



Artificial intelligence bot ChatGPT in medical research: the potential game changer as a double-edged sword

Jari Dahmen^{1,2,3} · M. Enes Kayaalp⁴ · Matthieu Ollivier⁵ · Ayoosh Pareek⁶ · Michael T. Hirschmann⁷ · Jon Karlsson⁸ · Philipp W. Winkler⁹

Received: 5 February 2023 / Accepted: 10 February 2023 / Published online: 21 February 2023

© The Author(s) under exclusive licence to European Society of Sports Traumatology, Knee Surgery, Arthroscopy (ESSKA) 2023

Keywords Plagiarism · ChatGPT · Artificial intelligence · Natural language processing

ChatGPT is a language generation model developed by OpenAI that is able to generate text based on the input it receives [1]. The new concern is an Artificial Intelligence (AI)-powered chatbot being able to generate intelligent “scientific” texts in response to prompts or questions from its users. As such, with its ability to understand and generate human language, ChatGPT can assist researchers in a variety of tasks such as literature review, data analysis and even the creation of new hypotheses.

When the chatbot is asked to tell us something about our journal, it responds with the following text—which we can say has been written in a technically well-designed manner with fluent language and a natural feel, although not fully comprehensive and not perfectly descriptive: “Knee Surgery, Sports Traumatology, Arthroscopy (KSSTA) is a peer-reviewed medical journal that covers research related to the diagnosis, treatment and rehabilitation of knee injuries and disorders. The journal is published by Springer and it is intended for orthopaedic surgeons, sports medicine physicians and other healthcare professionals who specialise in

knee surgery and sports traumatology. The journal publishes original research articles, review articles, case reports and surgical technique articles in the areas of arthroscopy, ligament and cartilage injuries, meniscus injuries, osteoarthritis and knee joint replacement. The KSSTA is an international journal and it publishes articles from authors all over the world. The journal is indexed in various databases such as PubMed, Scopus and Web of Science (answer by ChatGPT on 21 January 2023, <https://openai.com>).

One of the advantages of using ChatGPT in research may be its potential ability swiftly and accurately to analyse large amounts of data which may include scientific articles and medical reports, as well as patient records, all of which can provide novel insights into the causes, symptoms and treatment options for orthopaedic and sports medicine conditions. Through the use of natural language processing (NLP) techniques, ChatGPT is supposedly able to extract relevant information from the texts and present it in a structured format. Another different application of ChatGPT is the potential to assist in the creation of novel hypotheses for

✉ Jari Dahmen
j.dahmen@amsterdamumc.nl

¹ Department of Orthopaedic Surgery and Sports Medicine, Amsterdam Movement Sciences, Amsterdam UMC, Location AMC, University of Amsterdam, Meibergdreef 9, Amsterdam, The Netherlands

² Academic Center for Evidence Based Sports Medicine (ACES), Amsterdam, The Netherlands

³ Amsterdam Collaboration for Health and Safety in Sports (ACHSS), International Olympic Committee (IOC) Research Center Amsterdam UMC, Amsterdam, The Netherlands

⁴ Department for Orthopaedics and Traumatology, Istanbul Kartal Dr. Lutfi Kirdar Training and Research Hospital, Istanbul, Turkey

⁵ Aix Marseille Univ, CNRS, ISM, Inst Movement Sci, Marseille, France

⁶ Sports Medicine and Shoulder Service, Department of Orthopedic Surgery, Hospital for Special Surgery, 535 East 70th Street, New York, NY 10021, USA

⁷ Department of Orthopedic Surgery and Traumatology, Head Knee Surgery and DKF Head of Research, Kantonsspital Baselland, Bruderholz, 4101 Bottmingen, Switzerland

⁸ Department for Orthopaedics, Sahlgrenska University Hospital, Institute of Clinical Sciences, Sahlgrenska Academy, Gothenburg University, Gothenburg, Sweden

⁹ Department of Orthopaedics and Traumatology, Kepler University Hospital Linz, Linz, Austria

researchers. Through the analysis of existing research and the identification of gaps in knowledge, ChatGPT might be able to generate new ideas for further investigation which may save researchers time and resources.

Moreover, ChatGPT could also assist in the development of clinical decision, support systems through the analysis of patient records and the identification of common patterns. With the help of ChatGPT, a number of authors have already included the AI tool as a co-author as such crediting ChatGPT [2, 3]. However, not all scientists and authors agree, as Thorp [4] recently stated that “ChatGPT is fun, but not an author” [5].

However, scientists must be extremely aware of the potential problems and disadvantages of using ChatGPT in the field of research—also specifically in the case of orthopaedic and sports medicine research and science. It is a fact that ChatGPT could be used to generate manuscripts that are as such potentially seen as plagiarised content due to the authors not having written the text themselves. Another example of potential hazardous use is that researchers input a passage or passages of a text from a previously published article with ChatGPT generating a text that is highly similar to it. Additionally, ChatGPT could also be used to generate text that is almost identical to previously published research, which could be used to manipulate the results of a study or mislead researchers and readers. A recent study [6] published on a preprint server analysed whether human reviewers were able to distinguish original scientific abstracts from AI-generated abstracts written by ChatGPT. Blinded reviewers reported difficulty distinguishing between human-written and AI-generated abstracts. In fact, ChatGPT was able to mislead blinded reviewers in 32% of abstracts generated by the AI bot. The dangers and disadvantages—besides the plagiarism issue—are that the generated texts can consist of a lack of context (ChatGPT is trained on a large dataset of text but may not have enough information about a specific case), inaccuracy, bias (the data used to train ChatGPT may contain biases) and a lack of understanding of the nuances related to medical science(s) and language. The AI bot could underestimate the importance or novelty of articles on some issues, which are fewer in number. This could cause some potentially important aspects of new research findings to be overlooked or the hints obtained by novel research data not to be seen. The AI bot could therefore limit researchers with a more generalised perspective rather than a quality-based assessment of the present data. To avoid the above-mentioned pitfalls, it is important to advise the potential users of ChatGPT to utilise the tool responsibly and always correctly cite any sources used. The related concerns that are currently being raised are the extent to which current plagiarism checkers embedded in Editorial Managers of journals are able to identify

plagiarism arising from the use of ChatGPT. As the conventional plagiarism detection tools, these may not be sufficient and/or sensitive enough appropriately to detect plagiarism arising from chatbots. Consequently, as members of the scientific community and editors of the KSSTA journal, we should advocate the incorporation of more sophisticated tools for the determination of how likely it is that the submitted content is human generated or AI generated. Consequently, the KSSTA journal will include output detectors to detect AI-generated manuscript(s) in the near future. Moreover, at this stage of the initial discussion, we recommend not using ChatGPT in articles submitted to the KSSTA. If it is nevertheless used, it needs to be acknowledged when submitting a manuscript to a journal (of any kind). As a scientific community, we need to define the way research is added by AI in many ways, but in particular by chatbots. There are many good scientific practice questions to be considered, as well as ethical considerations. The discussion has just started and will continue.

Acknowledgements Not applicable.

Author contributions JD, MEK, MO, AP, MH, JK, and PWW drafted the manuscript and completed necessary revisions. All authors reviewed and edited the manuscript.

Funding Not applicable.

Data availability Data availability not applicable for this type of research.

Declarations

Conflict of interest Jari Dahmen, M. Enes Kayaalp, M. Ollivier, A. Pareek, Michael T. Hirschmann, Jon Karlsson, Philipp W. Winkler: not applicable.

Ethical approval Not applicable.

Informed consent Not applicable.

References

- Gordijn B, Have HT (2023) ChatGPT: evolution or revolution? *Med Health Care Philos.* <https://doi.org/10.1007/s11019-023-10136-0>
- King MR (2023) A conversation on artificial intelligence, Chatbots, and plagiarism in higher education. *Cell Mol Bioeng* 16:1–2
- Zhavoronkov A (2022) Rapamycin in the context of Pascal’s Wager: generative pre-trained transformer perspective. *Oncoscience* 9:82–84
- Thorp HH (2023) ChatGPT is fun, but not an author. *Science* 379:313
- Graham F (2023) Daily briefing: ChatGPT listed as author on research papers. *Nature.* <https://doi.org/10.1038/d41586-023-00188-w>

6. Gao CA, Howard FM, Markov NS, Dyer EC, Ramesh S, Luo Y et al (2022) Comparing scientific abstracts generated by Chat-GPT to original abstracts using an artificial intelligence output detector, plagiarism detector, and blinded human reviewers. *Biorxiv*. <https://doi.org/10.1101/2022.12.23.5216102022.2012.2023.521610>

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