

Letters

COMMENT & RESPONSE

Context for Large Language Model Evaluation in Ophthalmology

To the Editor We read with interest the recent article by Srinivasan et al,¹ which provides insights into the performance of large language models in ophthalmology, particularly the OpenAI o1 model. We wish to raise a minor point regarding the temperature parameter settings. The study notes that the o1 model was evaluated using its default temperature, whereas the other models were set to a temperature of 0. This discrepancy in temperature settings may give the implied framing that the authors intentionally adopted different decoding strategies to highlight the o1 model's reasoning capabilities. However, to our knowledge, o1 does not allow users to modify decoding parameters such as temperature—a capability available for the other evaluated models. Therefore, it seems possible that the use of o1's default temperature was not a conscious design choice but

rather a limitation imposed by the model's application programming interface at the time. We would appreciate clarification from the authors on this point, and we thank them for their valuable contribution to the field.

Yusuke Kameda, MD, PhD

Yutaka Kaneko, MD, PhD

Author Affiliations: Yotsuya-sanchome Ekimae Eye Clinic, Tokyo, Japan (Kameda); Department of Ophthalmology and Visual Sciences, Yamagata University Faculty of Medicine, Yamagata, Japan (Kaneko).

Corresponding Author: Yusuke Kameda, MD, PhD, Yotsuya-sanchome Ekimae Eye Clinic, 3-7-24 Yotsuya, Shinjuku-ku Tokyo 160-0004, Japan (y09025618059@leaf.ocn.ne.jp).

Published Online: January 22, 2026. doi:10.1001/jamaophthalmol.2025.5841

Conflict of Interest Disclosures: None reported.

1. Srinivasan S, Ai X, Zou M, et al. Ophthalmological question answering and reasoning using OpenAI o1 vs other large language models. *JAMA Ophthalmol.* 2025;143(9):740-748. doi:10.1001/jamaophthalmol.2025.2413