Review of: "How many papers are published each week reporting on trials of interventions involving behavioural aspects of health?"

Aliya Amirova[1]

1 King's College London

Thank you for inviting me to review this preprint. I support the open review initiative and the efforts of Qeios in promoting open science.

I thank the authors for conducting this review. It is fantastic to see the scale at which behavioural research is produced. I am very glad to see how prolific our field is today.

This bibliometric analysis estimating the number of studies reporting behavioural interventions relating to health is conducted accurately. As the authors pointed out, we cannot know the exact number of studies, but this bibliographic analysis provides the reader with an estimate.

I agree with the authors, that given this huge number of studies, some automation is required for efficient and timely integration of evidence and cumulative science. However, I am not sure if that implies that there is a need for an authoring tool to facilitate machine readability. It might be the case that currently existing large language models (e.g., GPT-3) are already capable of conducting human-level (and beyond) automated literature search and data extraction. These models do not need structured language. Perhaps today machine readability is no longer an issue since machines can read natural language.

I wonder if the authors could comment on the existence of such language models and their implications for the need to report studies in a more structured way.

Please see these papers:

- https://arxiv.org/abs/2108.07258
- Bommasani, R., Hudson, D. A., Adeli, E., Altman, R., Arora, S., von Arx, S., ... & Liang, P. (2021). On the opportunities and risks of foundation models. *arXiv preprint arXiv:2108.07258*.
- Ouyang, L., Wu, J., Jiang, X., Almeida, D., Wainwright, C. L., Mishkin, P., ... & Lowe, R. (2022). Training language models to follow instructions with human feedback. *arXiv preprint arXiv:2203.02155*.
- https://arxiv.org/abs/2203.02155