surgery procedures used to treat more advanced oral cancers. 14 How is radiation therapy utilized to treat oral cavity cancers? 15 What role does chemotherapy play in the treatment of oral cancers? 16 What are some of the newer targeted therapy options for oral cancer? 17 How can nutrition support services help oral cancer patients during treatment? 18 What dental rehabilitation may be required after oral cancer surgery? 19 How can speech therapy benefit oral cancer patients post-treatment? 20 What psychological support services should be offered to oral cancer patients?

scores were obtained by multiplying the average scores by 20 for each criterion. Identification of potential risks or limitations was based on responses receiving a score of 1 in the "potential risks/limitations" criterion, with the percentage calculated accordingly. The comparison between routine and complex procedures' accuracy was determined by analyzing score distributions, with scores of 4–5 indicating accuracy for routine procedures and approximately 3 for complex procedures. Weakness areas were identified through consensus discussions, categorizing common themes among lower-scoring responses.

To improve reliability in the evaluation process, a minimum of three panel members scored the responses for each question. In cases where there were inconsistencies in the scores given by different panel members, a consensus discussion was held. During these discussions, panel members revisited the response to clarify misunderstandings and reach a final agreed-upon score.

3. Results

The Table 3 provides a comprehensive overview of ChatGPT's performance in providing information related to oral surgery, preventive dentistry, and oral cancer, evaluated across several key criteria.

3.1. Accuracy

The accuracy provided by Chat-GPT on preventive dentistry topics was 4.3 out of 5, with oral surgery at 3.9 and oral cancer at 3.6. This trend presumes that the AI system did better in domains with well-defined guidelines and general recommendations, like measures of prevention and practices of oral hygiene, rather than more complex or specialized domains such as oral surgery and oral cancer, which presented higher challenges and therefore lower accuracy scores.

The accuracy scores for oral surgery differed tremendously between routine procedures, like extractions and implants, and complex surgeries like orthognathic and TMJ surgery. While the former was scored very high, at 4 or 5, the latter averaged about 3, indicative of limits in the ability to give accurate information for complex surgical interventions.

3.2. Completeness

Similarly, the completeness of the responses from ChatGPT followed the same trend, with preventive dentistry receiving the highest score of 4.1/5, followed by oral cancer and oral surgery at 3.3/5 and 3.2/5, respectively. The AI system provided full coverage for the large part of topics on preventive dentistry but showed gaps in addressing the individual risk assessments based on various factors.

In Oral Surgery, the AI did very well at covering the pre-operative considerations and surgical technique for oral surgery. However, it had some deficiencies in giving all the post-operative care instructions and potential complications.

Table 2
Standardized Scoring Rubric.

homogeneity and objectivity	The rubric evaluated responses
Accuracy	5 – Completely accurate information, aligned with established dental guidelines and best practices. 4 – Mostly accurate, with only minor inaccuracies or omissions. 3 – Moderately accurate, with some significant inaccuracies or gaps. 2 – Several major inaccuracies, but some correct information provided. 1 – Largely inaccurate, with very little correct information. 0 – Completely inaccurate information.
Completeness	5 – Comprehensive coverage of all relevant aspects of the topic. 4 – Covers most aspects, with only minor gaps or omissions. 3 – Moderate coverage, but with some significant gaps or missing information. 2 – Covers only a few aspects of the topic, major gaps present. 1 – Minimal information provided, largely incomplete. 0 – No relevant information provided, completely incomplete.
Relevance	5 – Highly relevant and directly applicable to the question/topic. 4 – Mostly relevant, with only minor tangential information. 3 – Moderately relevant, with some off-topic or less pertinent information. 2 – Somewhat relevant, but significant portions are tangential or off-topic. 1 – Minimally relevant, mostly tangential or unrelated information. 0 – Not relevant at all to the question/topic.
Clarity and Comprehensibility	5 – Highly clear, well-organized, and easily understandable for a general audience. 4 – Mostly clear, with only minor areas of confusion or technical jargon. 3 – Moderately clear, but some portions are confusing or overly technical. 2 – Somewhat unclear, significant portions are confusing or too technical. 1 – Largely unclear or incomprehensible for a general audience. 0 – Completely unclear and incomprehensible.
Potential Risks/ Limitations:	0 – No potential risks or limitations identified. 1 – Potential risks or limitations identified (e.g., lack of personalized advice, oversimplification, potential for misinterpretation).This scoring rubric provides a standardized way to evaluate ChatGPT's responses across the different criteria, with clear descriptors for each score level. The rubric allows for a structured and consistent assessment of the AI's performance, minimizing subjectivity and bias in the evaluation process.

Table 3
Summary Data from the Scoring Rubric.

Criteria	Oral Surgery	Preventive Dentistry	Oral Cancer
Accuracy (Average Score)	3.9/5 (78 %)	4.3/5 (86 %)	3.6/5 (72 %)
Completeness (Average Score)	3.2/5 (64 %)	4.1/5 (82 %)	3.3/5 (66 %)
Relevance (Average Score)	4.1/5 (82 %)	4.6/5 (92 %)	4.0/5 (80 %)
Clarity & Comprehensibility (Average Score)	4.5/5 (90 %)	4.7/5 (94 %)	4.4/5 (88 %)
Potential Risks/ Limitations Identified	40 % of questions	20 % of questions	53 % of questions
Accuracy for Routine vs. Complex Procedures	Routine: 4–5 Complex: ~3	Not provided	Not provided
Areas of Weakness	Post-op care, complications	Personalized risk assessment	Less common risk factors, specialized diagnostics

3.3. Relevance

Across all three domains, ChatGPT gave highly relevant responses, with preventive dentistry scoring the highest at 4.6/5, followed closely by oral surgery at 4.1/5 and lastly oral cancer at 4.0/5. This suggests that generally, the AI responses tended to be on point and applicable to the questions asked, though not without occasional tangents or less relevant facts.

3.4. Clarity and comprehensibility

One of the greatest features of ChatGPT was its capacity to connect with dental information clearly and at a level acceptable to the general audience. The clarity and comprehensibility scores were consistently high across all three areas: the highest score was for preventive dentistry, followed by oral surgery and oral cancer.

3.5. Potential risks/limitations

In a high percentage of questions, the panel noted potential risks or limitations related to solely relying on AI-generated information: 53 % for oral cancer and 40 % for oral surgery. These included the fact that advice on a personal level regarding specific cases of patients, potential misinterpretation or oversimplification of procedures or conditions that are complex, and lack of ability to replicate the holistic assessment and guidance of dental professionals.

The identified risks for preventive dentistry were much lower, around 20 %, generally related to the lack of personalized recommendations and inability to fully reproduce the consultation in person.

3.6. Areas of weakness

Table 3 highlights Areas of Weakness of ChatGPT: For oral surgery, weaknesses were related to post-operative care instructions and information on potential complications. In the area of preventive dentistry, the AI system showed signs of weakness related to the production of personalized risk assessments based on individual factors.

For oral cancer, the weaknesses were related to the provision of accurate information with regard to less common risk factors and specialized diagnostic techniques that could be important for the earlier detection and proper management of the condition.

The findings summarize the potential and limitations of AI-generated information for dental care. Whereas ChatGPT excelled in some areas, there were serious gaps and limitations where the subject matter is highly specialized or where the domain itself is relatively new or in constant evolution. Such results caution against the risk of exclusively relying on AI-generated information and demonstrate the need for professional supervision and personalized assessment within critical healthcare environments.

4. Discussion

The findings of this study have thrown up useful insights regarding the role of AI-driven language models, such as ChatGPT, in disseminating dental information and hence the need for critical evaluation and professional oversight. This research will be critical to assess the performance of ChatGPT under the three main domains of dentistry: oral surgery, preventive dentistry, and oral cancer. In this study, ChatGPT was chosen to be examined for several reasons, including advanced language understanding, high-quality responses, accessibility, and usability.

One area of strength demonstrated by ChatGPT was that of preventive dentistry, which scored highly on the criteria of accuracy, completeness, relevance, and clarity. This supports earlier indications that AI systems perform well in areas of well-established guidelines and general recommendations (Sarker 2022). The high scores the preventive dentistry topics such as oral hygiene practices, fluoride treatment, and dietary recommendations elicited showed that ChatGPT could be a resourceful complementary material in patient education and in raising awareness for oral health, as Birch et al. (2015) and Wells (2019) pointed out (Wells 2019; Birch et al. 2015).

However, the weakness of ChatGPT in giving personalized risk assessments from individual factors, very crucial for effective preventive care, has been revealed. This was supported by ten Cate in 2013, who cited that professional consultation and personalized advice from dental professionals are critical in taking into consideration a patient's unique circumstances and medical history (Ten Cate 2013).

ChatGPT showed a difference in performance between the routine and complex oral surgeries. While the app showed adequate and appropriate information for routine procedures such as extractions and implants, accuracy and completeness scores for the AI system were low with complex surgeries like orthognathic and TMJ treatments. This finding concurs with the observations by Lindahl and Ventä (2023) and Ngọc et al. (2016) that this type of surgical intervention does involve a high level of intricacy and specialized knowledge (Ngọc et al. 2016; Lindahl and Ventä 2023).

The findings of the study outlined weaknesses in the ability of ChatGPT to give information on post-surgery care instructions and possible complications, which Carlsson and Omar in 2010 and Neville et al. (2023) considered important in patient education and informed consent (Carlsson and Omar 2010; Neville et al. 2023). This is a limitation for which professional oversight is essential, and dental professionals should supplement information derived from AI sources with their expert opinion and tailor-made instructions.

In the area of oral cancer, ChatGPT performed relatively poor compared to the other two areas, since the three scores—accuracy, completeness—were lower in this area. The finding concurs with the observations of Abati et al. (2020) and Hadziabdic et al. (2017), indicating that early detection and precise information on risk factors, signs and symptoms, and diagnostic methods for proper management of oral cancer are highly important (Hadziabdic, Sulejmanagic, and Kurtovic-Kozaric 2017; Abati et al. 2020).

The study identified some weaknesses in ChatGPT regarding the less common risk factor information, such as HPV infection and genetic predisposition, and the specialized diagnostic techniques like brush biopsies and molecular testing. These areas form the basis for successful early detection and management of oral cancer, as identified by some authors, including Carreras-Torras and Gay-Escoda (2015), and Altan (2018) (Carreras-Torras and Gay-Escoda 2015; Akbulut et al. 2018).

Additionally, for a high number of questions, especially on oral

cancer and oral surgery, the panel felt that risks and limitations for totally AI-generated information might be failure to offer personalized advice relevant to individual patient cases, misuse or oversimplification of involved procedures or conditions, and inability to replicate a comprehensive assessment and guidance given by dental professionals.

While ChatGPT did quite well in some aspects, such as the effective communication of dental information, the study flagged deficiencies in complete reliance on AI-generated information, particularly in those domains that are specialized or evolutionarily fast. This confirms earlier concerns by researchers associated with the risks and limitations of artificial intelligence systems in healthcare (Akbulut et al. 2018; Carreras-Torras and Gay-Escoda 2015).

The performance of these AI systems, such as ChatGPT, is continually improving, and the results from this study present only one moment in time. As these systems are trained on larger datasets and embody new techniques, their capabilities and limitations may change. Continuous research and evaluation will be necessary in order to ensure that AI technologies become a responsible and effective aspect of dental practice, patient education, and general healthcare delivery.

4.1. Strengths and limitations

This comprehensive study employed a robust mixed-methods approach, utilizing a standardized rubric and input from multiple experts to systematically evaluate the quality of dental information provided by ChatGPT across oral surgery, preventive dentistry, and oral cancer. Its strengths lie in the well-designed methodology, thorough analysis covering aspects like accuracy, completeness, relevance, clarity, and potential risks/limitations, and the valuable findings highlighting ChatGPT's strengths in preventive dentistry and communicating clearly, as well as weaknesses in handling complex procedures, personalized risk assessments, and specialized diagnostics.

However, limitations include the narrow scope focused solely on dental information from ChatGPT, potential subjectivity or biases in evaluation, rapidly evolving AI technology necessitating regular re-evaluation, lack of comparison to other AI models or human experts, and limited details on specific question types. Nonetheless, the study underscores the need for professional oversight and responsible integration of AI-generated information in healthcare, particularly specialized domains requiring nuanced expertise.

5. Conclusion

The conclusions raise the need to responsibly and judiciously integrate AI technologies that maximize on their strengths but minimize the potential risks through professional oversight and adherence to established clinical guidelines. Although the performance of ChatGPT was quite good for areas of comparatively well-established practices like preventive dentistry, huge gaps and limitations still persist in specialized domains like oral surgery and oral cancer management. The study strongly suggests that healthcare providers exercise particular care when using AI-generated information, most especially for complex cases or for rapidly evolving fields where nuanced expertise and individualized assessment are most important.

Moreover, with weaknesses having been detected in the full information on post-operative care, potential complications, and individualized risk assessments, it goes to underline the irreplaceable role of human expertise in making critical decisions. Professionals in dentistry have to treat AI systems, including ChatGPT, as complementary tools, capable of enhancing patient education and raising awareness around oral health issues but definitely not replacing professional judgment and personalized care.

6 Objective

Assessing the quality of AI-generated information from ChatGPT in three critical areas: oral surgery, preventive dentistry, and oral cancer. In so doing, it will analyze how information given through the AI mode of operation will be accurate, complete, and relevant in clarity and risk or limitations to the client or patient. The study seeks to determine whether AI systems like ChatGPT can be effectively used as dental information for both professionals and patients.

7 Essence of our approach

The current study adopted quantitative methods approach. In assessing the quality of information generated through AI from ChatGPT, a total of 50 questions in the three areas of interest—oral surgery, preventive dentistry, and oral cancer—were developed by six consultants in each filed (AR, FR, MJ, HA, MD and MW) and tested via recent dentist graduated BZ. The development of questions was guided by an extensive literature review and cross-checked against relevant clinical practice guidelines. The questions were meant to encompass as much information as possible regarding the topics in each area, including procedures used, risk factors, preventive measures, diagnostic methods, and treatment options.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

Special thanks to Prof: Ann. Mari Glenny and Prof: Martian Tickle

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.sdentj.2024.09.009.

References

Abati, S., Bramati, C., Bondi, S., Lissoni, A., Trimarchi, M., 2020. Oral cancer and precancer: a narrative review on the relevance of early diagnosis. Int. J. Environ. Res. Public Health 17 (24), 9160. https://doi.org/10.3390/IJERPH17249160.

Aghiorghiesei, O., Zanoaga, O., Nutu, A., Braicu, C., Septimiu Campian, R., Lucaciu, O., Berindan Neagoe, I., 2022. The world of oral cancer and its risk factors viewed from the aspect of MicroRNA expression patterns. Genes 13 (4), 594. https://doi.org/10.3390/GENES13040594.

Akbulut, N., Altan, A., Akbulut, N., Altan, A., 2018. Early detection and multidisciplinary approach to oral cancer patients. In: Prevention, Detection and Management of Oral Cancer. https://doi.org/10.5772/INTECHOPEN.81126.

Alauddin, M.S., Syukran Baharuddin, A., Mohd Ghazali, M.I., 2021. The modern and digital transformation of oral health care: a mini review. Healthcare 9 (2), 118. https://doi.org/10.3390/HEALTHCARE9020118.

Alshammari, F.R., Alsayed, A.A., Albakry, M., Aljohani, M., Kabbi, H., Alamri, H., 2022. Evidence based recommendations to improve the children oral health in Saudi Arabia. Saudi Dental J. 34 (6), 431–444. https://doi.org/10.1016/J.SDENTJ.2022.06.003.

Balel, Y., 2023. Can ChatGPT be used in oral and maxillofacial surgery? J. Stomatol. Oral Maxillofacial Surg. 154 (5), 101471. https://doi.org/10.1016/j.jormas.2023.101471.

Birch, S., Bridgman, C., Brocklehurst, P., Ellwood, R., Gomez, J., Helgeson, M., Ismail, A., et al., 2015. Prevention in practice - a summary. BMC Oral Health 15 (1), 1–9. https://doi.org/10.1186/1472-6831-15-S1-S12/TABLES/4.

Carey, C.M., 2014. Focus on fluorides: update on the use of fluoride for the prevention of dental caries. J. Evid. Based Dent. Pract. 14 (SUPPL.), 95–102. https://doi.org/10.1016/J.JEBDP.2014.02.004.

Carlsson, G.E., Omar, R., 2010. The future of complete dentures in oral rehabilitation. A critical review. J. Oral Rehabil. 37 (2), 143–156. https://doi.org/10.1111/J.1365-2842.2009.02039.X.

Carreras-Torras, C., Gay-Escoda, C., 2015. Techniques for early diagnosis of oral squamous cell carcinoma: systematic review. Med. Oral Patol. Oral Cir. Bucal 20 (3), e305.

Cate, J.M.T., 2013. Contemporary perspective on the use of fluoride products in caries prevention. Br. Dent. J. 214 (4), 161–167. https://doi.org/10.1038/sj.bdj.2013.162.

Eggmann, F., Weiger, R., Zitzmann, N.U., Blatz, M.B., 2023. Implications of large language models such as ChatGPT for dental medicine. J. Esthet. Restor. Dent. 35 (7) https://doi.org/10.1111/jerd.13046.

Gilson, A., Safranek, C.W., Huang, T., Socrates, V., Chi, L., Taylor, R.A., Chartash, D., 2023. How Does ChatGPT perform on the united states medical licensing examination? The implications of large language models for medical education and knowledge assessment. JMIR Med. Educ. 9 (9) https://doi.org/10.2196/45312 p. e45312. [online].

Hadziabdic, N., Sulejmanagic, H., Kurtovic-Kozaric, A., 2017. The role of general dental practitioners in the detection of early-stage oral malignancies–a review. J. Oral Maxillofacial Surg. Med. Pathol. 29 (4), 363–376. https://doi.org/10.1016/J.AJOMS.2017.03.009.

Jevtić, M., Pantelinaci, J., Ilić, T.J., Petrović, V., Grgić, O., Blazić, L., 2015. The role of nutrition in caries prevention and maintenance of oral health during pregnancy. Med. Pregl. 68 (11–12), 387–393. https://doi.org/10.2298/MPNS1512387J.

Lindahl, O., Ventä, I., 2023. Level of difficulty of tooth extractions among roughly 100,000 procedures in primary care. Clin. Oral Invest. 27 (8), 4513–4520. https://doi.org/10.1007/S00784-023-05073-4.

Neville, B.W., Douglas, D.D., Carl, M.A., Angela, C.C., 2023. Oral and Maxillofacial Pathology.

Ngọc, H., Raucharernporn, S., Kiattavorncharoen, S., Boonsiriseth, K., Wongsirichat, K., 2016. Surgery First Approach in Orthognathic Surgery.

Orrù, G., Piarulli, A., Conversano, C., Gemignani, A., 2023. Human-like problem-solving abilities in large language models using ChatGPT. Front. Artif. Intell. 6 (May), 1199350. https://doi.org/10.3389/FRAI.2023.1199350/BIBTEX.

Sarker, I.H., 2022. AI-based modeling: techniques, applications and research issues towards automation, intelligent and smart systems. SN Comput. Sci. 3 (2), 1–20. https://doi.org/10.1007/S42979-022-01043-X.

Strunga, M., Urban, R., Surovková, J., Thurzo, A., 2023. Artificial intelligence systems assisting in the assessment of the course and retention of orthodontic treatment. Healthcare 11 (5), 683. https://doi.org/10.3390/healthcare11050683.

Tenore, G., Nuvoli, A., Mohsen, A., Cassoni, A., Battisti, A., Terenzi, V., Monaca, M. Della, et al., 2020. Tobacco, alcohol and family history of cancer as risk factors of oral squamous cell carcinoma: case-control retrospective study. Appl. Sci. 10 (11), 3896. https://doi.org/10.3390/APP10113896.

Wells, M.H., 2019. Pit and fissure sealants: scientific and clinical rationale. Pediatr. Dent. (January), 461–481.e2. https://doi.org/10.1016/B978-0-323-60826-8.00033-X.