TITLE

Implications of large language models such as ChatGPT for dental medicine

AUTHOR NAMES AND INSTITUTIONAL AFFILIATIONS

Florin Eggmann[1,2], Roland Weiger[1], Nicola U. Zitzmann[3], Markus B. Blatz[2]

[1]Department of Periodontology, Endodontology, and Cariology, University Center for Dental Medicine Basel UZB, University of Basel, Basel, Switzerland

[2]Department of Preventive and Restorative Sciences, Penn Dental Medicine, Robert Schattner Center, University of Pennsylvania, Philadelphia, PA, USA

[3]Department of Reconstructive Dentistry, University Center for Dental Medicine Basel UZB, University of Basel, Basel, Switzerland

CORRESPONDENCE

Dr. med. dent. Florin Eggmann, Klinik für Parodontologie, Endodontologie und Kariologie, Universitäres Zentrum für Zahnmedizin Basel UZB, Universität Basel, Mattenstrasse 40, CH-4058 Basel, Schweiz. Tel. +41 61 267 26 80; E-Mail: florin.eggmann@unibas.ch

KEYWORDS

Artificial intelligence, Dental care, Dental education, Evidence-based dentistry, Machine learning

SUMMARY

Artificial intelligence (AI) applications may offer several benefits for healthcare personnel, including dental practitioners. Large language models (LLMs) are AI applications trained on vast amounts of textual data. LLMs can perform various language-related tasks. ChatGPT, a LLM with a conversational interface, has been launched in November 2022 and is accessible online for free. Despite its impressive capabilities, the ChatGPT chatbot has serious limitations, such as occasionally giving incorrect answers, producing nonsensical content, and presenting misinformation as fact. LLMs are not expected to significantly impact the everyday work of dental practitioners, dental assistants, and hygienists. They could, however, change the work of administrative personnel and the provision of dental telemedicine. LLMs have the potential to become useful tools for clinical decision support, summarization of medical records and scientific articles, and efficient written communications. They can also facilitate multilingual communication. The number of people seeking health information from LLMs will likely increase soon. It is therefore crucial to safeguard against inaccurate, outdated, and biased responses to health-related queries. LLMs pose challenges for patient data confidentiality and cyber-security, which must be tackled. In dental education in Switzerland, where multiple-choice exams, practical assessments, and supervised treatments are primarily used for students' performance evaluation, LLMs present fewer challenges than in other academic fields. LLMs can help authors of academic articles write more fluently, but it is important to establish the boundaries of acceptable use of LLMs in science. This narrative review provides an overview of the possible applications and current limitations of LLMs in dentistry.

Introduction

Rapid advancements in artificial intelligence (AI) offer healthcare professionals numerous benefits, including improved diagnosis, prevention, and treatment of diseases and injuries (CHEN & ASCH 2017). In dental medicine, for example, machine learning applications are on the verge of causing a step change in radiographic caries detection (SCHWENDICKE ET AL. 2022, HUNG ET AL. 2023).

Large language models (LLMs) are AI applications trained on hundreds of terabytes of textual data. They are generative mathematical models of the statistical distribution of tokens in the vast public corpus of human generated text, where the tokens include words, graphemes, individual characters, and punctuation marks (SHANAHAN 2022). When given a prompt, such as “Who first described the use of phosphoric acid etching for dental bonding?”, the LLM provides the answer, “Dr. Michael Buonocore”, because the statistical assessment of the corpus of human generated text indicate that this is most likely the correct reply (NAUGHTON 2023). The LLM does not possess actual knowledge of adhesive dentistry. The capabilities of LLMs are nevertheless impressive. LLMs can generate fluent and coherent texts, answer questions, translate languages, and perform other language-related tasks.

ChatGPT, developed by the US company OpenAI, part-owned by Microsoft, is a LLM with a conversational interface that is free to use and easily accessible online. Since its launch in November 2022, the ChatGPT chatbot has been used by millions, sparking both considerable excitement and serious concern (CASTELVECCHI 2022, STOKEL-WALKER 2022, VAN DIS ET AL. 2023).

LLMs have numerous potential applications in dental medicine, from streamlined dental record keeping to AI support in clinical decision making (SHEN ET AL. 2023). While such applications may become available in the near future, the current iteration

of LLMs are already adept at text summarization and translation, making them useful tools for dental practitioners and students, especially those studying in a non-native language. However, LLMs can also give entirely wrong answers, produce nonsensical content, and present misinformation and disinformation as fact, which raises serious concerns in critical fields such as healthcare.

Given the importance of emergent AI applications for patient care, education, and dental research, the aim of this narrative review is to provide an overview of possible applications of LLMs such as ChatGPT in dentistry, their current limitations, and drawbacks.

Implications of large language models for dental medicine

Dental professions

The provision of dental care largely depends on face-to-face communication, clinical and radiographic examinations, and operative procedures. Therefore, LLMs will not substantially transform the everyday working life of dental practitioners, dental assistants, and hygienists. By contrast, advancements in LLMs could change the work and workload of administrative personnel—such as insurance claims administrators or clerical staff in dental offices—and alter the delivery of dental telemedicine.

Clinical decision support

The accuracy of answers provided by LLMs depends on the quantity, quality, and type of data used to train LLMs. The dataset to which a LLM has access needs to be comprehensive, up-to-date, and vetted to ensure that the whole body of evidence, including the latest scientific findings, is considered and to avoid bias from sources that contain misinformation or disinformation (PATCAS ET AL. 2022). The knowledge base of ChatGPT ends in 2021, rendering some prompts and queries useless. Moreover, ChatGPT is not designed to give medical guidance, which is why it is unsuitable for

clinical decision support today. However, LLMs have considerable potential to become an additional tool for clinical decision support. For instance, LLMs that also take account of patients' electronic health record data may soon be used to improve the evidence-based selection of appropriate medical imaging examinations (SHEN ET AL. 2023). Additionally, the use of LLMs for summarizing extensive medical records has the potential to aid healthcare professionals in quickly finding pertinent information, eliminating the need for time-intensive chart reviews (SHEN ET AL. 2023).

Administrative work

Dr. Clifford Stermer, a rheumatologist, has recently described how the use of ChatGPT makes writing preauthorization requests to insurance companies more efficient (SHEN ET AL. 2023). This demonstrates the potential of LLMs in supporting healthcare professionals in routine written communications and record keeping, leading to improved efficiency in administrative work. Such time savings could free up healthcare professionals' time for other important tasks, potentially enhancing the quality of patient care and reducing costs. LLMs can also facilitate multilingual communication by translating texts, which is helpful to overcome language barriers in patient communication and elsewhere.

Patient education

The internet has had a profound impact on patient education, making a vast amount of health information readily available. Wikipedia, for example, is the first port of call for many people worldwide seeking medical information (MASUKUME ET AL. 2016). ChatGPT and similar LLMs are predicted to challenge the domination of web-based search engines such as Google in online query traffic. The ascendancy of LLMs chatbots in the online search market has wide ramifications. ChatGPT is already capable of producing coherent and human-like conversational responses to various prompts and questions. In addition, unlike conventional search engines, it is not

necessary to click on a website to get an answer. The dialogue-based, human-like interaction of ChatGPT holds great appeal for users: Most simple prompts and queries will return an articulate and succinct reply. However, a major disadvantage of ChatGPT compared with web-based search engines is the impossibility to evaluate the credibility of the sources of its responses. Furthermore, ChatGPT sometimes wrongfully answers a question with complete confidence, its current knowledge base does not extend beyond the year 2021, and it cannot access the internet. Owing to these serious limitations in reliability, transparency, and knowledge, ChatGPT is not suitable for providing health advice and guidance today. The answers to health-related queries online can have immediate health consequences (MASUKUME ET AL. 2016). With the expected increase in the number of people requesting health information from LLMs, it is therefore vital to safeguard against inaccurate and biased responses to such queries.

Patient data privacy

The use of LLMs in dental medicine may involve the collection and storage of personal and medical information. This raises concerns about privacy and data security. Before LLMs are used in dental medicine, it is imperative to implement state-of-the-art data protection measures, including encryption of sensitive information, secure data transmission and storage, and access controls. The terms of service of ChatGPT allows OpenAI to collect and use log data, device information, and most importantly, usage data. To protect the privacy and confidentiality of patients' personal and medical information, no patient data whatsoever must therefore be entered into ChatGPT.

Cybersecurity

There has been a uptick in cyber-attacks targeted at healthcare providers (IACONO ET AL. 2022). The ransomware attack carried out against the United Kingdom’s National Health Service in August 2022 affected many of its services, including ambulance

dispatch, emergency prescriptions, mental health services, patient referrals, and out-of-hours appointment scheduling (MILMO 2022). Such malware attacks put a spotlight on the vulnerabilities of health organizations. LLMs can potentially be abused to assist in writing phishing messages and developing malware (KULESH 2022). The companies and agencies involved in building LLMs must therefore take concerted efforts to thwart the risks of malicious use. Additionally, considering that threat actors are adept at exploiting technologies such as LLMs, it is crucial to fortify defenses against malware attacks to protect patient data and the operational hardware and software of healthcare providers (SINGLETON ET AL. 2022).

Dental education

LLMs bring both opportunities and challenges to higher education, particularly in fields that heavily rely on written assignments (GRAHAM 2022, STOKEL-WALKER 2022, TAYLOR 2023). LLMs respond differently to each interaction, which is why plagiarism checkers are of hardly any use (TAYLOR 2023). To meet these challenges and to harness the potential of LLMs, some higher education curricula need to be tweaked and adapted. Given that dental schools in Switzerland primarily use multiple-choice exams, practical assessments, and supervised patient treatments to evaluate dental students' knowledge and skills, LLMs do not present the same challenges in dentistry as they do in other academic disciplines. Nonetheless, it is important to increase dental students' literacy in AI applications relevant to dental medicine. For this purpose, a special curriculum has recently been proposed (SCHWENDICKE ET AL. 2023). The implementation of this core curriculum, or a modification of it, at Swiss dental schools merits serious consideration.

Scientific writing

English is the lingua franca used in most dental and medical journals today. Drafting study reports and grant applications that are clear and concise can be challenging for

non-native English speakers and even some native English speakers (BARON 2012). LLMs may be a useful tool for authors to make the writing in such reports more fluent, thus levelling the playing field for non-native English speakers (KITAMURA 2023). There are, however, grave concerns that LLMs could promote flawed or even fabricated research (ELSE 2023, GORDIJN & TEN HAVE 2023). In a recent investigation, available as preprint, it has been shown that scientific abstracts written by ChatGPT based on completely generated data evaded plagiarism detection and often fooled human reviewers (GAO ET AL. 2022). Some publishers have already acted. For instance, Science, one of the most highly ranked academic journals, has banned the use of text written by AI, machine learning, or similar algorithmic tools (THORP 2023). Springer-Nature, a leading publishing house, has also amended its guidelines (NATURE EDITORIALS 2023). The amended guidelines require that researchers report their use of LLM tools in the methods or acknowledgments section. Additionally, they prohibit listing LLMs such as ChatGPT as authors because LLMs cannot accept accountability for the work produced, which is an important requirement of authorship.

To distinguish between text written by humans and text written by LLMs, OpenAI launched a classifier tool on January 31, 2023 (KIRCHNER ET AL. 2023). However, according to OpenAI, the classifier tool only correctly identifies 26% of English texts written by a LLM and incorrectly labels 9% of texts written by a human as probably written by a LLM (KIRCHNER ET AL. 2023). This underscores the importance of rigorous peer review to weed out unsound science and misinformation.

Conclusions

LLMs are likely to have significant impacts on various aspects of dental medicine in the near future. It is, however, challenging to envisage the exact nature and extent of these impacts. As AI technologies are still in their early stages, further research and

development is required to fully realize the potential benefits of LLMs for dental healthcare. Along with the potential benefits, it is crucial to consider and study the negative implications that LLMs may carry in dental medicine and beyond. To protect persons who seek medical information from LLMs, the factual accuracy of health-related content produced by LLMs must be scientifically evaluated and confirmed. Robust quality control measures must be implemented to monitor and assess the output of LLMs. The use of LLMs and other AI applications in dentistry needs to be carefully regulated, managed, and monitored to ensure that their use is safe, ethical, and beneficial for dental healthcare personnel and patients alike.

Acknowledgments

We recognize, with deep gratitude, the support of the Gottfried and Julia Bangerter-Rhyner Foundation and the Freiwillige Akademische Gesellschaft Basel (FAG), who funded scholarships for Florin Eggmann during this work. We disclose the use of ChatGPT during the revision of the draft of the report to make the text more fluent. The version of the draft that was revised using ChatGPT is available in an open repository (EGGMANN ET AL. 2023). The French and German summaries were made based on translations created with DeepL, a neural machine translation service (www.deepl.com). No LLM was used for conceptualization, literature review, data interpretation, and drawing conclusions.

Conflicts of interest

The authors report no competing financial and/or non-financial interests in relation to the work described.

Zusammenfassung

Anwendungen der künstlichen Intelligenz (KI) können dem Gesundheitspersonal, einschliesslich Zahnärztinnen/Zahnärzten verschiedene Vorteile bieten. Umfangreiche Sprachmodelle (Large Language Models, LLMs) sind KI-Anwendungen, die mit grossen Mengen von Textdaten trainiert werden und verschiedene sprachbezogene Aufgaben durchführen können. ChatGPT, ein LLM mit einer Konversationsschnittstelle, wurde im November 2022 auf den Markt gebracht und ist kostenlos online verfügbar. Trotz seiner beeindruckenden Fähigkeiten hat ChatGPT erhebliche Einschränkungen und Unzulänglichkeiten. Beispielsweise gibt ChatGPT teilweise fehlerhafte Antworten oder stellt Fehlinformationen als Fakten dar. Es ist nicht zu erwarten, dass LLMs den Arbeitsalltag von Zahnärztinnen/Zahnärzten, Dentalassistentinnen/Dentalassistenten und Dentalhygienikerinnen/Dentalhygienikern wesentlich verändern werden. LLMs könnten jedoch zu Veränderungen bei der Arbeit von Verwaltungspersonal und bei zahnärztlichen Telemedizindiensten führen. LLMs haben das Potenzial, sich zu nützlichen Instrumenten für die klinische Entscheidungsfindung, die Zusammenfassung umfangreicher Patientenakten und wissenschaftlicher Artikel sowie die Vereinfachung der schriftlichen Kommunikation zu entwickeln. LLMs können zudem die mehrsprachige Kommunikation insbesondere bei Patientinnen/Patienten mit Sprachbarrieren erleichtern. Die Zahl der Menschen, die LLMs nutzen, um Gesundheitsinformationen abzurufen, wird wahrscheinlich zeitnah steigen. Daher ist es von entscheidender Bedeutung, sicherzustellen, dass LLMs keine ungenauen, überholten oder verzerrten Antworten auf gesundheitsbezogene Anfragen liefern. LLMs bringen Herausforderungen in Bezug auf den Patientendatenschutz und die Cybersicherheit mit sich, die es zu bewältigen gilt. In der zahnmedizinischen Ausbildung in der Schweiz, in der für die Leistungsbewertung der Studierenden hauptsächlich Multiple-Choice-Prüfungen, praktische Beurteilungen und beaufsichtige

Patientenbehandlungen eingesetzt werden, stellen LLMs weniger Herausforderungen dar als in akademischen Fachbereichen, in denen schriftliche Aufgabenstellungen wie Essays eine hohe Bedeutung für die Leistungsbeurteilung haben. LLMs können Autoren von Artikeln akademischer Fachzeitschriften helfen, flüssiger zu schreiben. Es ist allerdings wichtig, die Grenzen der zulässigen Verwendung von LLMs in der Wissenschaft festzulegen. Dieser narrative Übersichtsartikel beschreibt die möglichen Anwendungen und die derzeitigen Limitationen von LLMs in der Zahnmedizin.

Résumé

Les applications d'intelligence artificielle (IA) peuvent offrir différents avantages aux professionnels de la santé, y compris aux dentistes. Les modèles linguistiques étendus (Large Language Models, LLMs) sont des applications d'IA entraînées avec de grandes quantités de données textuelles et capables d'effectuer différentes tâches liées à la langue. ChatGPT, un LLM doté d'une interface conversationnelle, a été lancé en novembre 2022 et est disponible gratuitement en ligne. Malgré ses capacités impressionnantes, ChatGPT présente des limitations et des insuffisances importantes. Par exemple, ChatGPT donne parfois des réponses erronées ou présente des informations erronées comme des faits. Il ne faut pas s'attendre à ce que les LLMs modifient considérablement le travail quotidien des dentistes, des assistants dentaires et des hygiénistes dentaires. Les LLMs pourraient toutefois entraîner des changements dans le travail du personnel administratif et dans les services de télémédecine dentaire. Les LLMs ont le potentiel de devenir des outils utiles pour la prise de décisions cliniques, la synthèse de dossiers de patients volumineux et d'articles scientifiques, ainsi que pour la simplification de la communication écrite. Les LLMs peuvent en outre faciliter la communication multilingue, notamment pour les patient(e)s ayant des barrières linguistiques. Le nombre de personnes utilisant les

LLMs pour accéder à des informations sur la santé va probablement augmenter rapidement. Il est donc essentiel de veiller à ce que les LLMs ne fournissent pas de réponses inexactes, obsolètes ou déformées aux demandes liées à la santé. Les LLMs présentent des défis en matière de protection des données des patients et de cybersécurité qu'il convient de relever. Dans l'enseignement dentaire en Suisse, où les examens à choix multiples, les évaluations pratiques et les soins supervisés aux patients sont les principaux outils utilisés pour évaluer les performances des étudiants, les LLMs posent moins de défis que dans les disciplines académiques, où les tâches écrites telles que les essais ont une grande importance pour l'évaluation des performances. Les LLMs peuvent aider les auteurs d'articles de revues académiques à écrire de manière plus fluide. Il est toutefois important de définir les limites de l'utilisation autorisée des LLMs dans le monde universitaire. Cet article de synthèse narratif décrit les applications possibles et les limites actuelles des LLMs en dentisterie.

References

BARON T H: ABC's of writing medical papers in English. Korean J Radiol 13 Suppl 1: S1-11 (2012)

CASTELVECCHI D: Are ChatGPT and AlphaCode going to replace programmers? Nature (2022)

CHEN J H, ASCH S M: Machine learning and prediction in medicine - beyond the peak of inflated expectations. N Engl J Med 376: 2507–2509 (2017)

EGGMANN F, WEIGER R, ZITZMANN N U, BLATZ M B: Implications of large language models such as ChatGPT for dental medicine. Zenodo (2023)

ELSE H: Abstracts written by ChatGPT fool scientists. Nature 613: 423 (2023)

GAO C A, HOWARD F M, MARKOV N S, DYER E C, RAMESH S, LUO Y, PEARSON A T: Comparing scientific abstracts generated by ChatGPT to original abstracts using an artificial intelligence output detector, plagiarism detector, and blinded human reviewers. Preprint (2022)

GORDIJN B, TEN HAVE H: ChatGPT: evolution or revolution? Med Health Care Philos (2023)

GRAHAM F: Daily briefing: Will ChatGPT kill the essay assignment? Nature (2022)

HUNG K F, YEUNG A W K, BORNSTEIN M M, SCHWENDICKE F: Personalized dental medicine, artificial intelligence, and their relevance for dentomaxillofacial imaging. Dentomaxillofac Radiol 52: 20220335 (2023)

IACONO L, WOJCIESZEK K, GLASS G: Q2 2022 threat landscape: ransomware returns, healthcare hit. Available at: https://www.kroll.com/-/media/kroll-images/pdfs/q2-2022-threat-landscape-report.pdf, [accessed on 01.30.2023] (2022)

KIRCHNER J H, AHMAD L, AARONSON S, LEIKE J: New AI classifier for indicating AI-written text. Available at: https://openai.com/blog/new-ai-classifier-for-indicating-ai-written-text/, [accessed on 02.01.2023] (2023)

KITAMURA F C: ChatGPT is shaping the future of medical writing but still requires human judgment. Radiology: 230171 (2023)

KULESH S: Why ChatGPT can be dangerous for every internet user. Available at: https://timesofindia.indiatimes.com/gadgets-news/why-chatgpt-can-be-dangerous-to-every-internet-user/articleshow/96393104.cms, [accessed on 01.30.2023] (2022)

MASUKUME G, KIPERSZTOK L, DAS D, SHAFEE T M A, LAURENT M R, HEILMAN J M: Medical journals and Wikipedia: a global health matter. Lancet Glob Health 4: e791 (2016)

MILMO D: NHS ransomware attack: what happened and how bad is it? Available at: https://www.theguardian.com/technology/2022/aug/11/nhs-ransomware-attack-what-happened-and-how-bad-is-it, [accessed on 01.30.2023] (2022)

NATURE EDITORIALS: Tools such as ChatGPT threaten transparent science; here are our ground rules for their use. Nature: 612 (2023)

NAUGHTON J: The ChatGPT bot is causing panic now – but it'll soon be as mundane a tool as Excel. Available at: https://www.theguardian.com/commentisfree/2023/jan/07/chatgpt-bot-excel-ai-chatbot-tech, [accessed on 01.30.2023] (2023)

PATCAS R, BORNSTEIN M M, SCHÄTZLE M A, TIMOFTE R: Artificial intelligence in medico-dental diagnostics of the face: a narrative review of opportunities and challenges. Clin Oral Investig 26: 6871–6879 (2022)

SCHWENDICKE F, CEJUDO GRANO DE ORO, J, GARCIA CANTU A, MEYER-LUECKEL H, CHAURASIA A, KROIS J: Artificial intelligence for caries detection: value of data and information. J Dent Res 101: 1350–1356 (2022)

SCHWENDICKE F, CHAURASIA A, WIEGAND T, URIBE S E, FONTANA M, AKOTA I, TRYFONOS O, KROIS J: Artificial intelligence for oral and dental healthcare: core education curriculum. J Dent 128: 104363 (2023)

SHANAHAN M: Talking about large language models. Available at: https://arxiv.org/pdf/2212.03551, [accessed on 01.31.2023] (2022)

SHEN Y, HEACOCK L, ELIAS J, HENTEL K D, REIG B, SHIH G, MOY L: ChatGPT and other large language models are double-edged swords. Radiology: 230163 (2023)

SINGLETON, CAMILLE, DEBECK C, CHUNG J, MCMILLEN D, CRAIG S, MOORE S, HAMMOND C, DWYER J, FRYDRYCH M, VILLADSEN O, EMERSON R, JOURDAN G-V, ONUT V, CARRUTHERS S,

345 LAURIE A, ALVAREZ M, WUTTKE S, PRASSINOS G, ZORABEDIAN J, MAYNE M, KESSEM K,
346 GALLAGHER I, EITAN A: X-force threat intelligence index 2022. Available at:
347 https://www.ibm.com/downloads/cas/ADLMYLAZ, [accessed on 02.04.2023] (2022)
348 STOKEL-WALKER C: AI bot ChatGPT writes smart essays - should professors worry? Nature
349 (2022)
350 TAYLOR P: On ChatGPT. London Review of Books 45. Available at: https://www.lrb.co.uk/the-
351 paper/v45/n01/paul-taylor/on-chatgpt, [accessed on 01.10.2023] (2023)
352 THORP H H: ChatGPT is fun, but not an author. Science 379: 313 (2023)
353 VAN DIS E A M, BOLLEN J, ZUIDEMA W, VAN ROOIJ R, BOCKTING C L: ChatGPT: five priorities
354 for research. Nature 614: 224–226 (2023)
355