We recently read with great interest the article by Sarangi and colleagues, published in the December 2023 issue of the Indian Journal of Radiology and Imaging that assessed ChatGPT, Google Bard, and Microsoft Bing’s diagnostic abilities in solving radiology vignettes.¹ The creative application of large language models by the authors holds promise for this field in radiology. We are grateful for their important contribution. The study’s overall methodology and conclusions are excellent; however, some methodological issues raised questions in our minds and made us look for further information.

Prompt selection may have an impact on large language models’ (LLMs) diagnostic success. Caruccio and colleagues demonstrated how varying prompts, ranging from simple to complex, might alter the chatbots’ performance.² First, could the authors kindly provide more details on the particular input prompt they used in their research? Second, have the chatbots been trained by the authors?

With its significant implications for integrating LLMs, we sincerely appreciate the chance to discuss this groundbreaking research. We would better grasp the complexities of the study if the previously mentioned points were clarified.

**Ethical Approval**
This article does not contain any studies with human participants or animals performed by any of the authors.

**Funding**
None.

**Conflict of Interest**
None declared.

**References**
