

# Application ChatGPT in conducting systematic reviews and meta-analyses

By Simpy Amit Mahuli,<sup>1</sup> Arpita Rai,<sup>2</sup> Amit Vasant Mahuli<sup>3</sup> and Ansul Kumar<sup>4</sup>

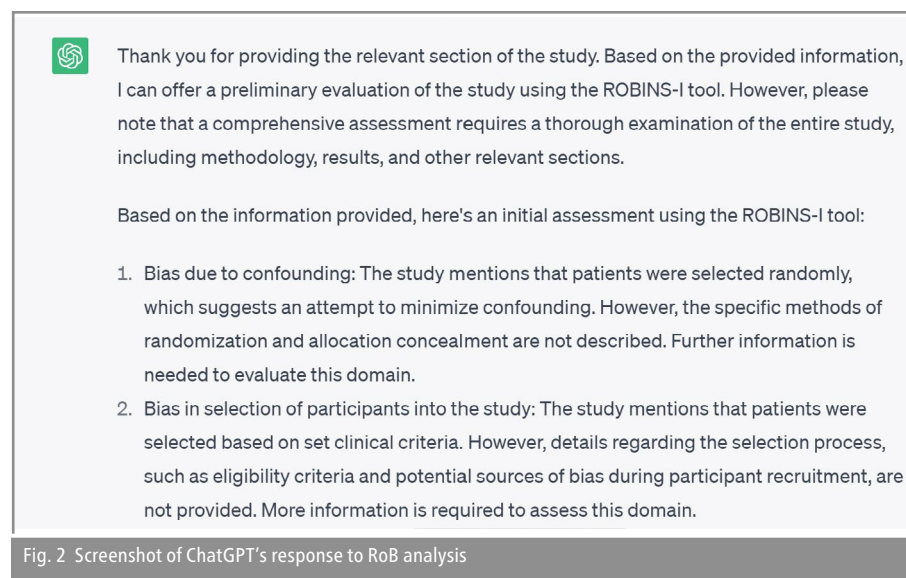
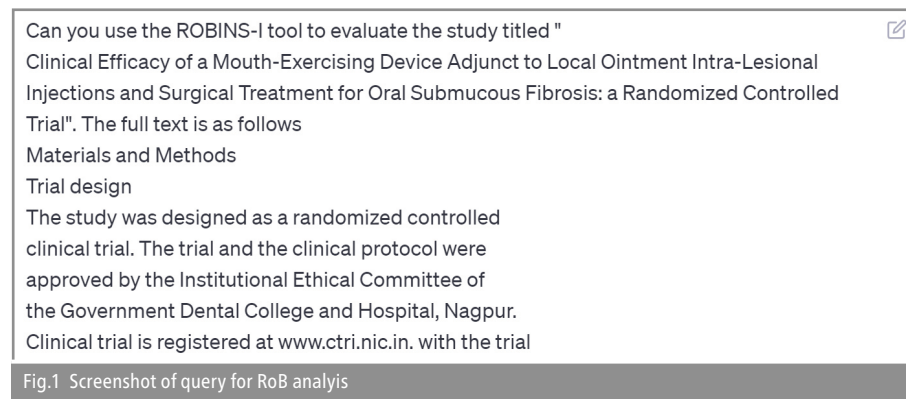
Systematic reviews and meta-analyses, which involve thorough literature searches and synthesis, are essential for evidence-based decision-making. With the growing number of publications, thorough review takes more time. A systematic review often takes more than 15 months to complete, and the long conception-completion interval may render a systematic review outdated by the time they are ready to be submitted and published.<sup>1,2</sup>

Artificial intelligence (AI)-based quick approaches are being developed in response to the growing demand for screening and data extraction techniques that are more effective. These techniques automate decision-making by utilising machine learning (ML) algorithms and natural language processing (NLP) to extract pertinent information from textual material. These AI-based methods have the potential to accelerate the processes of systematic reviews and evidence synthesis while requiring less human labour.<sup>3,4</sup>

ChatGPT is a Large Language Model (LLM) developed by OpenAI (OpenAI, L.L.C., San Francisco, CA, USA) which has tremendous potential to revolutionise the healthcare industry. ChatGPT has been effectively used for a variety of tasks in healthcare, including conducting literature reviews, analysing datasets, writing academic papers, documentation, and enhancing clinical workflow. ChatGPT is a significant advancement in the field of NLP. It generates human-like text by mimicking human language processing abilities using deep learning techniques and neural networks. Owing to its training on diverse datasets, it is able to provide pertinent responses to human prompts by processing large amounts of information. The easy accessibility and free availability have contributed to the success ChatGPT version 3.5 platform.<sup>5</sup>

We tried to use ChatGPT for conducting Risk of Bias analysis and data extraction from a randomised controlled trial. The response of ChatGPT was surprising, to say the least. Conducting such processes through ChatGPT is simple and straightforward.

Firstly, to conduct a ROB analysis effectively using AI, obtaining the full-text version of the article under evaluation is crucial. This ensures that all relevant information and details necessary for the assessment are available. As an AI tool, it does not have real-time browsing capabilities, so it cannot search the internet or access the latest research articles. Additionally, it being a text-based AI, we cannot upload or import text/pdf files into the chatbox. However, ChatGPT can still try to assist us with general questions or provide guidance based on the information we provide from the study. If we have any specific questions or excerpts from the study that we would like ChatGPT to analyse or discuss, we can paste them into the chat box, and ChatGPT will do its best to assist. ►►



<sup>1</sup>PhD Scholar, Rajendra Institute of Medical Sciences (RIMS), Bariatu, Ranchi – 834009, India; <sup>2</sup>Associate Professor, Oral Medicine and Radiology, Rajendra Institute of Medical Sciences (RIMS), Bariatu, Ranchi – 834009, India; <sup>3</sup>Associate Professor, Public Health Dentistry, Rajendra Institute of Medical Sciences (RIMS), Bariatu, Ranchi – 834009, India; <sup>4</sup>Associate Professor, Cardiothoracic and Vascular Surgery, Dental College, Rajendra Institute of Medical Sciences, Bariatu, India

Secondly, we need to provide clear and specific instructions to the tool for it to provide an appropriate response. Hence when prompted ‘Can you use the ROBINS-I tool to evaluate “(Materials and methods from the full-text article)”’, it generated a detailed response highlighting the various lacunae in the methodology in context of the Biases included in the ROBINS-I tool (Fig. 1 and Fig. 2). This assessment does not eliminate the need for a human expert; however, it considerably reduces the time and effort required for the process. Also, AI provides validation to the information obtained through human effort and may even catch some details that are missed by the human eye. Additionally, ChatGPT is extremely helpful in the interpretation of studies for those who are not so fluent in the language. While ChatGPT can comprehend and generate text in languages other than English to some extent, its proficiency may vary across different languages. It’s worth noting that ChatGPT’s performance is generally more robust in English, as the majority of its training data is in English.

Next, we tried to use ChatGPT for data extraction. By sharing the complete text and specifying the desired information or data to be extracted, AI can be of assistance. Hence, following various prompts from us, ChatGPT extracted data like mean values, standard deviation and sample size for study groups and sub-groups and presented it. Also, it was able to extract information from the tables provided, though it is not possible to paste a table in the chatbox. We can paste the data in the table en masse into the chatbox and it was able to interpret and extract information from it (Fig. 3 and Fig. 4). Needless to say, this will reduce the time required for data extraction as well as ensure accuracy of the values, thereby reducing human errors.

AI methods are increasingly integrated into healthcare reviews to streamline processes. A recent systematic review by Blaizot *et al.* identified and examined automated tools and platforms that utilise artificial intelligence (AI) techniques. It assessed the pros and cons associated with the use of these methods in conducting systematic reviews. They identified various tools like Rayyan, Robot Reviewer, EPPI-reviewer, SWIFT-review, and Abstracker which were employed for different review stages. Most included reviews

Can you do data extraction for groupwise Mouth opening scores from the given text?After six months of recall visits, average improvement in IID in the subgroup A1 (n=53) was 8.4 mm and in subgroup B1 (n=50) was 5.5 mm (p<0.01) (Table 3 and 4). In patients treated with intra-lesional injections, average improvement in IID for subgroup A2 (n=46) was 9.3 mm and for subgroup B2 (n=48) was 5.1 mm (p<0.01) (Table 3 and 4). In surgery group, the improvement was measured in comparison with the immediate post-surgical opening, and the average improvement after six month recall was found to be 9.6 mm for subgroup A3(n=18) and 4.8 mm for subgroup B3(n=16) (p<0.01) (Table 3 and 4).

Fig. 3 Screenshot of query for data extraction from text

Certainly! Based on the provided text, here is the data extraction for groupwise mouth opening scores:

Group A (Device Users):

- Subgroup A1: Improvement in IID (Inter-incisal Distance) after 6 months recall: 8.4 mm
- Subgroup A2: Improvement in IID after 6 months recall: 9.3 mm
- Subgroup A3: Improvement in IID after 6 months recall: 9.6 mm

Group B (Non-Device Users):

- Subgroup B1: Improvement in IID after 6 months recall: 5.5 mm
- Subgroup B2: Improvement in IID after 6 months recall: 5.1 mm
- Subgroup B3: Improvement in IID after 6 months recall: 4.8 mm

Can you tell me the number of participants in each group and subgroup?

Certainly! Based on the provided text, here is the information on the number of participants in each group and subgroup:

Group A (Device Users):

- Subgroup A1: n = 53 participants
- Subgroup A2: n = 46 participants
- Subgroup A3: n = 18 participants

Group B (Non-Device Users):

- Subgroup B1: n = 50 participants
- Subgroup B2: n = 48 participants
- Subgroup B3: n = 16 participant

Can u extract the mean values for each subgroup from the given data?  
Device users (Group A) Non Device users (Group B)  
A1 A2 A3 B1 B2 B3  
Mean 8.4038 9.263 9.6167 5.51 5.1438 4.8313  
SD 1.7174 1.6329 1.3853 1.4376 1.4057 1.3671

Certainly! Based on the provided data, here are the mean values for each subgroup:

Device Users (Group A):

- Subgroup A1: Mean = 8.4038
- Subgroup A2: Mean = 9.263
- Subgroup A3: Mean = 9.6167

Non-Device Users (Group B):

- Subgroup B1: Mean = 5.51
- Subgroup B2: Mean = 5.1438

Fig. 4 Response from ChatGPT for various queries for data extraction

► acknowledged that screening with AI methods generally demonstrates low risks and enhances efficiency. However, possibility of missing relevant articles and

evolving through regular updates and fine-tuning based on user feedback and new training data, enhancing its performance and accuracy over time.<sup>5</sup>

'AI provides validation to the information obtained through human effort and may even catch some details that are missed by the human eye.'

issues in data extraction required human intervention.<sup>4</sup>

ChatGPT possesses several distinct advantages in comparison to other AI tools. Its ability to retain context and coherence during conversations equips it to comprehend and respond to complex queries. It is extremely versatile which is demonstrated through its proficiency in handling various tasks, including factual inquiries, explanatory responses, creative content generation, and problem-solving assistance. ChatGPT is continuously

It is crucial to acknowledge that ChatGPT also has limitations. These include the potential for generating incorrect or biased responses and its reliance on pre-existing data. Users should exercise critical thinking and independently verify information obtained from any AI tool. Currently, it is unable to process visual information; however, this issue has been addressed in ChatGPT 4.0. Additionally, the model's output heavily relies on user input; minor changes in the query can lead to significant alterations in the response.<sup>5</sup>

This was the first attempt to test the feasibility of using ChatGPT in the tedious task of conducting systematic reviews and meta-analyses. It shows promise in reducing workload and time, but careful implementation and validation are necessary. We as researchers need to interact more with this tool to completely understand its applicability in generating evidence. ■

#### References

1. Borah R, Brown A W, Capers P L, Kaiser K A. Analysis of the time and workers needed to conduct systematic reviews of medical interventions using data from the PROSPERO registry. *BMJ Open* 2017; doi: 10.1136/bmjopen-2016-012545.
2. Yaffe J, Montgomery P, Hopewell S, Shepard L D. Empty reviews: a description and consideration of Cochrane systematic reviews with no included studies. *PLoS One* 2012; doi: 10.1371/journal.pone.0036626.
3. Chollet F. *Deep learning with Python*. pp 361. Shelter Island, New York: Manning Publications Co, 2018.
4. Blaizot A, Veettil S K, Saidoung P, Moreno-Garcia C F *et al*. Using artificial intelligence methods for systematic review in health sciences: A systematic review. *Res Synth Methods* 2022; **13**: 353–362.
5. OpenAI. Introducing ChatGPT. Available at: <https://openai.com/blog/chatgpt> (accessed 19 May 2023).

## BSPD President supports Free School Meals for All

Dr Jenny Harris, President of the British Society of Paediatric Dentistry (BSPD) (pictured), joined a group of supporters of the Free School Meals for All campaign outside Downing Street on 29 June 2023.

The 'No Child Left Behind' group responsible for the Free School Meals for All initiative delivered its open letter, signed by Dr Harris on behalf of the Society, calling for an extension of free school meals to all children in primary schools in England. Primary schools in England do not currently receive funding for Free School Meals for All, despite the scheme being announced in Scotland, Wales and most recently in London as part of a one-year scheme to support families with the cost-of-living crisis.

The open letter, signed by more than 240 community organisations, faith groups, charities and trade unions, and over 90,000 individuals, came on the penultimate day of the Free School Meals for All campaign's National Week

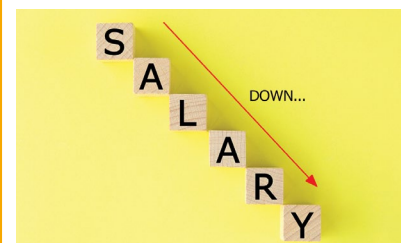


of Action (24–30 June), which has a series of local and national events and activities. The Free School Meals for All campaign is supported by nearly 90 MPs, Peers, local councils and mayors, who are putting their weight behind the call to ensure no child goes too hungry to learn at school.

A full list of organisations supporting the campaign can be found at: <https://freeschoolmealsforall.org.uk/take-action>.

## Pay cut for uniformed dentists

British Medical Association (BMA) Armed Forces Committee Chair Colonel Mark Weir and British Dental Association (BDA) Armed Forces Committee Chair Surgeon Captain (D) (Retired) Mike Gall have responded with dismay to the pay award announced by the Government for doctors and dentists in the Armed Forces, which is 5% for 2023/24 and a £1,000 consolidated sum, and said:



'Once again there has been universal failure from the review body to reflect the detailed recommendations made by the BMA and the BDA and the result is a pay cut in real terms during an extended inflationary period.'