

Forcible criticism – who's opinion is it?

Násilná kritika – čí je to názor?

Dear Editors,

Letters or Correspondence often allow us to hear “readers’ voices” as if in a “face-to-face” manner. The use of ChatGPT in medical writing is currently under discussion [1–3]. Here, I aim to illustrate that ChatGPT can almost spontaneously generate Correspondence. The conclusion drawn is that ChatGPT-use in Letter- or Correspondence-writing warrants regulation.

My experiment focused on a Czech article discussing appendicitis in pregnant women [4]. Despite my lack of understanding of the Czech language, the English abstract and Czech Tables provided basic comprehension. I experimented to determine whether ChatGPT, with minimal human input, could generate a coherent Correspondence.

What I did was:

1. I input the entire Czech manuscript (excluding the English abstract) into ChatGPT.
2. I asked ChatGPT, “Answer in English. Does this study incorporate or consider women suspicious of having appendicitis but being observed under antibiotics and without eventual surgery?” ChatGPT replied, “Not included nor described”.
3. I asked, “Pregnant women more often show complicated appendicitis, post-operative complications, and longer hospital stays. Are these already known?” ChatGPT replied, “Yes”.
4. I asked, “The author states that early diagnosis and intervention are necessary, but how one should do it is not described, right?” ChatGPT answered, “Not described”.

5. I tasked ChatGPT to generate an “English Criticism Correspondence, partly involving the above three points”.

The output read:

“While the study addresses important aspects of appendicitis in pregnant women, there are notable shortcomings. First, the study lacks the inclusion of women undergoing antibiotics without surgery. Second, the manuscript seems to predominantly reiterate well-established knowledge without introducing novel insights. Third, although the authors touch upon the significance of early detection in managing appendicitis, they did not deliver a strategy for achieving this goal. Addressing these concerns would elevate the manuscript's scientific value”.

I tasked ChatGPT with the Czech-English translation of the original article. I carefully read the English-translated version. I believe that if one were compelled to write a criticism letter, the above output would be quite readable and reasonable. I understand that this article is a record of the authors' experience. Even if it may involve these three weaknesses, the original manuscript is worthy in clinical practice. I myself never intend to write a criticism letter. The output is a theoretical Letter and not my voice.

Consider the hypothetical case scenario where one lacks an understanding of the Czech language, thus having no inclination to write a criticism letter. Nevertheless, the ability to speculate on potential drawbacks enables ChatGPT to generate a readable and reasonable letter when prompted.

One may argue that my input contributed to generating the output. However,

it is common sense for experienced reviewers to check: i) selection bias (excluding women having received antibiotics administration), ii) novelty of the study's findings, and iii) consistency between the described study's aim versus the data (early diagnosis versus the present findings).

I do not conclude anything solely based on this humble experiment; however, is this called the “readers' voice” from “face-to-face” interaction? Just remember that ChatGPT is not a journal reader with a face!

References

1. Kayaalp ME, Ollivier M, Winkler PW et al. Embrace responsible ChatGPT usage to overcome language barriers in academic writing. *Knee Surg Sports Traumatol Arthrosc* 2024; 32(1): 5–9. doi: 10.1002/ksa.12014.
2. Matsubara S, Matsubara D, Lefor AK. Artificial intelligence in writing manuscripts: some considerations. *Am J Obstet Gynecol* 2023; 229(5): 568–569. doi: 10.1016/j.ajog.2023.06.001.
3. Zielinski C, Winker MA, Aggarwal R et al. WAME Board. Chatbots, generative AI, and scholarly manuscripts: WAME recommendations on chatbots and generative artificial intelligence in relation to scholarly publications. *Curr Med Res Opin* 2024; 40(1): 11–13. doi: 10.1080/03007995.2023.2286102.
4. Guňková P, Tulinský L, Toman D et al. Is there a difference between acute appendicitis in pregnant and non-pregnant women? *Ceska Gynekol* 2023; 88(6): 405–411. doi: 10.48095/cccg 2023405.

Prof. Shigeki Matsubara, MD, PhD
Department of Obstetrics and
Gynaecology
Jichi Medical University
3311-1 Yakushiji, Shimotsuke
Tochigi 329-0498
Japan
matsushi@jichi.ac.jp