Correspondence

INTERNATIONAL JOURNAL OF SURGERY

OPEN

Towards clinically robust AI for CcRCC nuclear grading perspectives on validation, imaging standardization, and prospective translation

Jiale Wang, MD, Hongjin Shi, MD, Jinsong Zhang, MD*

Dear Editor,

Artificial intelligence is increasingly being integrated into clinical practice. A recent study by Zheng *et al*, titled "Artificial Intelligence-Based Multi-Modal Prediction for Nuclear Grading Status and Prognosis of Clear Cell Renal Cell Carcinoma: A Multicenter Cohort Study," has garnered significant attention[1]. The multimodal AI model proposed in this study demonstrates improved accuracy in predicting the nuclear grading status of clear cell renal cell carcinoma, and its contributions are widely recognized. Nevertheless, we believe that several key issues warrant further investigation to advance research in this field and facilitate its clinical application.

Firstly, the training set and internal validation set of this model are both derived from Chinese population data, which may confer population-specific characteristics to the model. The external validation set was constructed using 217 cases from the TCGA and CPTAC databases, with the CPTAC cohort comprising only 42 cases (19.4% of the external validation set). Such limited representation may be insufficient to comprehensively assess the model's generalization capacity, particularly regarding its predictive performance for rare subgroups or clinical outcomes, thereby potentially compromising its robustness. Therefore, future studies should prioritize expanding multi-center collaborations, augmenting validation datasets, and conducting model training and evaluation across diverse ethnic populations to enhance their generalizability and broad applicability. Secondly, during the analysis of CTU images, although a consistent methodology (Common Medication Practice, CMP) was employed for feature extraction from regions of interest (ROIs), variations in CT scanning protocols across different centers were not adequately addressed. This discrepancy may lead the artificial intelligence system to misinterpret technical artifacts as biological patterns, thereby introducing potential bias into the research outcomes. It is therefore recommended that standardized preprocessing be implemented on retrospective multi-center CT data to harmonize and normalize the datasets. Furthermore, prior to initiating prospective multi-center cohort studies, a unified CT scanning protocol should be established to mitigate the impact of inter-center technical variability on model accuracy. Thirdly, regions of interest (ROIs) are currently delineated manually on CTU images, a process that is not only time-consuming but also susceptible to subjective variability. To mitigate these limitations, artificial intelligence has the potential to enable automated segmentation of tumor regions in CTU images, with human oversight limited to quality assurance tasks. Such an approach would enhance both operational efficiency and measurement consistency[2,3]. Additionally, the study period spans a considerable duration (2016–2024), and the potential influence of technological advancements in imaging equipment and improvements in computational data processing capabilities—such as enhanced image resolution—on the interpretation of historical data has not been adequately considered, which may introduce temporal bias. Fifthly, regarding clinical translation, the authors assert that the model has the potential to provide real-time support for intraoperative decision-making. However, the model was developed

HIGHLIGHTS

- Limited External Validation Scope: Model generalizability constrained by small external validation set (n = 217), particularly the CPTAC cohort (n = 42, 19.4%), and reliance on predominantly Chinese population data, raising concerns about performance in diverse ethnicities and rare subgroups.
- Technical Variability & Temporal Bias Risk: Unaddressed differences in multi-center CT scanning protocols may lead AI to misinterpret technical artifacts as biological signals. Significant study period (2016-2024) risks temporal bias due to unaccounted advancements in imaging technology and computational power.
- Manual ROI Delineation Inefficiency: Current dependency on time-consuming, subjective manual segmentation of Regions of Interest (ROIs) on CTU images hinders scalability and consistency; AI-powered automation is a critical unmet need.
- Unvalidated Clinical Translation Claim: Asserted potential for real-time intraoperative decision support lacks prospective validation; model developed solely on retrospective data, with performance and speed unproven in actual surgical (e.g., frozen section) settings.

Department of Urology, The Second Affiliated Hospital of Kunming Medical University, Yunnan, China

**Corresponding author. Address: Department of Urology, The Second Affiliated Hospital of Kunming Medical University, Yunnan 650032, China. Tel.: +86 13987681926. E-mail: 945933392zjs@sina.com (J. Zhang).*

International Journal of Surgery (2025) 111:9924–9925

Received 24 July 2025; Accepted 27 July 2025

Published online 6 August 2025

http://dx.doi.org/10.1097/JS9.0000000000003183

exclusively using retrospective pathological images, and its efficacy has not been validated through prospective clinical cohort studies. Further research is warranted to conduct prospective trials in intraoperative frozen section settings in order to assess the model's performance in real-time tumor grading and enhance its processing speed. Finally, enhancing patient education is essential to improve adherence to follow-up protocols and minimize data attrition. The implementation of these strategies will contribute to the establishment of a robust evidence base for clinical practice and support the effective translation of research findings into practical applications.

We sincerely appreciate the valuable research conducted by Zheng et al., which, through innovative technical approaches and analytical models, has established a precise assessment system for nuclear grade status and prognosis in clear cell renal cell carcinoma. This work provides an important practical pathway for optimizing patient outcomes and the efficient allocation of clinical resources, thereby facilitating the implementation of individualized diagnosis and treatment strategies. Concurrently, it is hoped that the authors will thoroughly consider and incorporate these recommendations in future research: expanding multi-center collaboration, standardizing CT scanning protocols across institutions, and applying artificial intelligence for image segmentation to improve the generalizability and accuracy of the findings. Finally, prospective trials should be conducted in intraoperative settings to validate and enhance the model's practical applicability. Implementing these measures will further strengthen the study's scientific rigor and its potential contributions to academic knowledge and societal benefit. All this was noted in accordance with the TITAN guidelines, which mandate transparency when reporting artificial intelligence use[4].

Ethical approval

None.

Consent

None.

Sources of funding

None.

Author contributions

J.W., H.S.: data curation, investigation, writing – original draft. J.Z.: writing – review and editing, supervision, and project administration. J.W. and H.S. made equal contributions to this study.

Conflicts of interest disclosure

The authors declare that they have no conflicts of interest.

Guarantor

Jinsong Zhang.

Research registration unique identifying number (UIN)

None.

Provenance and peer review

Not commissioned.

Data availability statement

Data sharing is not applicable to this article as no new data were created or analyzed in this study.

References

[1] Zheng Q, Mei H, Weng X, *et al*. Artificial intelligence-based multimodal prediction for nuclear grading status and prognosis of clear cell renal cell carcinoma: a multicenter cohort study. Int J Surg 2025;111:3722–30.
[2] Lai C, Hu Z, Zhu J, *et al*. Development and validation of a deep learning-based automated computed tomography image segmentation and diagnostic model for infectious hydronephrosis: a retrospective multicentre cohort study. EClinicalMedicine 2025;82:103146.
[3] Karunanayake N, Lu L, Yang H, *et al*. Dual-stage AI model for enhanced CT imaging: precision segmentation of kidney and tumors. Tomography 2025;11:3.
[4] Agha RA, Mathew G, Rashid R, *et al*. Transparency in the reporting of artificial intelligence: the TITAN guideline. Prem J Sci 2025;10:100.