Oral Oncology Reports 10 (2024) 100225

Contents lists available at ScienceDirect

Oral Oncology Reports

journal homepage: www.journals.elsevier.com/oral-oncology-reports

A review: Exploring the role of ChatGPT in the diagnosis and treatment of oral pathologies

Payal Panwar [1], Shalini Gupta [*]

Department of Oral Pathology & Microbiology, King George's Medical University, Lucknow, 226003, UP, India

ARTICLE INFO

Keywords:
ChatGPT
Oral pathologies
Deep learning
Health care
Artificial intelligence

ABSTRACT

This review article delves into the evolving landscape of artificial intelligence (AI), specifically focusing on the integration of ChatGPT in the domain of oral pathologies. ChatGPT, a language model powered by deep learning, has demonstrated significant potential in providing insights and support in various fields. Seeking the capabilities of ChatGPT, several researches have been explored in the context of ChatGPT's effectiveness in oral pathologies in this review article. In the context of oral health, the model's capabilities and limitations are scrutinized, offering a critical analysis of its current applications and future prospects in different sections of this review article.

1. Introduction

The phrase artificial intelligence (AI) was first introduced in the 1950s, referring to the capability of machines to perform tasks traditionally carried out by the person [1]. Artificial Intelligence is also recognized as a branch of computer science, dedicated to crafting intelligent systems that demonstrate traits akin to human manners and behavior. As humanity enters a new era characterized by noteworthy progress in AI, its incorporation into the field of computational molecular biology presents a substantial potential for advancements in the medical field. In a short period, Artificial Intelligence (AI) has brought about a revolution in the medical and scientific domains, facilitating substantial transformations and the seamless incorporation of therapeutic, diagnostic, and patient care courses [2]. Initially dominated by the evolution of Machine Learning (ML) models, subsequent breakthroughs, including Deep Learning (DL) and technologies like Convolutional Neural Networks (CNN), quickly emerged [3]. These algorithms excel in recognizing features and intricate patterns within photos, such as the detection of cancerous cells, achieving high certainty and thereby diminishing the risk of incorrect diagnosis. Fig. 1 shows the working of the AI model and its ability to accommodate a diverse range of inputs that provide an added advantage in transforming the landscape of medical, dental, and healthcare delivery.

Some other application of AI in prognostic pathology involves employing natural language processing (NLP) algorithms to scrutinize illness records. Large language models (LLMs) are artificial intelligence applications trained on vast volumes of text-based data. To capture the statistical distribution of tokens found in extensive collections of publicly available human-generated texts, generative mathematical models are employed in those programs [4]. These tokens encompass individual features, words, graphemes, and punctuation symbols. These algorithms extract pertinent information, aiding in disease diagnosis by identifying major comorbidities, symptoms, and demographic details from pathology records, empowering diagnosticians to make detailed diagnoses [5].

The effectiveness of Artificial Intelligence in addressing more sophisticated reasoning queries within the field of pathology is contingent upon the complications of the queries and the breadth of the training data the AI-based system has undergone. For straightforward or elementary queries, AI-based systems can furnish precise and pertinent responses simultaneously. For instance, a pathology-trained chatbot may proficiently answer queries about physiology, anatomy, ordinary illness, and their symptoms. However, when confronted with more intricate questions demanding a profound comprehension of healthcare expertise and pathology, AI-based systems may not match the efficacy of human experts. Tasks requiring critical thinking, cognitive skills, and nuanced interpretation may currently surpass the abilities of existing AI-based systems [6]. The Chat Generative Pre-trained Transformer (ChatGPT) is a modern Artificial Intelligence model crafted to produce similar-to-person communication. It achieves this by predicting responses from an extensive repository of publicly unpublished resources, comprising books, articles, and websites up to the year 2021. Refined for communicational tasks through reinforcement learning from personal

* Corresponding author.
E-mail addresses: panwarpayal07@gmail.com (P. Panwar), drshalni@gmail.com (S. Gupta).
[1] Dr. Payal Panwar Junior Resident, should be considered first author.

https://doi.org/10.1016/j.oor.2024.100225
Received 6 February 2024; Accepted 18 February 2024
Available online 11 March 2024

observation, ChatGPT enhances precision and consistency in generating responses [7].

This review seeks to offer a comprehensive examination of ChatGPT's application in oral pathologies, providing insights into its impact on diagnostics, treatment planning, patient education, and the overall advancement of oral healthcare. Through the utilization of natural language understanding and generation capabilities, ChatGPT has the potential to revolutionize how dental professionals engage with patients, access information, and make well-informed decisions within the field of oral health.

Within the confines of this article, we will delve into specific use cases of ChatGPT in oral pathology, scrutinizing its proficiency in analyzing clinical data, generating differential diagnoses, and aiding in the formulation of individualized treatment plans. Furthermore, our exploration will extend to the examination of ChatGPT's role in patient communication and education, assessing its capacity to enhance health literacy and facilitate collaborative decision-making between practitioners and those seeking oral healthcare.

Navigating the diverse dimensions of ChatGPT's integration into oral pathology requires a critical evaluation of the model's strengths, limitations, and ethical considerations. This approach aims to contribute to a nuanced comprehension of the dynamic interplay between AI and oral healthcare, fostering informed dialogues regarding the responsible and effective implementation of state-of-the-art technologies.

2. Methodologies used

In this review article, different research articles have been explored for the effective use of ChatGPT in the treatment of oral pathologies. The researchers have used different data sets and methodologies for their studies which are as follows.

- *Sinha* et al. [8] employed ChatGPT in engaging with a set of 100 more complex reasoning questions. These queries were chosen at random from the institution's repository and were organized based on various systems. The questions were systematically classified into 11 pathology systems, such as general, cardiovascular, or gastrointestinal pathology. These queries were intentionally designed to be of a higher order, necessitating a profound understanding of the subject field. They emphasized grasping primary beliefs and philosophies in preference to relying on rote memorization of actuality. To ensure their quality, an expert pathologist with over 10 years of teaching and research experience validated both the face and text of the queries. The responses to each query were systematically gathered and archived for subsequent analysis. A panel of three expert pathologists assessed these responses using a 0–5 scale, categorizing them within the framework of the Structure of Observed Learning Outcome (SOLO) taxonomy, as shown in Fig. 2. The responses were systematically categorized into five distinct groups, each representing a different level of understanding:
  - *Pre-structural*: Characterized by a lack of comprehension of the job.
  - *Unistructural*: Indicates a limited grasp of the job, focusing on a single aspect.
  - *Multi-structural*: Reflects an understanding of multiple aspects without establishing connections among those.
  - *Relational*: Demonstrates an understanding of links and associations between multiple features.
  - *Extended Abstract*: This represents an extensive and advanced understanding, incorporating abstract and theoretical convictions.

This categorization aimed to assess the depth and complexity of the responses, providing a nuanced evaluation based on the Structure of Observed Learning Outcome (SOLO) taxonomy.

The researchers employed descriptive statistical tests to present the data, reporting figures such as numbers, means, medians, standard deviations, and first and third quartiles. The data distribution was assessed using the Shapiro-Wilk test, revealing a departure from normality. To assess response accuracy, a one-sample median test with hypothetical anticipated results was conducted, comparing observed scores to a hypothetical value of 4. The median score is assumed 4 in the case of statistically insignificant different values.

Scores, categorized by diverse pathology systems, underwent analysis using Friedman's test with a posthoc assessment. The agreement among the 3 raters was assessed using the intraclass correlation coefficient (ICC). Classificational data, with frequencies less than five in a category, were collated using Fisher's exact assessment [9]. Statistical inspections were carried out using GraphPad Prism 7, and a significance level of p-value <0.05 was considered statistically significant.

- *Albagieh* et al. [10] performed a comparative inspection of answers from 3rd- and 4th-year inhabitants undergoing training in Oral Pathology and Medicine at King Saud University, College of Dentistry. A close-ended MCQ test, comprising 19 queries with 4 options labeled A to D and one query with 5 options labeled A to E, was administered to the inhabitants. The MCQ was conducted through Google Forms, and individual responses were electronically recorded in a spreadsheet.

To evaluate the inhabitants' performance, their responses were juxtaposed with responses produced by 3 primary language models:

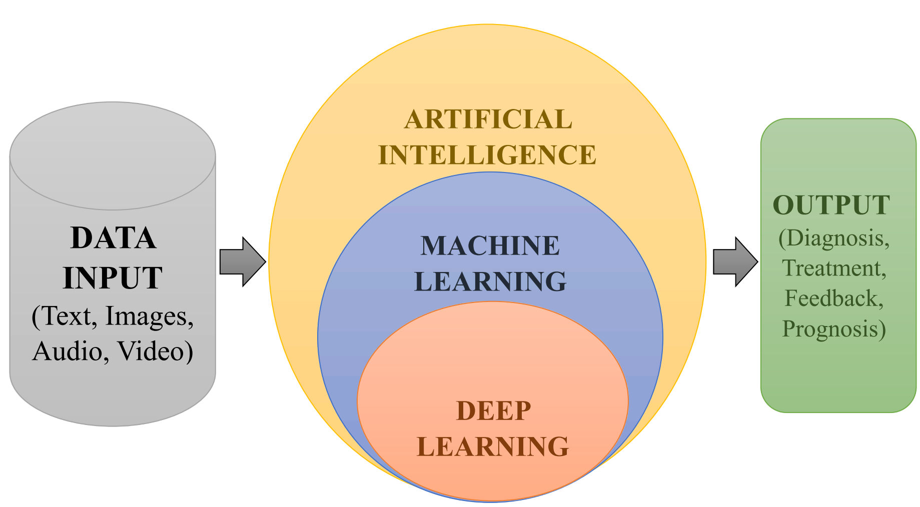

Fig. 1. Working of artificial intelligence model.

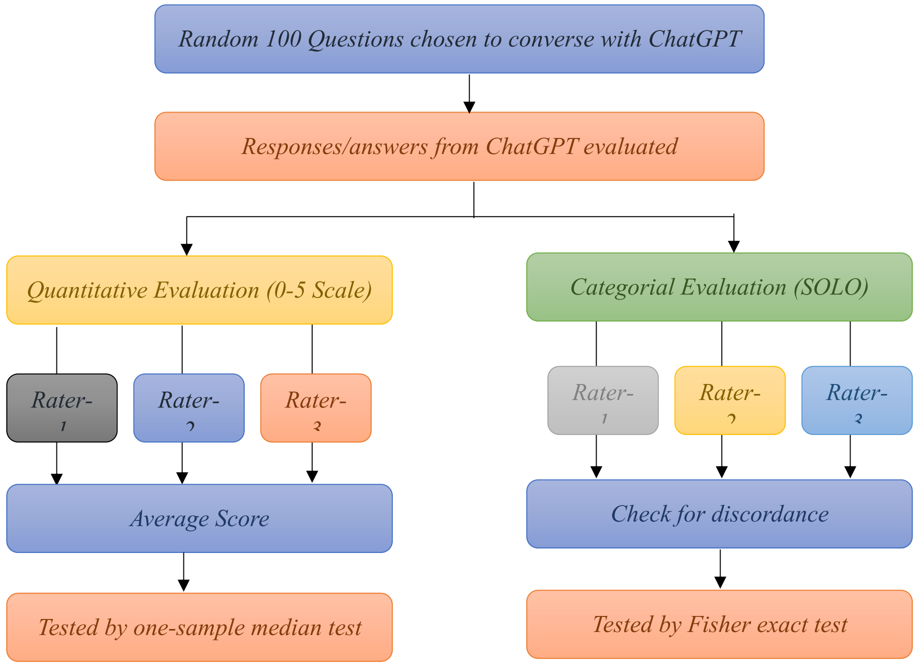

Fig. 2. Brief study method flow-chart used.

PopAI, Stablediffusion, and OpenAI. The queries were fed into the language models in a similar set out as the actual assessment, and to minimize memory retention bias, an AI-based chat session was initiated before every query. This process took place on Nov. 19, 2023, coinciding with the day of the formal MCQ test administration.

The research encompassed a specimen proportion of 20 inhabitants specializing in Oral Pathology and Medicine at King Saud University, College of Dentistry, including both 3rd-year and 4th-year inhabitants.

The statistical analysis was conducted utilizing RStudio. Categorical variables were presented using frequencies and percentages. Specifically centered on clinical case queries, a scoring system was established, assigning a value of 1 for right responses and 0 for wrong ones. An all-inclusive knowledge grade was then calculated by adding up the right response values from both inhabitants and Language Models (LLMs). For every subject (inhabitant or LLM), the all-inclusive grade was scaled from 0 to 20, with excessive scores indicating a greater knowledge level and more right answers. To aid the explanation, raw scores were normalized to a percentage score using:

p_grade = (r_grade * 100)/20

Where p_ grade is the percentage score and r_ grade represents the raw score.

The accomplishment of LLMs and inhabitants was compared by analyzing the p_grades through a Wilcoxon rank-sum examination. Median and interquartile ranges (IQRs) were used to express the scores. To further explore the differences between LLMs and residents, Fisher's precise test was employed. Additionally, a response agreement investigation was conducted using Light's Kappa as a variable of interrater agreement between the various subjects.

- The dataset comprising questions was curated by *Kuscu* et al. [11] from inquiries frequently posed by reputable professional organizations and communities, including the National Cancer Institute, the Medline Plus Medical Encyclopaedia, and the American Head & Neck Society (AHNS). To foster universality and patient depiction, queries sourced from patient recourse clusters and internet community posts were also integrated. A meticulous screening process, conducted by 3 writers, was employed to assess the appropriateness of these questions for inclusion in the study.

A total of 154 queries were selected to elicit replies from ChatGPT. These queries underwent systematic categorization into separate groups based on their respective subjects, encompassing: (i) basic knowledge, (ii) prognosis, (iii) therapeutics, (iv) healing, operative complications, risks, and consequentiality, (v) cancer eradication, and (vi) other. To ensure clarity and precision, certain questions underwent grammatical adjustments. The exactness and reproducibility of replies to queries were evaluated by 2 skillful head and neck surgeons actively engaged in academic enactment. Each surgeon conducted an independent review and provided individual grades, contributing to a thorough and unbiased assessment process. This dual-review approach aimed to enhance the reliability and credibility of the evaluation, ensuring a robust analysis of the responses from ChatGPT. The reviewers assessed the correctness of replies based on the following order.

- *Correct/Comprehensive*: The information is precise and thorough; an H&N surgeon would not need to offer additional details if queried by a patient.
- *Partially Correct/Incomplete*: The presented data is accurate, but supplementary relevant information could be offered upon inquiry by a patient.
- *Mixed*: The dataset contains both accurate and inaccurate information, posing a potential risk of being misleading.
- *Misleading/Completely Inaccurate*: The provided data is entirely incorrect or irrelevant. Reproducibility was evaluated by examining the steadiness of the 2 replies to every query. If the replies were alike, the ChatGPT's initial reply was assessed. In instances where the replies varied, replies were separately scored by the reviewers. In cases where there were discrepancies in the scores, the replies were considered irreproducible. Any inconsistencies in the correctness and reproducibility of replies between the 2 reviewers were examined and settled by a 3rd observer, a highly skilled scholastic H&N surgeon. This third reviewer was rendered unsighted to the initial assessments. Reproducibility was assessed for every query classification and compared across classifications.

The distribution of scores among the replies was computed and expressed as percentage scores. Classification variables were scrutinized using both the chi-square and Fisher's exact test. To gauge inter-rater agreement, the kappa statistic was applied, revealing a substantial level of concordance between 1st and 2nd Reviewers (Kappa value 0.657, $p < 0.001$). All statistical computations were carried out utilizing SPSS v.25.0 by IBM.

3. Results analysis and discussions

This section of the study discusses the results achieved by the researchers aforementioned in the methodologies.

- *Sinha* et al. [8] found the all-inclusive median grade as 4.08 (Q1-Q3:4–4.33), which was remarkably lesser than the maximum accomplishable grade of 5 but comparable to a grade of 5. Specifically, in hepatobiliary and nervous system pathology, the grade closely resembled 5, while for the remaining systems, it was akin to 4. The grades across distinct pathology systems displayed no remarkable dissimilarity in the Kruskal-Wallis test ($p = 0.55$). Out of 100 responses, 86 responses were classified as "*relational*", 12 were classified as "*multi-structural*", and 2 were categorized as "*pre-structural*" ($p < 0.0001$). The grades produced by the 3 raters exhibited an outstanding level of inter-rater trustability. The ICC was 0.975, between an interval of 0.965–0.983 ($F = 40.26$, $p < 0.0001$) with a 95% confidence.
- Among the 20 questions, *Albagieh* et al. [10] found notable dissimilarities in responses for only 2 queries (10%) in their research. In the case of the diagnosis of ectodermal dysplasia, no Language Models (LLMs) produced righteous replies (0.0%), incorrectly identifying every case as xeroderma pigmentosum. In contrast, 85.0% of superior dental inhabitants accurately recognized the prognosis ($p = 0.011$). Another notable performance difference occurred in the administration of a huge tender ulcer post-kidney transplant, where all 3 LLMs advocated topical corticosteroid (100%), while 70.0% of superior dental inhabitants rightly leaned towards intralesional corticosteroid injection ($p = 0.022$). For the persisting queries, no notable dissimilarities were observed in the segments of right answers between LLMs and inhabitants. Based on percentage knowledge grades, the median (IQR) grade of LLMs was 50.0 (45.0–60.0), with a least of 40 (for Stablediffusion) and a highest of 70 (for OpenAI). The median (IQR) grade of superior inhabitants was 65.0 (55.0–75.0), with the greatest and least inhabitant grades being 40 and 90, accordingly. No notable distinction in the percentage grades of inhabitants and LLMs ($p = 0.211$) was found. The agreement level, computed using the Kappa value, indicated diverse degrees of accordance. The agreement among superior dental inhabitants was low, with a 0.396 Kappa value, mirroring challenges in accomplishing consensus among inhabitants. In contradiction, the agreement among LLMs presented an ordinary level, with a 0.622 Kappa value, proposing an additional connected alignment in replies among the AI models. When collating inhabitants' replies with the replies produced by distinct AI models, consisting of PopAI, Stablediffusion, and OpenAI, the agreement levels were relevantly classified as low, with Kappa values of 0.392, 0.381, and 0.402, respectively.
- A total of 154 inquiries related to Head and Neck Cancer (HNC) were directed to ChatGPT by *Kuscu* et al. [11] in their study. ChatGPT produced "*correct/comprehensive*" replies to 133 out of 154 queries, accounting for 86.4%. Meanwhile, "*incomplete/partially correct*" responses were noted at a rate of 11%, and responses that were "*misleading/mixed with accurate and inaccurate data*" were observed at a rate of 2.6%. Significantly, no occurrences of "*irrelevant/completely inaccurate*" responses were noticed. The queries were classified into distinct groups: basic knowledge (23.4%, 36 queries), prognosis (17.5%, 27 queries), therapeutic (17.5%, 27 queries), healing - operative risks - complications - consequentiality (26%, 40 queries), cancer eradication (6.5%, 10 queries), and other (9.1%, 14 queries). The inquiries related to cancer prevention achieved the topmost rate of "*correct/comprehensive*" replies at 100%. Remarkably, no notable distinctions were observed between the classifications regarding the scores of ChatGPT replies ($p = 0.88$). In terms of overall reproducibility, the model exhibited a rate of 94.1% (145 questions). Reproducibility rates were consistently high, reaching 100% for the classifications of basic knowledge, cancer eradication, and others. However, this rate slightly declined to 88.9% for prognosis, 92.5% for healing - operative risks - complications - follow-up, and 88.9% for therapeutic classifications. Nonetheless, no remarkable distinctions were noticed between the classifications agitating reproducibility ($p = 0.309$).

The results from the studies present that ChatGPT can be used as an advanced tool in oral pathologies as the researchers have found significantly accurate responses for the medical inquiries fed into it.

4. Applications and implications

Malik and Zaheer [12] have presented the applications of AI and ChatGPT for the pathological diagnosis of cancer. They have talked about advanced techniques that can be used in diagnosing and in different processes. Whole Slide Imaging (WSI) facilitates the digital transformation of traditionally analog microscopic figure information with outstanding quantity and standard. The mentioned scanned images can be saved and maintained electronically for future perspectives. The use of digital slides provides numerous advantages, including distant primary prognostic assessments, teleconsultation, virtual learning, enhanced workload efficiency, collective opinions and image investigation, and support for cutting-edge research [13–20]. Below are some of the major employments of artificial intelligence enactment in pathology.

- *Crucial role played by AI in the analysis and description of digitized images*: AI can swiftly and adequately measure digitally transformed photos of Papanicolaou-stained cervical smears, cytological smears, encompassing fine needle aspiration cytology (FNAC), and cytocentrifuged smears of bodily fluids. The AI-based system aids in the categorization of individual cells, assessment of margin status, identification of tumor arrangement patterns, and recognition of the presence of linked invasion into neighboring structures too. Consequently, this amplifies the proficiency to offer suitable differentials for the case, resulting in an additional precise prognosis and accelerated turnaround times [21].
- *Identification and classification of diseases*: Certain researchers have designed a CNN model to classify separate image patches pulled out from H&E-stained lung ADC WSIs as either non-malignant or malignant. This methodology supports tumor recognition and delegates the analysis of spatial distribution, boundary features, and shape in the research of tumors [22].
- *Modeling for predictions and prognosis*: AI is not only instrumental in discovering malignancies but has showcased its efficacy in exploring the intricate facts of neoplasms and forecasting sick-person diagnosis too. The research encompassed the examination of 182 elucidated WSIs of prostatic core biopsies to identify the cribriform ordering of malignant cells. Remarkably, the outcomes were outstanding, demonstrating the 0.82 ROC curve for cribriform pattern recognition [23–25].
- *Automation and efficiency in workflows*: By analyzing digitized images and sick-person information, AI can generate initial discoveries that accelerate the reporting procedure, significantly decreasing the time expended. This, in turn, allows pathologists to dedicate their efforts to more intricate cases requiring in-depth evaluation. It also plays a role in expediting the quick and uncomplicated description of essential information from complicated pathology reports. Additionally, artificial intelligence-powered slide scanners robotize the digital transformation of glass slides, modernizing the saving, retrieving, and measuring of pathology photos [26–28].

- *Integration of data and support for decision-making*: Artificial intelligence can integrate information from diverse origins, including clinical records, laboratory investigations, examination findings, imaging and radiology reports, previous electronic medical history, and cytopathological and histopathological information [29,30].
- *Development and research in pharmaceuticals*: AI aids in discovering novel prospective molecular biomarkers and cancer immunohistochemicals, contributing to targeted treatment and the evolution of medications [31,32].

The incorporation of digitally transformed slides into the pathology workflow, along with modern algorithms and computer-aided prognostic techniques, extends the pathologist's outlook on the farther side of a microscopic slide. This enhances the effective implementation and amalgamation of knowledge, surpassing personal boundaries and restraints. Accordingly, there is a notable possibility for the use of AI, such as ChatGPT, in the cancer pathological prognosis. When presented with optical data from pathological slides, ChatGPT can at first measure the visual characteristics, consisting of the patterns, size, and shape of cell order. It can then support its ML algorithms to recognize patterns indicative of cancer within the slide. The specific procedure and analytics through which ChatGPT contributes to description and pathological prognosis can be outlined as follows.

- *Acquisition and digitization of data*: ChatGPT can offer advice on incorporating digital pathology platforms and AI-driven slide scanners.
- *Preparation of training data*: Diagnosticians can utilize ChatGPT to aid in the righteous and efficient annotation of histopathological photos. This involves accommodating contextual data, focusing on specific characteristics, and ensuring all-inclusive data preparation.
- *Extraction of features and preprocessing*: ChatGPT plays a crucial role in incorporating CNNs, ANNs, or any deep learning frameworks employed by AI models for analyzing electronic pathology images. It supports pathologists in pointing photos by accommodating natural language interpretations for patterns, certain characteristics, or abnormalities.
- *Training of the model*: ChatGPT can assist in annotating labeled datasets by producing expressive and contextually admissible annotations. This accords with the development of all-inclusive datasets essential for effectively training AI-based models. Throughout the training procedure, ChatGPT can provide explanations for the selections and forecasts made by the AI-based model [33,34].
- *Drawing conclusions and interpretation*: Collaborating with ChatGPT, the AI model can commence the investigation by partitioning the digitally transformed photo into smaller tiles or patches. This division facilitates an additional detailed investigation of the tissue specimen. Leveraging the knowledge gained during its training step, the AI-based model, under the guidance of ChatGPT, can scrutinize these tiny patches.
- *Localization and Categorization*: During training, ChatGPT-supported models can learn patterns linked with distinct tissues, abnormalities, and cell types by being exposed to labeled datasets.
- *Joint diagnosis or Collaborative diagnostic process*: ChatGPT can serve as an invaluable aid to pathologists, significantly enhancing diagnostic exactness and reducing the time consumed on an individual patient's prognosis. This, in turn, facilitates the prompt development of accurate therapeutic plans and the timely inception of therapy [35].
- *Retrieval of information and research*: Diagnosticians can avail of ChatGPT's capability to quickly get relevant information and summarize study insights. This keeps them notified about the latest evolutions in their field, facilitating continuous learning and staying up-to-date with the most recent developments [36,37].
- *Discussion and Consultation on Cases*: With its extensive healthcare knowledge, ChatGPT can function as a remote fellow, offering advice and findings. It can contribute by generating plans for potential differentials, suggesting courses of therapeutic, and providing a firm perspective necessary for diagnosing challenging instances [38].
- *Resource for Education*: By offering interactive learning experiences, clarifying complex notions, and responding to questions, ChatGPT can support diagnosticians in their training. It can produce detailed interpretations of several illnesses, elucidate their diagnostical features, and explain relevant prognostic ways [39].
- *Communication with Patients*: Using ChatGPT to provide explanations in a patient-friendly manner makes discussing diagnoses, potential treatments, and possible outcomes more accessible. This simplification can enhance communication and understanding between healthcare professionals and patients [40].
- *Optimization of Workflow*: Through the integration of ChatGPT into the center for research information systems, diagnosticians can streamline their workflow by automating repetitive tasks such as summarizing patient histories, producing preliminary reports, and cross-referencing with prior cases. This automation enhances efficiency in their work processes [41].

5. Challenges and limitations

Several challenges must be addressed before seamless integration into diagnosis and healthcare can occur. Machine learning relies on extensive datasets, often held by private dental practices and institutions. Overcoming challenges related to data sharing and privacy is crucial, and the development of federated guidelines and laws is essential to address these concerns. This can help rectify a common drawback observed in many studies: the shortage of datasets for training and refining AI models in dentistry [42]. As of now, the application of ChatGPT in pathology is still in its initial phases. Specifically, in ChatGPT 3.5, it is evident that the data in total on which the algorithm has been instructed plays a crucial role in its capability to furnish righteous responses to specific prompts. Various research articles have raised concerns about potential bias, transparency issues, and the harm that could arise from inexactness or completely wrong content. The phenomenon of hallucination is one significant concern. Presently, ChatGPT appears capable of generating scientifically accurate content, but there is a challenge in ensuring that the content is appropriately sourced or referenced. This issue underscores the importance of ongoing efforts to improve the transparency, reliability, and accountability of AI models, especially when applied in critical domains such as pathology [43]. Some of the other issues with ChatGPT are as follows.

- *Excessive dependence on machine learning and the resulting bias poses significant challenges*: ChatGPT produces predictions by analyzing data through ML algorithms. The effectiveness of these procedures (algorithms) is heavily reliant on the standardized training dataset. If the data is insufficient or biased, it can impede ChatGPT's capacity to generate robust results. Therefore, while ML-based models may perform well with the training data, their achievements might not be as reliable when employed with unseen, fresh data.
- *Complexity of Pathological Details*: Oral pathologies often involve intricate details and nuances. ChatGPT may struggle to comprehend and analyze complex pathological information, leading to potential inaccuracies or oversimplifications.
- *Insufficient contextual knowledge*: Due to the absence of experiential and contextual knowledge found in human pathologists, ChatGPT may encounter difficulties in recognizing nuanced descriptions or making associations between diverse pieces of data without further refinements in its present state.
- *Image Analysis Limitations*: Oral pathologies frequently require visual examination of images, such as X-rays or histological slides. ChatGPT, primarily a text-based model, may not be equipped to analyze visual data effectively.

- *Absence of image integration*: The current ChatGPT interface lacks a feature for uploading digitally transformed slides. It provides a prognosis based on the analysis of text-based history, along with cytological or histopathological and clinical information.
- *Lack of Real-time Interaction*: In certain clinical scenarios, real-time interaction and dynamic exchange of information with healthcare professionals are crucial. ChatGPT, as a static text-based model, may not provide the immediacy required for effective communication.
- *Constrained interpretability*: Due to the limited interpretability of ChatGPT, diagnosticians may find it challenging to believe the results and subsequently incorporate them into patient care decisions.
- *Data Privacy Concerns*: Integration of ChatGPT into healthcare systems may raise concerns regarding the privacy and security of patient data. Adhering to strict data protection regulations becomes paramount.
- *Integration with Existing Systems*: Seamless integration with existing medical information systems and workflows is crucial. Ensuring compatibility and smooth collaboration with other tools used in oral pathology diagnosis is a practical challenge.
- *Need for Continuous Updates*: Medical knowledge is dynamic, with ongoing advancements. ChatGPT may require regular updates to stay current with the latest developments in oral pathology, posing a challenge in maintaining relevance over time.
- *User Acceptance and Trust*: Achieving acceptance and trust from healthcare professionals, especially pathologists, is vital. Demonstrating the reliability and efficacy of ChatGPT in contributing meaningfully to oral pathology diagnosis is essential for widespread adoption.

6. Conclusion and future scope

In conclusion, while ChatGPT presents potential benefits in the field of oral pathologies, several challenges need to be addressed for its effective integration into clinical practice. The model's limitations in specificity, handling complex pathological details, and the absence of real-time interaction may impact its utility in oral pathology diagnosis. The inability to analyze visual data, interpretability challenges, data privacy concerns, the need for continuous updates, and ensuring seamless integration with existing systems are additional factors that demand careful consideration.

To unlock the full potential of ChatGPT in oral pathology, collaborative efforts between AI researchers, healthcare professionals, and technology developers are essential. Overcoming these challenges requires a commitment to refining the model's capabilities, ensuring its alignment with real-world clinical needs, and addressing ethical considerations such as data privacy. While ChatGPT may offer valuable insights, its successful implementation in oral pathology diagnosis will depend on the careful navigation of these challenges and the establishment of trust among healthcare practitioners.

The future development and integration of ChatGPT in oral pathologies will likely involve a collaborative approach, engaging stakeholders from the fields of artificial intelligence, dentistry, and healthcare. As technology advances and research progresses, ChatGPT holds the potential to become an important tool in aiding oral pathology prognosis and contributing to improved patient care.

CRediT authorship contribution statement

Payal Panwar: Conceptualization, Writing – original draft. **Shalini Gupta:** Supervision.

Declaration of competing interest

The authors declare that they have no conflict of interest.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.oor.2024.100225.

References

[1] Schwendicke F, Samek W, Krois J. Artificial intelligence in dentistry: chances and challenges. J Dent Res 2020;99:769–74.
[2] Xiang Y, Zhao L, Liu Z, Wu X, Chen J, Long E, Lin D, Zhu Y, Chen C, Lin Z, et al. Implementation of artificial intelligence in medicine: status analysis and development suggestions. Artif Intell Med 2020:102.
[3] Haug CJ, Drazen JM. Artificial intelligence and machine learning in clinical medicine, 2023. N Engl J Med 2023;30:1201–8.
[4] Talking about large language models. https://arxiv.org/pdf/2212.03551.pdf. [Accessed 3 January 2024].
[5] Niazi MK, Parwani AV, Gurcan MN. Digital pathology and artificial intelligence. Lancet Oncol 2019;20:e253–61.
[6] Ouyang L, Wu J, Jiang X, Almeida D, Wainwright CL, Mishkin P, et al. Training language models to follow instructions with human feedback. Adv Neural Inf Process Syst 2022;35:27730.
[7] Korteling JE, van de Boer-Visschedijk GC, Blankendaal RA, Boonekamp RC, Eikelboom AR. Human- versus artificial intelligence. Front Artif. Intell. 2021;4: 622364.
[8] Sinha RK, Roy AD, Kumar N, Mondal H. Applicability of ChatGPT in assisting to solve higher order problems in Pathology. Cureus 2023;15(2):1–9.
[9] Mondal S, Saha S, Mondal H, De R, Majumder R, Saha K. How to conduct inferential statistics online: a brief hands-on guide for biomedical researchers. Indian J VascEndovascSurg 2022;9:54–62.
[10] Albagieh H, Alzeer ZO, Alasmari ON, et al. Comparing artificial intelligence and senior residents in oral lesion diagnosis: a comparative study. Cureus January 03, 2024;1 6(1):e51 584. 1 0.7759/cureus.51 584.
[11] Kuscu O, Pamuk AE, Sütay Süslü N, Hosal S. Is ChatGPT accurate and reliable in answering questions regarding head and neck cancer? Front Oncol 2023;13: 1256459. https://doi.org/10.3389/fonc.2023.1256459.
[12] Malik S, Zaheer S. ChatGPT as an aid for pathological diagnosis of cancer. Pathol Res Pract 2024;253.
[13] Tizhoosh HR, Pantanowitz L. Artificial intelligence, and digital pathology: challenges and opportunities. J Pathol Inf 2018;9:38.
[14] Saltz J, Gupta R, Hou L, Kurc T, Singh P, Nguyen V, et al. Spatial organization and molecular correlation of tumor-infitrating lymphocytes using deep learning on pathology images. Cell Rep 2018;23(1):181–93.
[15] Iyengar JN. Whole slide imaging: the future scape of histopathology. Indian J Pathol Microbiol 2021;64(1):8–13.
[16] Jain E, Patel A, Parwani AV, Shafi S, Brar Z, Sharma S, et al. Whole slide imaging technology and its applications: current and emerging perspectives. Int J Surg Pathol 2023:1–16.
[17] Pallua JD, Brunner A, Zelger B, Schirmer M, Haybaeck J. The future of pathology is digital. Pathol Res Pract 2020;216(9).
[18] Jahn SW, Plass M, Moinfar F. Digital pathology: advantages, limitations and emerging perspectives. J Clin Med 2020;9(11).
[19] Betmouni S. Diagnostic digital pathology implementation: learning from the digital health experience. Digit Health 2021;7.
[20] Biswas S. ChatGPT and the future of medical writing. Radiology 2023;307(2).
[21] Sandbank J, Bataillon G, Nudelman A, Krasnitsky I, Mikulinsky R, Bien L, et al. Validation and real-world clinical application of an artificial intelligence algorithm for breast cancer detection in biopsies. NPJ Breast Cancer 2022;8(1):129.
[22] Wang S, Chen A, Yang L, Cai L, Xie Y, Fujimoto J, Gazdar A, Xiao G. Comprehensive analysis of lung cancer pathology images to discover tumor shape and boundary features that predict survival outcome. Sci Rep 2018;8(1).
[23] Sebastian AM, Peter D. Artificial intelligence in cancer research: trends, challenges, and future directions. Life 2022;12(12).
[24] Iqbal MJ, Javed Z, Sadia H, Qureshi IA, Irshad A, Ahmed R, et al. Clinical applications of artificial intelligence and machine learning in cancer diagnosis: looking into the future, 21. Cancer Cell Int 2021;21(1).
[25] Silva-Rodríguez J, Colomer A, Sales MA, Molina R, Naranjo V. Going deeper through the Gleason scoring scale: an automatic end-to-end system for histology prostate grading and cribriform pattern detection. Comput Methods Progr Biomed 2020;195.
[26] Bohr A, Memarzadeh K. The rise of artificial intelligence in healthcare applications. Artif. Intell. Healthc. 2020:25–60.
[27] Stenzinger A, Alber M, Allgäuer M, Jurmeister P, Bockmayr M, Budczies J, et al. Artificial intelligence and pathology: from principles to practice and future applications in histomorphology and molecular profiling. Semin Cancer Biol 2022; 84:129–43.
[28] Zuraw A, Aeffner F. Whole-slide imaging, tissue image analysis, and artificial intelligence in veterinary pathology: an updated introduction and review. Vet Pathol 2022;59(1):6–25.
[29] Hosny A, Parmar C, Quackenbush J, Schwartz LH, Aerts HJWL. Artificial intelligence in radiology. Nat Rev Cancer 2018;18(8):500–10.
[30] Lee S, Kim HS. Prospect of artificial intelligence based on electronic medical record. J. Lipid Atheroscler. 2021;10(3):282–90.
[31] Michelhaugh SA, Januzzi Jr JL. Using artificial intelligence to better predict and develop biomarkers. Heart Fail Clin 2022;18(2):275–85.

[32] Chen ZH, Lin L, Wu CF, Li CF, Xu RH, Sun Y. Artificial intelligence for assisting cancer diagnosis and treatment in the era of precision medicine. Cancer Commun 2021;41(11):1100–15.
[33] Xue VW, Lei P, Cho WC. The potential impact of ChatGPT in clinical and translational medicine. Clin Transl Med 2023;13(3).
[34] Arshad HB, Butt SA, Khan SU, Javed Z, Nasir K. ChatGPT and artificial intelligence in hospital level research: potential, precautions, and prospects. Methodist Debakey Cardiovasc J 2023;19(5):77–84.
[35] Sarker IH. Machine learning: algorithms, real-world applications and research directions. SN Comput. Sci. 2021;2(3).
[36] Sallam M. ChatGPT utility in healthcare education, research, and practice: systematic review on the promising perspectives and valid concerns. Health 2023; 11(6).
[37] Raghupathi W, Raghupathi V. Big data analytics in healthcare: promise and potential. Health Inf Sci Syst 2014;2:3.
[38] Cazzato G, Capuzzolo M, Parente P, Arezzo F, Loizzi V, Macorano E, et al. ChatGPT in diagnostic human pathology: will it Be useful to pathologists? A preliminary review with 'query session' and future perspectives, AI, vol. 4; 2023. p. 1010–22. 4.
[39] Javaid M, Haleem A, Singh RP, Khan S, Khan IH. Unlocking the opportunities through ChatGPT Tool towards ameliorating the education system. BenchCouncil Trans. Benchmarks, Stand. Eval. 2023;3(2).
[40] Sallam M. ChatGPT utility in healthcare education, research, and practice: systematic review on the promising perspectives and valid concerns. Health 2023; 11(6):887.
[41] Cadamuro J, Cabitza F, Debeljak Z, De Bruyne S, Frans G, Perez SME, et al. Potentials and pitfalls of ChatGPT and natural-language artificial intelligence models for the understanding of laboratory medicine test results. An assessment by the European Federation of Clinical Chemistry and Laboratory Medicine (EFLM) Working Group on Artificial Intelligence (WG-AI). Clin Chem Lab Med 2023;61(7): 1158–66.
[42] Patil S, Albogami S, Hosmani J, Mujoo S, Kamil MA, Mansour MA, Abdul HN, Bhandi S, Ahmed SSSJ. Artificial Intelligence in the diagnosis of oral diseases: applications and pitfalls. Diagnostics 2022;12:1029.
[43] Cazzato G, Capuzzolo M, Parente P, Arezzo F, Loizzi V, Macorano E, Marzullo A, Cormio G, Ingravallo G. Chat gpt in diagnostic human pathology: will it Be useful to pathologists? A preliminary review with 'query session' and future perspectives. AI, vol. 4; 2023. p. 1010–22.