Am J Clin Exp Immunol 2026;15(4):333-339
www.ajcei.us /ISSN:2164-7712/AJCEI0176023

Commentary

Multimodal AI in precision medicine: linking omics, imaging and clinical decisions

Rong Wei[1,2], Qiping Zheng[1,2]

[1]*The Molecular Oncology Laboratory, Department of Orthopedic Surgery and Rehabilitation Medicine, The University of Chicago Medical Center, Chicago, IL 60637, USA;* [2]*Department of Hematology and Hematological Laboratory Science, School of Clinical Medicine & Laboratory Medicine, Jiangsu University, Jiangsu 212013, China*

Received August 15, 2026; Accepted August 25, 2026; Epub August 25, 2026; Published August 30, 2026

Abstract: Precision medicine is shifting from a single-biomarker paradigm toward multimodal integration of molecular, imaging, and clinical data, with artificial intelligence (AI) serving as a key enabling technology. Multimodal imaging, computational pathology, spatial omics and liquid biopsy have all become more adept at capturing disease and biological variation, and foundation models, especially vision-language models, are now creating a common language between histology, molecular and biomedical text. AI-powered pipelines are also improving data quality upstream, which ultimately means that multimodal inference can be applied to tasks of real clinical importance: biomarker screening, modelling treatment outcomes, even autonomous coordination among diagnostic systems, rather than just retrospective prediction. But none of this means more modalities necessarily translate into more value. Additional data may introduce redundancy, technical artefacts, institutional bias, and missingness. The future of multimodal AI should therefore emphasize sufficiency rather than maximalism, with systems designed to accommodate incomplete or asynchronous inputs, quantify uncertainty, and demonstrate added value over established clinical standards. Ultimately, clinical utility will depend on external and prospective validation showing improvements in calibration, efficiency, net clinical benefit, and patient outcomes. The promise of AI-enabled multimodal precision medicine lies not in combining every available data stream, but in determining which molecular, morphological, and clinical information is necessary, complementary, and actionable for an individual patient at a specific decision point.

Keywords: Artificial intelligence, multimodal learning, precision medicine, multi-omics, medical imaging

Quantitative computational models of tissue and molecular states are increasingly complementing conventional single diagnostic markers in precision medicine [1, 2]. BioCompNet, for example, extracts automated body-composition features from magnetic resonance imaging for cardiometabolic stratification [3]. This shift toward reusable representations is also evident in disease-specific and cross-task imaging models: RETFound supports label-efficient transfer across retinal disease tasks, UMedPT learns transferable features across heterogeneous biomedical imaging tasks, and MedFMC provides a real-world benchmark for adapting foundation models to medical image classification [4-6]. In computational pathology, systematic evaluation of foundation-model feature extractors using whole-slide multiple-instance learning has demonstrated that pretrained representations can transfer across grading and biomarker-prediction tasks [7], while CHIEF extends this approach toward pan-cancer detection and prognosis-oriented representation learning [8]. Doppler ultrasonography combined with magnetic resonance imaging has also been used to predict blood flow under simulated microgravity [9], whereas RadFM extends generalist multimodal modeling across web-scale two- and three-dimensional radiological data [10]. Similarly, Prov-GigaPath scales self-supervised learning to gigapixel whole-slide pathology while incorporating both local and global tissue context [11], and FedM2CT addresses cross-institutional variability through federated computed-tomography reconstruction without pooling patient data [12]. Collectively, these advances are moving medical imaging beyond passive visual documentation toward scalable computational phenotyping.

https://doi.org/10.62347/URCI9453

Molecular profiling augments this imaging dimension by providing mechanistic insight into disease and its diversity. With the support of machine learning, single-cell and spatial technologies are increasingly able to map molecular drivers within local tissue environments; in lung adenocarcinoma, for example, such approaches have revealed MAPK-driven transcriptional programs within specific immune niches [13]. The same convergence now extends to spatial modeling: STAIG integrates expression, spatial position and image information [14], while GHIST and iSCALE use histology or tissue-scale structure to infer or extend molecular states beyond the native assay resolution [15, 16]. Recent approaches have further bridged morphology and molecular state. OmiCLIP links histopathology with spatial transcriptomics to enable cell-type decomposition and gene-expression prediction [17], while SEQUOIA and mSTAR provide complementary routes from histology to transcriptomic or multimodal slide representations [18, 19]. A third, less invasive layer is added by liquid biopsy: plasma and tissue DNA methylation signatures can detect early-stage non-small-cell lung cancer [20], and combining liquid-biopsy signals with radiology has improved non-invasive discrimination of gallbladder cancer from benign disease [21]. Vision-language models extend this integration further. MCPL aligns medical text and image representations through collaborative multimodal prompting across downstream medical tasks [22], while VLM-CPL, generative and foundation-model approaches in digital pathology, and PathChat illustrate complementary strategies for annotation-efficient classification, language-guided representation learning, and interactive multimodal reasoning [23-25]. Similar principles apply beyond oncology; in breast capsular contracture, transcriptomic and machine-learning analyses have identified metabolic and immune mechanisms associated with disease progression [26]. More broadly, pan-cancer histology-genomic fusion, multi-omic prediction of breast-cancer therapy response and systematic multimodal fusion studies show how complementary modalities can sharpen patient stratification and biomarker discovery [27-29].

The integration of molecular and morphological information is being further accelerated by foundation models and artificial intelligence (AI)-enabled analytical pipelines. One example of AI's upstream contribution is mtFFPECleaner, which filters spurious mitochondrial DNA variants caused by formalin-fixation artefacts in sequencing data, highlighting that reliable multimodal inference depends on analytical accuracy before integration [30]. At the molecular-integration level, Flexynesis provides a reusable framework for bulk multi-omics fusion, feature prioritization, and clinical-outcome modeling [31]. MUSK advances shared representation learning by combining large-scale pathology images and biomedical text within a unified vision-language model for cancer detection, biomarker prediction, and outcome modeling [32]. A systematic assessment of multimodal generative AI in clear-cell renal cell carcinoma similarly showed that general-purpose models can capture diagnostically relevant pathological features, although their performance varies substantially [33]. Recent work on AI and digital tools in cancer pathology extends this transition from image classification toward language-aware, generative, and workflow-oriented systems [34]. Another important interface between molecular measurement and patient care is clinical laboratory medicine, where AI is being investigated for quality control, interpretation, risk stratification, and clinical decision-making [35]. HONeYBEE demonstrates how foundation-model-derived embeddings can organize clinical, radiological, pathological, and molecular information at scale, including when modality availability is heterogeneous [36]. Together, these developments indicate that multimodal AI is moving beyond simple feature concatenation toward shared representations that preserve and connect molecular measurements, morphology, imaging, and language.

The central question, however, is whether multimodal integration can improve clinical decision-making rather than simply increase retrospective discrimination on benchmark datasets. Explainable AI applied to more than 15,000 patients with multiple solid cancers has integrated clinical variables, image-derived body-composition features, and tumor mutational profiles to identify patient-specific predictors of treatment outcome [37]. Across cancer types, multimodal systems have integrated radiology, pathology, and genomics to predict response to PD-(L)1 blockade [38], predicted

prognosis in adult and pediatric brain tumors [39], inferred transcriptomic information from histology for treatment-response modeling [40], used self-supervised histopathology to predict molecular classes in endometrial cancer [41], and derived pathomics signatures associated with prognosis and chemotherapy benefit in stage III colon cancer [42]. Histology-derived AI biomarkers have also approached the performance of transcriptomic predictors of anti-angiogenic treatment response in renal cancer [43]. More recent models incorporate clinical-information prompting [44], prior knowledge-guided multimodal learning [45], explicit robustness to missing modalities [46], and deep fusion of patho-radiomic and clinical data [47]. EAGLE moves closer to clinical implementation by predicting EGFR mutation status from routine lung adenocarcinoma histology; external validation and a prospective silent trial suggest that AI-assisted screening may reduce the need for rapid molecular testing while maintaining high screening accuracy [48]. An even broader transition is emerging from predictive models toward systems capable of orchestrating multiple specialized tools. An autonomous precision-oncology agent combining a large language model with multimodal diagnostic and analytical tools has already been tested in clinically representative decision-making scenarios [49]. These studies redefine the goal itself: the clinically meaningful endpoint is not maximal predictive accuracy, but the ability to judge when further testing is required, synthesize complementary evidence and support an actionable decision for an individual patient.

Multimodality, however, does not necessarily imply clinical value. Additional data can just as easily add redundancy, technical artefacts, institutional signatures and population-specific biases, and complete molecular, imaging and clinical measurements are seldom all available in routine care. The future of multimodal systems should therefore focus on sufficiency rather than maximalism: models that can cope with missing and asynchronous modalities, preserve modality-specific provenance, quantify uncertainty and show incremental benefit over robust clinical baselines rather than merely increasing complexity. Recent frameworks that explicitly evaluate heterogeneous or missing modalities illustrate one route towards this goal [36, 46], but robustness must ultimately be demonstrated prospectively.

Validation is as important as design: external validation across institutions, devices and populations should precede claims of generalizability, and prospective evaluation should include calibration, net clinical benefit, workflow efficiency and patient outcomes rather than isolated gains in area under the curve. Reporting and evaluation frameworks such as DECIDE-AI [50], TRIPOD+AI [51], STARD-AI [52] and FUTURE-AI [53] provide increasingly concrete standards for this transition. Methodological work has also converged on practical design principles: cross-modal interactions should be modeled explicitly rather than reduced to simple feature concatenation [54-56], while multi-omics integration should preserve biologically meaningful structures across data types. Translation also depends on demonstrating that added modalities improve biomarker discovery, diagnosis or prognosis in clinically relevant cohorts rather than only on retrospective benchmarks. The actual potential of AI-powered precision medicine is therefore not the combination of every data stream that happens to exist. It lies in determining which combination of molecular, morphological and clinical evidence is necessary, complementary and actionable at a specific decision point for an individual patient.

Disclosure of conflict of interest

None.

Address correspondence to: Qiping Zheng, The Molecular Oncology Laboratory, Department of Orthopedic Surgery and Rehabilitation Medicine, The University of Chicago Medical Center, Chicago, IL 60637, USA. E-mail: qp_zheng@hotmail.com

References

[1] Acosta JN, Falcone GJ, Rajpurkar P and Topol EJ. Multimodal biomedical AI. Nat Med 2022; 28: 1773-1784.

[2] Moor M, Banerjee O, Abad ZSH, Krumholz HM, Leskovec J, Topol EJ and Rajpurkar P. Foundation models for generalist medical artificial intelligence. Nature 2023; 616: 259-265.

[3] Wei J, Chen H, Yao L, Hou X, Zhang R, Shi L, Sun J, Hu C, Wei X and Jia W. BioCompNet: a deep learning workflow enabling automated body composition analysis toward precision management of cardiometabolic disorders. Cyborg Bionic Syst 2025; 6: 0381.

[4] Zhou Y, Chia MA, Wagner SK, Ayhan MS, Williamson DJ, Struyven RR, Liu T, Xu M, Lozano MG, Woodward-Court P, Kihara Y; UK Biobank Eye & Vision Consortium, Altmann A, Lee AY, Topol EJ, Denniston AK, Alexander DC and Keane PA. A foundation model for generalizable disease detection from retinal images. Nature 2023; 622: 156-163.

[5] Schäfer R, Nicke T, Höfener H, Lange A, Merhof D, Feuerhake F, Schulz V, Lotz J and Kiessling F. Overcoming data scarcity in biomedical imaging with a foundational multi-task model. Nat Comput Sci 2024; 4: 495-509.

[6] Wang D, Wang X, Wang L, Li M, Da Q, Liu X, Gao X, Shen J, He J, Shen T, Duan Q, Zhao J, Li K, Qiao Y and Zhang S. A real-world dataset and benchmark for foundation model adaptation in medical image classification. Sci Data 2023; 10: 574.

[7] Xu H, Wang M, Shi D, Qin H, Zhang Y, Liu Z, Madabhushi A, Gao P, Cong F and Lu C. When multiple instance learning meets foundation models: advancing histological whole slide image analysis. Med Image Anal 2025; 101: 103456.

[8] Wang X, Zhao J, Marostica E, Yuan W, Jin J, Zhang J, Li R, Tang H, Wang K, Li Y, Wang F, Peng Y, Zhu J, Zhang J, Jackson CR, Zhang J, Dillon D, Lin NU, Sholl L, Denize T, Meredith D, Ligon KL, Signoretti S, Ogino S, Golden JA, Nasrallah MP, Han X, Yang S and Yu KH. A pathology foundation model for cancer diagnosis and prognosis prediction. Nature 2024; 634: 970-978.

[9] Cai L, Liu Y, Li K, Xing C, Xu Z, Zhao L, Lv K, Li Z, Wang H, Wang L, Luo D, Yuan L, Qu L, Li Y, Wang Z and Ren P. Multimodal imaging-based cerebral blood flow prediction model development in simulated microgravity. Cyborg Bionic Syst 2025; 6: 0448.

[10] Wu C, Zhang X, Zhang Y, Hui H, Wang Y and Xie W. Towards generalist foundation model for radiology by leveraging web-scale 2D&3D medical data. Nat Commun 2025; 16: 7866.

[11] Xu H, Usuyama N, Bagga J, Zhang S, Rao R, Naumann T, Wong C, Gero Z, González J, Gu Y, Xu Y, Wei M, Wang W, Ma S, Wei F, Yang J, Li C, Gao J, Rosemon J, Bower T, Lee S, Weerasinghe R, Wright BJ, Robicsek A, Piening B, Bifulco C, Wang S and Poon H. A whole-slide foundation model for digital pathology from real-world data. Nature 2024; 630: 181-188.

[12] Wang H, Zhang X, Guo H, Ren X, Wang S, Fan F, Ma J and Zeng D. Federated metadata-constrained iRadonMAP framework with mutual learning for all-in-one computed tomography imaging. Cyborg Bionic Syst 2025; 6: 0376.

[13] Lin M, Zheng R, Liang P, Huang J, Ke X, Zhang W and Shang P. Exploring biomarkers of MAPK pathway co-expression in lung adenocarcinoma and their functions based on machine learning algorithms and single-cell analysis. Genes Dis 2025; 12: 101222.

[14] Yang Y, Cui Y, Zeng X, Zhang Y, Loza M, Park SJ and Nakai K. STAIG: Spatial transcriptomics analysis via image-aided graph contrastive learning for domain exploration and alignment-free integration. Nat Commun 2025; 16: 1067.

[15] Fu X, Cao Y, Bian B, Wang C, Graham D, Pathmanathan N, Patrick E, Kim J and Yang JYH. Spatial gene expression at single-cell resolution from histology using deep learning with GHIST. Nat Methods 2025; 22: 1900-1910.

[16] Schroeder A, Loth ML, Luo C, Yao S, Yan H, Zhang D, Piya S, Plowey E, Hu W, Clemenceau JR, Jang I, Kim M, Barnfather I, Chan SJ, Reynolds TL, Carlile T, Cullen P, Sung JY, Tsai HH, Park JH, Hwang TH, Zhang B and Li M. Scaling up spatial transcriptomics for large-sized tissues: uncovering cellular-level tissue architecture beyond conventional platforms with iSCALE. Nat Methods 2025; 22: 1911-1922.

[17] Chen W, Zhang P, Tran TN, Xiao Y, Li S, Shah VV, Cheng H, Brannan KW, Youker K, Lai L, Fang L, Yang Y, Le NT, Abe JI, Chen SH, Ma Q, Chen K, Song Q, Cooke JP and Wang G. A visual-omics foundation model to bridge histopathology with spatial transcriptomics. Nat Methods 2025; 22: 1568-1582.

[18] Pizurica M, Zheng Y, Carrillo-Perez F, Noor H, Yao W, Wohlfart C, Vladimirova A, Marchal K and Gevaert O. Digital profiling of gene expression from histology images with linearized attention. Nat Commun 2024; 15: 9886.

[19] Xu Y, Wang Y, Zhou F, Ma J, Jin C, Yang S, Li J, Zhang Z, Zhao C, Zhou H, Li Z, Lin H, Wang X, Wang J, Han A, Chan RCK, Liang L, Zhang X and Chen H. A multimodal knowledge-enhanced whole-slide pathology foundation model. Nat Commun 2025; 16: 11406.

[20] Li L, Fu K, Cai X, Liu D, Zhu Y, Wang W, Tian P, Wang Y, Xue H, Snyder MP and Li W. Diagnosis of early-stage non-small cell lung cancer using DNA methylation in tissue and plasma. Genes Dis 2025; 12: 101548.

[21] Yang M, Zhao Y, Li C, Weng X, Li Z, Guo W, Jia W, Feng F, Hu J, Sun H, Wang B, Li H, Li M, Wang T, Zhang W, Jiang X, Zhang Z, Liu F, Hu H, Wu X, Gu J, Yang G, Li G, Zhang H, Zhang T, Zang H, Zhou Y, He M, Yang L, Wang H, Chen T, Zhang J, Chen W, Wu W, Li M, Gong W, Lin X, Liu F, Liu Y and Liu Y. Multimodal integration of liquid biopsy and radiology for the noninvasive diagnosis of gallbladder cancer and benign disorders. Cancer Cell 2025; 43: 398-412, e394.

[22] Wang P, Zhang H and Yuan Y. MCPL: multimodal collaborative prompt learning for medi-

cal vision-language model. IEEE Trans Med Imaging 2024; 43: 4224-4235.
[23] Zhong L, Huang Z, Liu Y, Liao W, Zhang S, Wang G and Zhang S. VLM-CPL: consensus pseudo-labels from vision-language models for annotation-free pathological image classification. IEEE Trans Med Imaging 2025; 44: 4023-4036.
[24] Waqas A, Bui MM, Glassy EF, El Naqa I, Borkowski P, Borkowski AA and Rasool G. Revolutionizing digital pathology with the power of generative artificial intelligence and foundation models. Lab Invest 2023; 103: 100255.
[25] Lu MY, Chen B, Williamson DFK, Chen RJ, Zhao M, Chow AK, Ikemura K, Kim A, Pouli D, Patel A, Soliman A, Chen C, Ding T, Wang JJ, Gerber G, Liang I, Le LP, Parwani AV, Weishaupt LL and Mahmood F. A multimodal generative AI copilot for human pathology. Nature 2024; 634: 466-473.
[26] Mao Y, Hou X, Fu S and Luan J. Transcriptomic and machine learning analyses identify hub genes of metabolism and host immune response that are associated with the progression of breast capsular contracture. Genes Dis 2024; 11: 101087.
[27] Chen RJ, Lu MY, Williamson DFK, Chen TY, Lipkova J, Noor Z, Shaban M, Shady M, Williams M, Joo B and Mahmood F. Pan-cancer integrative histology-genomic analysis via multimodal deep learning. Cancer Cell 2022; 40: 865-878, e866.
[28] Sammut SJ, Crispin-Ortuzar M, Chin SF, Provenzano E, Bardwell HA, Ma W, Cope W, Dariush A, Dawson SJ, Abraham JE, Dunn J, Hiller L, Thomas J, Cameron DA, Bartlett JMS, Hayward L, Pharoah PD, Markowetz F, Rueda OM, Earl HM and Caldas C. Multi-omic machine learning predictor of breast cancer therapy response. Nature 2022; 601: 623-629.
[29] Steyaert S, Pizurica M, Nagaraj D, Khandelwal P, Hernandez-Boussard T, Gentles AJ and Gevaert O. Multimodal data fusion for cancer biomarker discovery with deep learning. Nat Mach Intell 2023; 5: 351-362.
[30] Li S, Guo W, Sun T, Xie F, Peng F, Zhou K, Lei Z, Guo X, Liu Y and Xing J. mtFFPECleaner: a high-fidelity machine learning approach for artifact removal in high-throughput mitochondrial DNA sequencing data of archival formalin-fixed paraffin-embedded specimens. Interdiscip Med 2025; 3: e70062.
[31] Uyar B, Savchyn T, Naghsh Nilchi A, Sarigun A, Wurmus R, Shaik MM, Grüning B, Franke V and Akalin A. Flexynesis: a deep learning toolkit for bulk multi-omics data integration for precision oncology and beyond. Nat Commun 2025; 16: 8261.
[32] Xiang J, Wang X, Zhang X, Xi Y, Eweje F, Chen Y, Li Y, Bergstrom C, Gopaulchan M, Kim T, Yu KH, Willens S, Olguin FM, Nirschl JJ, Neal J, Diehn M, Yang S and Li R. A vision-language foundation model for precision oncology. Nature 2025; 638: 769-778.
[33] Lu R, Shen J, Jiang A, Chen W, Qi C, Chen L, Zhu L, Mou W, Gan W, Zeng D, Tang B, Xiao M, Chu G, Peng S, Wong HZH, Zhang L, Zhang H, Deng X, Cheng Q, Cui X, Lin A and Luo P. Towards artificial intelligence-assisted digital pathology: a systematic evaluation of multimodal generative artificial intelligence in clear cell renal cell carcinoma assessment. Interdiscip Med 2025; 3: e70047.
[34] Shaktah LA, Carrero ZI, Hewitt KJ, Gustav M, Cecchini M, Foersch S, Berezowska S and Kather JN. Application of artificial intelligence and digital tools in cancer pathology. Lancet Digit Health 2025; 7: 100933.
[35] Xie H, Jia Y and Liu S. Integration of artificial intelligence in clinical laboratory medicine: advancements and challenges. Interdiscip Med 2024; 2: e20230056.
[36] Tripathi A, Waqas A, Schabath MB, Yilmaz Y and Rasool G. HONeYBEE: enabling scalable multimodal AI in oncology through foundation model-driven embeddings. NPJ Digit Med 2025; 8: 622.
[37] Keyl J, Keyl P, Montavon G, Hosch R, Brehmer A, Mochmann L, Jurmeister P, Dernbach G, Kim M, Koitka S, Bauer S, Bechrakis N, Forsting M, Führer-Sakel D, Glas M, Grünwald V, Hadaschik B, Haubold J, Herrmann K, Kasper S, Kimmig R, Lang S, Rassaf T, Roesch A, Schadendorf D, Siveke JT, Stuschke M, Sure U, Totzeck M, Welt A, Wiesweg M, Baba HA, Nensa F, Egger J, Müller KR, Schuler M, Klauschen F and Kleesiek J. Decoding pan-cancer treatment outcomes using multimodal real-world data and explainable artificial intelligence. Nat Cancer 2025; 6: 307-322.
[38] Vanguri RS, Luo J, Aukerman AT, Egger JV, Fong CJ, Horvat N, Pagano A, Araujo-Filho JAB, Geneslaw L, Rizvi H, Sosa R, Boehm KM, Yang SR, Bodd FM, Ventura K, Hollmann TJ, Ginsberg MS, Gao J; MSK MIND Consortium, Hellmann MD, Sauter JL and Shah SP. Multimodal integration of radiology, pathology and genomics for prediction of response to PD-(L)1 blockade in patients with non-small cell lung cancer. Nat Cancer 2022; 3: 1151-1164.
[39] Steyaert S, Qiu YL, Zheng Y, Mukherjee P, Vogel H and Gevaert O. Multimodal deep learning to predict prognosis in adult and pediatric brain tumors. Commun Med (Lond) 2023; 3: 44.
[40] Hoang DT, Dinstag G, Shulman ED, Hermida LC, Ben-Zvi DS, Elis E, Caley K, Sammut SJ,

Sinha S, Sinha N, Dampier CH, Stossel C, Patil T, Rajan A, Lassoued W, Strauss J, Bailey S, Allen C, Redman J, Beker T, Jiang P, Golan T, Wilkinson S, Sowalsky AG, Pine SR, Caldas C, Gulley JL, Aldape K, Aharonov R, Stone EA and Ruppin E. A deep-learning framework to predict cancer treatment response from histopathology images through imputed transcriptomics. Nat Cancer 2024; 5: 1305-1317.

[41] Fremond S, Andani S, Barkey Wolf J, Dijkstra J, Melsbach S, Jobsen JJ, Brinkhuis M, Roothaan S, Jurgenliemk-Schulz I, Lutgens LCHW, Nout RA, van der Steen-Banasik EM, de Boer SM, Powell ME, Singh N, Mileshkin LR, Mackay HJ, Leary A, Nijman HW, Smit VTHBM, Creutzberg CL, Horeweg N, Koelzer VH and Bosse T. Interpretable deep learning model to predict the molecular classification of endometrial cancer from haematoxylin and eosin-stained whole-slide images: a combined analysis of the PORTEC randomised trials and clinical cohorts. Lancet Digit Health 2023; 5: e71-e82.

[42] Jiang W, Wang H, Dong X, Yu X, Zhao Y, Chen D, Yan B, Cheng J, Zhuo S, Wang H and Yan J. Pathomics signature for prognosis and chemotherapy benefits in stage III colon cancer. JAMA Surg 2024; 159: 519-528.

[43] Jasti J, Zhong H, Panwar V, Jarmale V, Miyata J, Carrillo D, Christie A, Rakheja D, Modrusan Z, Kadel EE 3rd, Beig N, Huseni M, Brugarolas J, Kapur P and Rajaram S. Histopathology based AI model predicts anti-angiogenic therapy response in renal cancer clinical trial. Nat Commun 2025; 16: 2610.

[44] Hou J, Zhang R, Xie Y, Li C and Qin W. Multimodal deep learning for cancer prognosis prediction with clinical information prompts integration. NPJ Digit Med 2025; 9: 76.

[45] He Q, Tan H, Xiao B, Tan Y, Peng X, Peng C, Yue X, Jiang L, Cao Y, Lv FJ, Zhao W, Yi H, Liu Y, He W and Xiao M. Prior knowledge-guided multimodal deep learning system for biomarker exploration and prognosis prediction of urothelial carcinoma. NPJ Digit Med 2025; 9: 53.

[46] Xiao Y, Yang S, He M, Chen L, Wu Y and Zhong L. UroFusion-X: a unified multimodal deep learning framework for robust diagnosis, subtyping, and prognosis of urological cancers. NPJ Digit Med 2026; 9: 117.

[47] Shi R, Sun J, Zhou Z, Su Q and Shu Y. Deep multimodal fusion of patho-radiomic and clinical data for enhanced survival prediction for colorectal cancer patients. NPJ Digit Med 2025; 9: 86.

[48] Campanella G, Kumar N, Nanda S, Singi S, Fluder E, Kwan R, Muehlstedt S, Pfarr N, Schüffler PJ, Häggström I, Neittaanmäki N, Akyürek LM, Basnet A, Jamaspishvili T, Nasr MR, Croken MM, Hirsch FR, Elkrief A, Yu H, Ardon O, Goldgof GM, Hameed M, Houldsworth J, Arcila M, Fuchs TJ and Vanderbilt C. Real-world deployment of a fine-tuned pathology foundation model for lung cancer biomarker detection. Nat Med 2025; 31: 3002-3010.

[49] Ferber D, El Nahhas OSM, Wölflein G, Wiest IC, Clusmann J, Leßmann ME, Foersch S, Lammert J, Tschochohei M, Jäger D, Salto-Tellez M, Schultz N, Truhn D and Kather JN. Development and validation of an autonomous artificial intelligence agent for clinical decision-making in oncology. Nat Cancer 2025; 6: 1337-1349.

[50] Vasey B, Nagendran M, Campbell B, Clifton DA, Collins GS, Denaxas S, Denniston AK, Faes L, Geerts B, Ibrahim M, Liu X, Mateen BA, Mathur P, McCradden MD, Morgan L, Ordish J, Rogers C, Saria S, Ting DSW, Watkinson P, Weber W, Wheatstone P and McCulloch P; DECIDE-AI expert group. Reporting guideline for the early-stage clinical evaluation of decision support systems driven by artificial intelligence: DECIDE-AI. Nat Med 2022; 28: 924-933.

[51] Collins GS, Moons KGM, Dhiman P, Riley RD, Beam AL, Van Calster B, Ghassemi M, Liu X, Reitsma JB, van Smeden M, Boulesteix AL, Camaradou JC, Celi LA, Denaxas S, Denniston AK, Glocker B, Golub RM, Harvey H, Heinze G, Hoffman MM, Kengne AP, Lam E, Lee N, Loder EW, Maier-Hein L, Mateen BA, McCradden MD, Oakden-Rayner L, Ordish J, Parnell R, Rose S, Singh K, Wynants L and Logullo P. TRIPOD+AI statement: updated guidance for reporting clinical prediction models that use regression or machine learning methods. BMJ 2024; 385: e078378.

[52] Sounderajah V, Guni A, Liu X, Collins GS, Karthikesalingam A, Markar SR, Golub RM, Denniston AK, Shetty S, Moher D, Bossuyt PM, Darzi A and Ashrafian H; STARD-AI Steering Committee. The STARD-AI reporting guideline for diagnostic accuracy studies using artificial intelligence. Nat Med 2025; 31: 3283-3289.

[53] Lekadir K, Frangi AF, Porras AR, Glocker B, Cintas C, Langlotz CP, Weicken E, Asselbergs FW, Prior F, Collins GS, Kaissis G, Tsakou G, Buvat I, Kalpathy-Cramer J, Mongan J, Schnabel JA, Kushibar K, Riklund K, Marias K, Amugongo LM, Fromont LA, Maier-Hein L, Cerdá-Alberich L, Martí-Bonmatí L, Cardoso MJ, Bobowicz M, Shabani M, Tsiknakis M, Zuluaga MA, Fritzsche MC, Camacho M, Linguraru MG, Wenzel M, De Bruijne M, Tolsgaard MG, Goisauf M, Cano Abadía M, Papanikolaou N, Lazrak N, Pujol O, Osuala R, Napel S, Colantonio S, Joshi S, Klein S, Aussó S, Rogers WA, Salahuddin Z and Starmans MPA; FUTURE-AI Consortium. FUTURE-AI: international consensus guideline for trust-

worthy and deployable artificial intelligence in healthcare. BMJ 2025; 388: e081554.
[54] Lipkova J, Chen RJ, Chen B, Lu MY, Barbieri M, Shao D, Vaidya AJ, Chen C, Zhuang L, Williamson DFK, Shaban M, Chen TY and Mahmood F. Artificial intelligence for multimodal data integration in oncology. Cancer Cell 2022; 40: 1095-1110.
[55] Chen RJ, Lu MY, Wang J, Williamson DFK, Rodig SJ, Lindeman NI and Mahmood F. Pathomic fusion: an integrated framework for fusing histopathology and genomic features for cancer diagnosis and prognosis. IEEE Trans Med Imaging 2022; 41: 757-770.
[56] Qiu L, Zhao L, Hou R, Zhao W, Zhang S, Lin Z, Teng H and Zhao J. Hierarchical multimodal fusion framework based on noisy label learning and attention mechanism for cancer classification with pathology and genomic features. Comput Med Imaging Graph 2023; 104: 102176.