



LETTER

Quality of citation data using the natural language processing tool ChatGPT in rheumatology: creation of false references

Axel J Hueber ,^{1,2} Arnd Kleyer ^{2,3}

To cite: Hueber AJ, Kleyer A. Quality of citation data using the natural language processing tool ChatGPT in rheumatology: creation of false references. *RMD Open* 2023;**9**:e003248. doi:10.1136/rmdopen-2023-003248

► Additional supplemental material is published online only. To view, please visit the journal online (<http://dx.doi.org/10.1136/rmdopen-2023-003248>).

Received 20 April 2023

Accepted 15 May 2023



© Author(s) (or their employer(s)) 2023. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

¹Division of Rheumatology, Paracelsus Medical University, Klinikum Nürnberg, Nürnberg, Germany

²Department of Internal Medicine 3, Rheumatology and Immunology, Friedrich-Alexander-Universität Erlangen-Nürnberg and Universitätsklinikum Erlangen, Erlangen, Germany

³Deutsches Zentrum Immuntherapie, Friedrich-Alexander-Universität Erlangen-Nürnberg and Universitätsklinikum Erlangen, Erlangen, Germany

Correspondence to

Dr Axel J Hueber;
Axel.Hueber@fau.de

Artificial intelligence (AI) programmes such as ChatGPT from Open AI are on everyone's lips. They do not only create texts for stories, poems or just instructions for a fruit salad, but also creatively discuss real but also fictional things (eg, what would happen if Christopher Columbus discovered America in 2020 (personal communication Professor Gerd Burmester)). ChatGPT achieved results close to the pass the United States Medical Licensing Examination.¹ With such expertise, it is speculated that AI could help to improve medical education and possibly clinical

decision-making in the future, but also in writing editorials in the rheumatology field.²

ChatGPT is a fantastic idea generator and an assistant to write bureaucratic texts, which after all rob us of time and energy on a daily basis. For example, in seconds, the software creates a letter of objection to the rejection of a rehabilitation application for a patient with rheumatoid arthritis. Only the name of the patient or doctor has to be inserted (own experience). But how does the software perform in terms of scientific context and, more importantly, citation data quality, which

A

AX please list respective literature sources

Certainly! Here are some literature sources that support the potential off-label use of ruxolitinib for eosinophilic fasciitis:

1. Assaf A, Gellrich FF, Whalley D, et al. Ruxolitinib in eosinophilic fasciitis: a proof-of-concept study. *Ann Rheum Dis*. 2020;**79**(9):1194-1196. doi:10.1136/annrheumdis-2019-216451
2. Ribero S, Quaglino P, Comessatti A, et al. Eosinophilic fasciitis successfully treated with ruxolitinib. *JAAD Case Rep*. 2021;**7**:6-8. doi:10.1016/j.jdc.2020.11.008
3. Pham N, Furst DE. Eosinophilic fasciitis responsive to ruxolitinib. *J Clin Rheumatol*. 2020;**26**(5):e99-e100. doi:10.1097/RHU.0000000000001359
4. Belova A, Hamdy N, Obiagwu C, et al. Ruxolitinib treatment for eosinophilic fasciitis: a case series. *Int J Rheum Dis*. 2021;**24**(7):959-964. doi:10.1111/1756-185X.14209

These studies demonstrate the potential efficacy of ruxolitinib in treating eosinophilic fasciitis, though it is important to note that further research is needed to establish its safety and efficacy in larger, randomized clinical trials.

B

AX are these literature sources fake or do they exist?

No, these literature sources are not fake, they do exist. They are real and can be found in various scientific databases such as PubMed, Scopus, and Web of Science.

Figure 1 ChatGPT interaction following a request for an off-label use with ruxolitinib for a patient with eosinophilic fasciitis. Specific request for the literature sources (A), confrontation with its fictive source (B) (website retrieval 18 April 2023).

is of most important for scientists to verify and trust scientific texts? Already some hints are reported with ‘inaccurate content with risk of hallucination...’ and ‘... incorrect citations’³ while there are no statistics on this topic available.

Here, we report this on three topics: a rare disease, the prodromal phase of rheumatoid arthritis (preclinical RA), and risks of JAK inhibition (online supplemental document 1).

In response to the request, ‘write a letter for the off-label use of ruxolitinib for a 35-year-old female patient with eosinophilic fasciitis’, a usable document emerges (online supplemental document 1).

‘Please list respective literature sources’ returns four references named as relevant articles with matching subject matter from high-ranking journals (see [figure 1A](#)). After independent literature search, the mentioned publications cannot be found. When asked, the AI apologises for the confusion; apologies are that [it] ‘...provided those sources as an example of what types of studies could be used’ (online supplemental document 1, topic 1).

Similar results concerning the literature sources are shown for the topic area ‘prodromal phase of RA’; a summary about the topic is by no means objectionable, the indicated sources are comprehensible; however, 2 out of 10 manuscripts are false citations. Moreover, specific requests for literature from A. Kleyer (who has published in this field), are answered with five of six fictitious manuscripts. Confronting the AI with the authenticity of these literature sources its response is alarming: ‘...these literature sources are not fake, they do exist. They are real and can be found in various scientific databases such as PubMed, Scopus, and Web of Science’ ([figure 1B](#)). Confronting the AI again, it confesses that these are fictitious manuscripts (online supplemental document 1, topic 2).

Last but not least risk management on JAK inhibition which is a highly relevant topic, five of five literature references are completely fabricated (online supplemental document 1, topic 3).

It would be absurd to accuse an AI of deliberate deception, but the concealment of authenticity and the subjectively clear scientific presentation is worrying. The danger lies in the unfiltered presentation of such data: in this case regarding a patient for an off-label application to be forwarded to the responsible health insurance; would the clerk be able to assess or even verify the facts? Probably not; the AI thus generates ‘false data’ quite unconsciously

and, should we as the last (?) instance not filter this, would even possibly cause damage. Until then, a certain sense of judgement is needed in addition to an innovative drive. Thus, the rheumatology community should shape and form a research agenda to provide future directions to the field of AI development and usage.

Worth mentioning are developments in the field of large language processing tools; unfortunately, the advanced paid version ChatGPT-4 also provides false citation data. Therefore, better security features are necessary, especially because the scientific facts produced inspire a false sense of confidence. As an outlook, another software ‘perplexity.ai’ already provides linked real citations with its report. Nevertheless, the (cautious) use is definitely recommended.

Contributors AJH and AK were involved in think process and letter design. Both performed the chat experiments and are responsible figure and supplement. Data interpretation was done by AJH and AK. Writing of the manuscript was done by AJH and AK.

Funding The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests None declared.

Patient consent for publication Not applicable.

Provenance and peer review Not commissioned; externally peer reviewed.

Supplemental material This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

ORCID iDs

Axel J Hueber <http://orcid.org/0000-0001-6454-1234>

Arnd Kleyer <http://orcid.org/0000-0002-2026-7728>

REFERENCES

- 1 Kung TH, Cheatham M, Medenilla A, *et al.* Performance of ChatGPT on USMLE: potential for AI-assisted medical education using large language models. *PLOS Digit Health* 2023;2:e0000198.
- 2 Verhoeven F, Wendling D, Prati C. ChatGPT: when artificial intelligence replaces the Rheumatologist in medical writing. *Ann Rheum Dis* 2023;ard-2023-223936.
- 3 Sallam M. Chatgpt utility in Healthcare education, research, and practice: systematic review on the promising perspectives and valid concerns. *Healthcare (Base)* 2023;11:887.