

Letter to the Editor

Same-Patient, Same-Time Point Evidence for Radiomics Benchmarking

Wenze Kan^{1,2}, PhD; Haoliang Pei^{1,2}, PhD; Mingxin Zhang³, PhD

¹School of Clinical Medicine, Jiamusi University, Jiamusi, China

²Department of Radiology, The First Affiliated Hospital of Jiamusi University, Jiamusi, China

³Department of Radiology, Zaoyang First People's Hospital, Zaoyang, Hubei, China

Corresponding Author:

Mingxin Zhang, PhD
Department of Radiology
Zaoyang First People's Hospital
183 Zhongxing Avenue
Zaoyang, Hubei 441200
China
Phone: 86 17346379074
Email: zhangmx199908@163.com

Related Articles:

Comment on: <https://www.jmir.org/2026/1/e93892>

Comment in: <https://www.jmir.org/2026/1/e110607>

J Med Internet Res 2026;28:e110058; doi: [10.2196/110058](https://doi.org/10.2196/110058)

Keywords: radiomics; non-small cell lung carcinoma; neoadjuvant therapy; pathological response; meta-analysis; NSCLC

Jiang et al [1] provide a timely synthesis of studies on radiomics-based AI for predicting a pathological response after neoadjuvant immunochemotherapy in resectable non-small cell lung cancer. The pooled discrimination is encouraging. However, the comparison architecture and clinical estimand support a narrower conclusion than comparative superiority or surgery-modifying utility.

For pathological complete response (pCR), Table S11 in their paper [1] contrasts 16 AI validation datasets (n=1031 patients) with 2 conventional-criteria datasets (reported sample size: n=72). Table S7, however, reports PERCIST (Positron Emission Tomography Response Criteria in Solid Tumors) and RECIST 1.1 (Response Evaluation Criteria in Solid Tumors version 1.1) results in the same Deng et al [2] cohort, each totaling 36 participants; the corresponding AI entry in Table S5 is the separate 8-patient external-validation cohort [1]. Thus, the reported denominator sums 2 correlated test assessments rather than 72 unique participants and compares them with unmatched AI cohorts; this does not estimate relative accuracy [3]. The direction is also a trade-off: AI sensitivity was higher (0.77 vs 0.42), whereas specificity was lower (0.79 vs 0.97). The reported Z tests therefore cannot establish overall superiority without same-patient data and prespecified costs for false-positive and false-negative decisions.

Clinical use is further constrained by the target population. Eligibility required neoadjuvant therapy followed by radical resection, while included models used pretreatment, during-treatment, or pre-to-posttreatment imaging. The meta-analysis therefore conditions the pathological reference standard on reaching surgery and pools predictions made at different decision points. Patients who did not reach surgery because of progression, toxicity, or unresectability—the group most relevant to early redirection—are absent. In the CheckMate 816 trial [4], by contrast, pCR required no viable tumor in both lung and sampled nodes, and patients without surgery were counted as nonresponders in the intention-to-treat analysis. The pooled 68% positive posttest probability therefore does not establish readiness to guide surgical timing or organ preservation.

Two clinical examples also require correction. First, a 25% diameter decrease is stable disease under RECIST 1.1, not a partial response, which requires at least 30% [5]. Second, if pCR or a major pathological response is the positive condition, fibrosis mistaken for residual tumor creates a false negative, not a false positive.

A useful next step would be a treatment-initiation cohort with a locked model, fixed imaging time and threshold, explicit primary-tumor and nodal labels, and failure to reach surgery included in the estimand. AI, RECIST, and PERCIST should be compared in the same patients with dependence

preserved, then assessed for calibration and net benefit at explicit surgical thresholds [6]. Such a design would test whether radiomics improves decisions, while the present synthesis more appropriately supports promising pathological-response prediction among patients who reach resection.

Acknowledgments

During preparation of this letter, the authors used ChatGPT (OpenAI; accessed August 20, 2026) to assist with language editing. The authors independently verified all article-specific facts and references, revised the output, and take full responsibility for the manuscript.

Funding

This correspondence received no specific funding.

Data Availability

No new data were generated or analyzed.

Conflicts of Interest

None declared.

References

1. Jiang Z, Xu Y, Jia S, Liu H. Radiomics-based AI for predicting neoadjuvant immunotherapy pathological response in non-small cell lung cancer: systematic review and meta-analysis. *J Med Internet Res*. Aug 13, 2026;28:e93892. [doi: [10.2196/93892](https://doi.org/10.2196/93892)] [Medline: [42596052](https://pubmed.ncbi.nlm.nih.gov/42596052/)]
2. Deng Y, Zhang X, Hu F, Lan X. Quantitative ^{18}F -FDG PET/CT model for predicting pathological complete response to neoadjuvant immunotherapy in NSCLC: comparison with RECIST 1.1 and PERCIST. *Eur J Nucl Med Mol Imaging*. Nov 2025;52(13):4806-4819. [doi: [10.1007/s00259-025-07342-8](https://doi.org/10.1007/s00259-025-07342-8)] [Medline: [40418330](https://pubmed.ncbi.nlm.nih.gov/40418330/)]
3. Nikoloulopoulos AK. Joint meta-analysis of two diagnostic tests accounting for within and between studies dependence. *Stat Methods Med Res*. Oct 2024;33(10):1800-1817. [doi: [10.1177/09622802241269645](https://doi.org/10.1177/09622802241269645)] [Medline: [39262159](https://pubmed.ncbi.nlm.nih.gov/39262159/)]
4. Forde PM, Spicer J, Lu S, et al. Neoadjuvant nivolumab plus chemotherapy in resectable lung cancer. *N Engl J Med*. May 26, 2022;386(21):1973-1985. [doi: [10.1056/NEJMoa2202170](https://doi.org/10.1056/NEJMoa2202170)] [Medline: [35403841](https://pubmed.ncbi.nlm.nih.gov/35403841/)]
5. Ruchalski K, Braschi-Amirfarzan M, Douek M, et al. A primer on RECIST 1.1 for oncologic imaging in clinical drug trials. *Radiol Imaging Cancer*. May 2021;3(3):e210008. [doi: [10.1148/rycan.2021210008](https://doi.org/10.1148/rycan.2021210008)] [Medline: [33988475](https://pubmed.ncbi.nlm.nih.gov/33988475/)]
6. Vasey B, Nagendran M, Campbell B, et al. Reporting guideline for the early-stage clinical evaluation of decision support systems driven by artificial intelligence: DECIDE-AI. *Nat Med*. May 2022;28(5):924-933. [doi: [10.1038/s41591-022-01772-9](https://doi.org/10.1038/s41591-022-01772-9)] [Medline: [35585198](https://pubmed.ncbi.nlm.nih.gov/35585198/)]

Abbreviations

pCR: pathological complete response

PERCIST: Positron Emission Tomography Response Criteria in Solid Tumors

RECIST: Response Evaluation Criteria in Solid Tumors

Edited by Harshita Pawar; This is a non-peer-reviewed article; submitted 20.Aug.2026; accepted 31.Aug.2026; published 06.Oct.2026

Please cite as:

Kan W, Pei H, Zhang M

Same-Patient, Same-Time Point Evidence for Radiomics Benchmarking

J Med Internet Res 2026;28:e110058

URL: <https://www.jmir.org/2026/1/e110058>

doi: [10.2196/110058](https://doi.org/10.2196/110058)

© Wenze Kan, Haoliang Pei, Mingxin Zhang. Originally published in the Journal of Medical Internet Research (<https://www.jmir.org>), 06.Oct.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.