

Letter to the Editor

Authors' Reply: Beyond Visual Consensus: Tiered Reference Framework for AI Cystoscopy Studies

Chung-You Tsai^{1,2,3}, MD, PhD; Shi-Wei Huang^{3,4}, MD, PhD

¹Division of Urology, Department of Surgery, Far Eastern Memorial Hospital, New Taipei, Taiwan

²Department of Electrical Engineering, Yuan Ze University, Taoyuan, Taiwan

³Department of Urology, College of Medicine, National Taiwan University, Taipei, Taiwan

⁴Department of Urology, National Taiwan University Hospital, Yunlin Branch, Yunlin, Taiwan

Corresponding Author:

Shi-Wei Huang, MD, PhD

Department of Urology

National Taiwan University Hospital, Yunlin Branch

No. 579, Sec. 2, Yunlin Rd., Douliou City

Yunlin

Taiwan

Phone: 886 55323911

Email: will6438.huang@gmail.com

Related Articles:

Comment on: <https://www.jmir.org/2026/1/e87193>

Comment on: <https://www.jmir.org/2026/1/e101910>

J Med Internet Res 2026;28:e103335; doi: [10.2196/103335](https://doi.org/10.2196/103335)

Keywords: multimodal; large language model; AI; cystoscopy; diagnostic reasoning; finding description; biopsy indication; bladder tumor; artificial intelligence

We thank the reader [1] for their thoughtful and constructive comments on our study [2]. We would like to clarify an important methodological point regarding the reference standard used in our study.

The 401 cystoscopic images were curated from established educational and clinical sources, including reference atlases, PubMed-indexed articles and repositories, clinical websites, industry archives, and Creative Commons–licensed educational videos. These images were originally accompanied by source-provided diagnoses, educational labels, captions, or contextual information.

As stated in the Study Design section, “The reference standard diagnoses were determined through a multiphase, consensus-based process.” To further clarify how source-provided diagnostic information was incorporated, this process can be described in two key phases. In the initial independent review phase, source-provided diagnostic labels were withheld, and two urological experts independently inspected each image and answered the prespecified Q1 to Q5 framework, including anatomic site, cystoscopic findings, lesion detection, lesion reasoning, and final diagnosis. In the subsequent consensus phase, the available source-provided diagnoses or educational labels were disclosed to the experts and considered together with the cystoscopic appearance

to resolve discrepancies and establish the final reference answers for Q1 to Q5.

Therefore, the final consensus diagnosis for each image was not based solely on independent visual inspection by two urologists. Rather, the reference standard was established through a source-informed expert consensus process, integrating source-provided diagnostic information with expert review of the cystoscopic appearance.

Thus, the reference standard in our study was designed to support the evaluation of model-generated cystoscopic interpretation and reasoning, rather than to independently re-establish or revalidate the original pathological diagnosis of each image. We agree that histopathology remains the definitive standard for pathological lesion classification, particularly for entities such as carcinoma in situ, papilloma, and papillary urothelial carcinoma. Accordingly, our study should be read as an image interpretation and reasoning benchmark based on images with pre-existing source-provided diagnostic or educational information, with final reference labels established through expert consensus after consideration of both the cystoscopic appearance and the diagnoses provided by the original educational or clinical sources.

Acknowledgments

Generative artificial intelligence assistance was used for language editing. The authors reviewed, revised, and approved the final content and take full responsibility for the submitted manuscript.

Funding

The authors declared no financial support was received for this work.

Conflicts of Interest

None declared.

References

1. Bayraktar AM, İşler B. Beyond visual consensus: tiered reference framework for AI cystoscopy studies. *J Med Internet Res*. 2026;28:e101910. [doi: [10.2196/101910](https://doi.org/10.2196/101910)]
2. Shih YC, Wu CY, Huang SW, Tsai CY. Multimodal large language models for cystoscopic image interpretation and bladder lesion classification: comparative study. *J Med Internet Res*. Jan 28, 2026;28:e87193. [doi: [10.2196/87193](https://doi.org/10.2196/87193)] [Medline: [41605505](https://pubmed.ncbi.nlm.nih.gov/41605505/)]

Edited by Tiffany Leung; This is a non-peer-reviewed article; submitted 03.Jun.2026; accepted 04.Jun.2026; published 18.Jun.2026

Please cite as:

Tsai CY, Huang SW

Authors' Reply: Beyond Visual Consensus: Tiered Reference Framework for AI Cystoscopy Studies

J Med Internet Res 2026;28:e103335

URL: <https://www.jmir.org/2026/1/e103335>

doi: [10.2196/103335](https://doi.org/10.2196/103335)

© Chung-You Tsai, Shi-Wei Huang. Originally published in the Journal of Medical Internet Research (<https://www.jmir.org>), 18.Jun.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research (ISSN 1438-8871), is properly cited. The complete bibliographic information, a link to the original publication on <https://www.jmir.org/>, as well as this copyright and license information must be included.