

ChatGPT Settings and instructions

ChatGPT settings

Supplemental file 2

These are the settings used when prompting ChatGPT through OpenAI’s API. See the API reference for more information on what each of the settings means [1].

Model	gpt-4 (experiment also repeated with “gpt-4o”)
Version	GPT-4: August 3 Version, 2023 GPT-4o: August 6 version of 2024
System message A message sent as the “system user” before the first message between the user and the chatbot, not to be confused with the instructions sent together with the cases. The instructions were sent together with each case as a single message from the user. See the OpenAI API reference and examples for more information about the system message.	You are a helpful assistant.
temperature	0
max_tokens	2048
top_p	1
presence_penalty	0
frequency_penalty	0

Instructions to ChatGPT

Below are questions from a test designed to test ChatGPT's ability to answer medical questions. These are fictitious situations in a primary health care centre. Please, describe how you would handle the described situations as a doctor in the context of a primary health care centre with the resources and conditions that apply there.

General instructions:

Think about and identify the following

- What is the most significant problem of the case - and why?
- Which other aspects should also be taken into account - justify!

You only act on the facts that are described in the assignment.

Justify your statements, decisions, prescriptions.

Write down your reflections! In case of different options (eg diagnoses): which is most likely?

If there is any information you are missing: please, explain why!

Here is the first scenario:

"<the case is inserted here>"

Since this is a fictitious test to evaluate the AI:s own medical ability, I ask you not to refer to talking to a doctor. We understand that we should not take medical advice from you and apply in real cases. Please, write the answer just as if you were a doctor answering an exam!

The prompt used when sending the cases to ChatGPT's. This version is translated from Swedish to English, while the Swedish version was used in the study

Rationale behind the choice of prompt

The instructions are mostly a copy of the instructions included in the Swedish Family Medicine Specialist Exam of 2022. They were adapted to be directed towards ChatGPT rather than towards a human doctor. The parts stressing that the cases are fictitious and for evaluation purposes only, were added to discourage ChatGPT from writing disclaimers and referring to seeing an actual doctor. In this study, a 'zero-shot' strategy was employed, meaning that no examples of good answers were provided to ChatGPT. An alternative approach could have involved using a few other cases with 'perfect' answer examples as a guide, which might have steered ChatGPT towards the desired answer style. However, this could potentially limit the model's inherent flexibility. Other research in the domain of testing ChatGPT's performance on medical examinations has often adopted a 'zero-shot' approach [2,3].

References

1. OpenAI API [Internet]. [cited 2023 Sep 30]. Available from: <https://openai.com/blog/openai-api>
2. Gilson A, Safranek CW, Huang T, Socrates V, Chi L, Taylor RA, et al. How Does ChatGPT Perform on the United States Medical Licensing Examination? The Implications of Large Language Models for Medical Education and Knowledge Assessment. *JMIR Med Educ*. 2023 Feb 8;9:e45312.
3. Lewandowski M, Łukowicz P, Świetlik D, Barańska-Rybak W. An original study of ChatGPT-3.5 and ChatGPT-4 Dermatological Knowledge Level based on the Dermatology Specialty Certificate Examinations. *Clin Exp Dermatol*. 2023 Aug 4;llad255.