Cancer Biol Med 2026. doi: 10.20892/j.issn.2095-3941.2025.0636

PERSPECTIVE

Beyond origin: multimodal AI synthesis to resolve cancers of unknown primary

Hongru Shen, Xiangchun Li

Key Laboratory of Cancer Prevention and Therapy, Tianjin's Clinical Research Center for Cancer, Tianjin Medical University Cancer Institute & Hospital, National Clinical Research Center for Cancer, Tianjin Medical University, Tianjin 300060, China

For decades, the central dogma of oncology has been that a cancer's identity is inextricably linked to its anatomical origin. This principle underpins the entire diagnostic and therapeutic framework, from histology-based classification to site-specific treatment guidelines. Yet, this framework catastrophically fails for a substantial population of patients diagnosed with cancer of unknown primary (CUP). These patients present metastatic disease, yet their primary tumors remain elusive despite exhaustive clinical workup[1]. CUP, accounting for 1%–3% of all cancer diagnoses, is an enigma with devastating consequences; the median overall survival is only 2–12 months[2-4]. The inability to pinpoint an origin forces clinicians to rely on broad-spectrum empirical chemotherapy, such as taxane-carboplatin regimens, which have limited efficacy and exclude patients from the promise of targeted therapies and clinical trials[5]. CUP is not only a diagnostic challenge but also an indictment of the siloed approach to understanding malignancy: this cancer highlights the limitations of origin-based diagnostic frameworks. However, the confluence of high-dimensional biological data and advanced artificial intelligence (AI) is now poised to address this long-standing diagnostic limitation and to herald a new era for not only CUP but also oncology as a whole (**Figure 1**).

A revolution in silos: the rise of unimodal predictors

The past decade has witnessed a remarkable, albeit fragmented, revolution in the ability to infer a tumor's tissue of origin (TOO). AI and machine learning have been the driving force enabling the extraction of subtle, predictive patterns from complex biological data that were previously invisible to human experts.

In genomics, comprehensive profiling has moved beyond identifying single driver mutations to providing rich molecular fingerprints of individual tumors. Gene expression profiling and next-generation sequencing have become foundational. AI models such as CUP-AI-Dx have demonstrated high accuracy in predicting TOO from RNA expression data[6]. More recently, a new generation of deep learning models such as OncoNPC and GDD-ENS has enabled patterns in somatic mutations from clinically standard gene panels to be leveraged to classify tumor types with impressive performance, thus offering a more accessible route to genomic diagnosis[7,8]. The emergence of large language models such as OncoChat, trained on vast genomic datasets, has further advanced the frontier of classification and diagnostic support[9]. These tools are beginning to deliver on the promise of molecularly guided therapy. The landmark CUPISCO and Fudan CUP-001 clinical trials have recently demonstrated that using molecular information to guide treatment, through either histology-agnostic targeted therapy or site-specific therapy, results in longer progression-free survival than standard empirical chemotherapy, thus providing definitive evidence of the clinical utility of this approach[4,10,11].

Simultaneously, a parallel revolution has unfolded in pathology. Deep learning models have shown that cellular morphology and tissue architecture within standard images of hematoxylin and eosin staining contain extensive information regarding tumor lineage. Algorithms such as TOAD and TORCH have achieved high accuracy in TOO prediction from pathology images alone, thus demonstrating the untapped diagnostic potential of routine histology[12,13]. This capability is being markedly scaled by new pathology foundation models. Systems such as CONCH and PathChat, which integrate visual and language data, can interpret and even discuss complex histopathological features in natural language, thus providing a powerful new interface that augments pathologists' diagnostic capabilities[14,15].

Radiomics applies AI to extract thousands of quantitative features from medical images such as computed tomography

Correspondence to: Xiangchun Li
E-mail: lixiangchun@tmu.edu.cn
ORCID ID: https://orcid.org/0000-0002-2230-2152
Received October 15, 2025; accepted December 22, 2025.
Available at www.cancerbiomed.org
©2026 The Authors. Creative Commons Attribution-NonCommercial 4.0 International License (CC BY-NC 4.0)

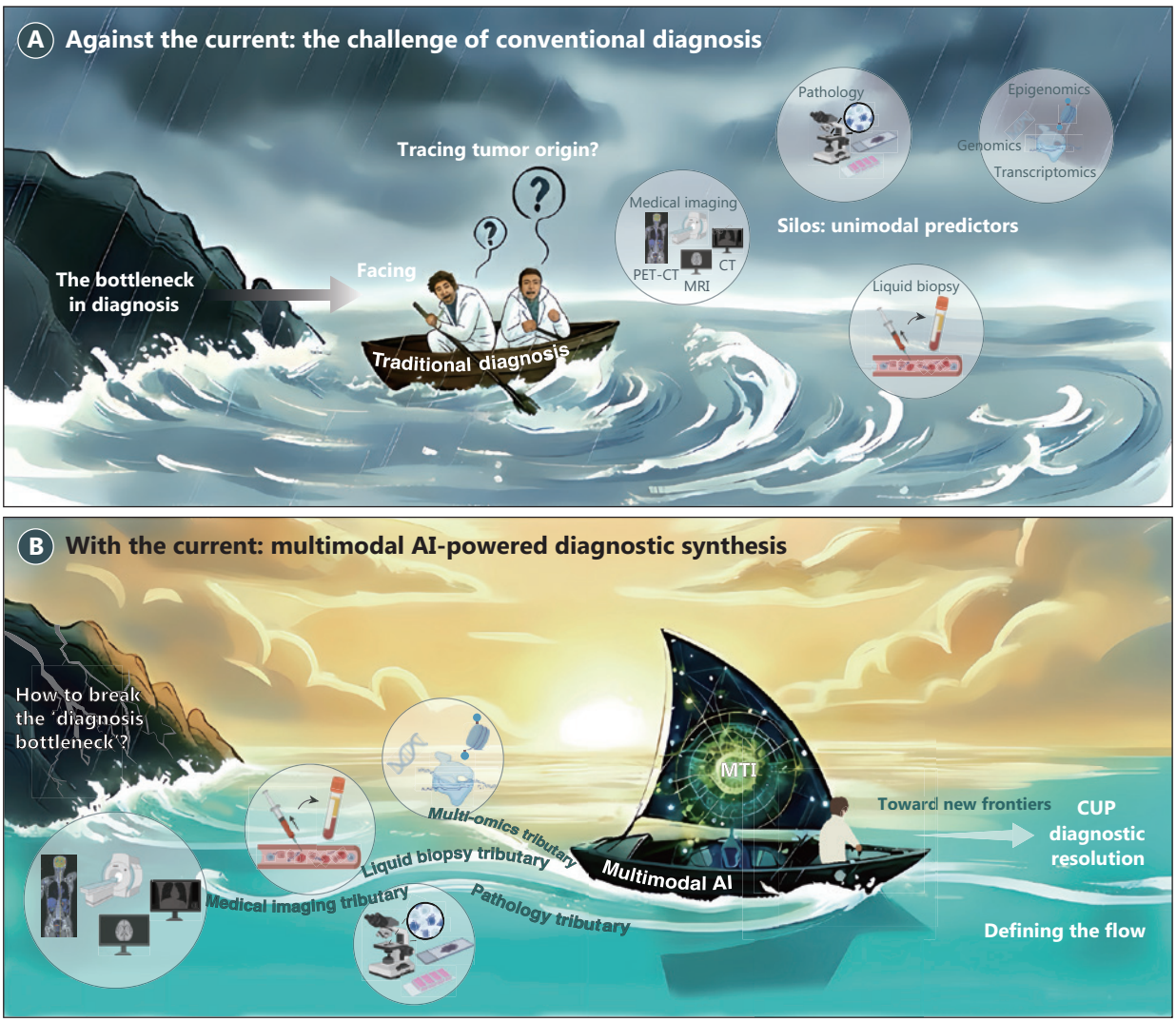


Figure 1 A paradigm shift in the diagnosis of CUP: from retrograde inference to prospective multimodal synthesis. (A) The conventional paradigm: a retrograde investigative process. The standard diagnostic workflow for CUP is fundamentally a retrograde endeavor akin to navigating against a current. This approach attempts to infer an often-obscured anatomical tissue of origin from the molecular and pathological features of a metastatic lesion. Key data modalities, including histopathology, genomics, medical imaging, and liquid biopsy analyses, are frequently assessed in isolation or in a sequential manner. This fragmented analysis of data silos creates a major diagnostic bottleneck characterized by high uncertainty and inefficiency. The clinical consequence of this retrograde strategy is often diagnostic ambiguity, thus leading to prolonged workups, empirical treatment, and delays in the initiation of targeted, site-specific therapy. (B) A multimodal AI-driven paradigm: prospective synthesis of a multimodal tumor identity. We propose a paradigm shift toward a prospective synthesis that leverages a multimodal AI model to integrate comprehensive patient data. Instead of tracing an anatomical origin, this approach treats the metastatic tumor as the primary source of information. Disparate, high-dimensional data streams are co-analyzed as integrated inputs for a foundation model. This model learns the deep, cross-modal relationships latent within the data to construct a MTI, a consolidated biological representation that defines the tumor's intrinsic lineage, molecular drivers, and therapeutic vulnerabilities. This prospective synthesis resolves the diagnostic impasse by defining what the cancer is rather than inferring where it came from. By providing an actionable tumor classification, the MTI can directly inform therapeutic decision-making and has promise in accelerating the path to effective precision medicine for patients with CUP. AI, artificial intelligence; CT, computed tomography; CUP, cancer of unknown primary; MRI, magnetic resonance imaging; MTI, multimodal tumor identity; PET, positron emission tomography.

(CT) and positron emission tomography (PET) scans, thereby revealing textural and spatial patterns that correlate with tumor biology and origin. In parallel, liquid biopsies offer a window into tumors through the bloodstream. Analysis of circulating tumor DNA (ctDNA) methylation patterns or fragmentation profiles can provide crucial clues regarding the TOO without

a need for tissue biopsy, thus conferring a critical advantage when tissue is scarce or inaccessible[16,17]. Multi-analyte tests such as CancerSEEK combine ctDNA analysis with protein biomarkers to enhance both detection and localization[18]. In each of these domains, AI is essential for separating the faint, tumor-derived signal from the overwhelming background noise of the body, as exemplified by language model-based tools such as ACID for cell-free DNA (cfDNA) analysis[19].

The need for synthesis: from data points to a biological portrait

Each of these unimodal advances represents a major step forward, yet each operates within the confines of its own data type and has intrinsic limitations. Traditional pathology, the cornerstone of diagnosis, falters with poorly differentiated tumors, in which morphological clues are lost. Genomic profiling, although powerful, requires sufficient high-quality tumor tissue, is costly, and provides no information regarding the tumor's spatial architecture or macroscopic presentation. Imaging captures this spatial context but lacks molecular depth, whereas liquid biopsies have drawbacks of low signal-to-noise ratios and a complete absence of anatomical information.

In essence, clinicians receive a series of disconnected, often contradictory "eyewitness" accounts of the tumor's identity. The challenge is not to choose the most reliable witness but to synthesize these accounts into a single, coherent narrative. A cancer's identity is not encoded solely in its DNA, its morphology, its metabolic activity, or its secreted biomarkers but instead is expressed through the complex interplay of all these features. A specific genomic alteration in a tumor may manifest as a subtle textural pattern on a PET-CT scan, a distinct cellular arrangement on a hematoxylin and eosin slide, or a characteristic methylation signature in the blood. To date, tools to learn and interpret this complex, cross-modal language have been lacking. Therefore, the necessary paradigm shift will lie in moving from unimodal classification to multimodal synthesis. The field must transition from developing isolated AI predictors to building integrated systems that can construct a comprehensive biological portrait of a patient's cancer from all available data (**Table 1**).

The dawn of the multimodal AI clinician

This vision is no longer speculative: agentic multimodal AI can integrate pathology, imaging, and molecular data to support evidence-based diagnostic prediction and decision-making. Architectures inspired by models such as Contrastive Language-Image Pre-training (CLIP) have demonstrated remarkable ability to learn shared representations between images and text, thus creating a common language for disparate data types[20]. This concept is now being extended into the medical domain. Models such as Med-PaLM Multimodal, and the forthcoming Med-Gemini are being trained on vast, diverse datasets of biomedical information, and are learning to connect clinical language, imaging, and genomics within a single, unified framework[21].

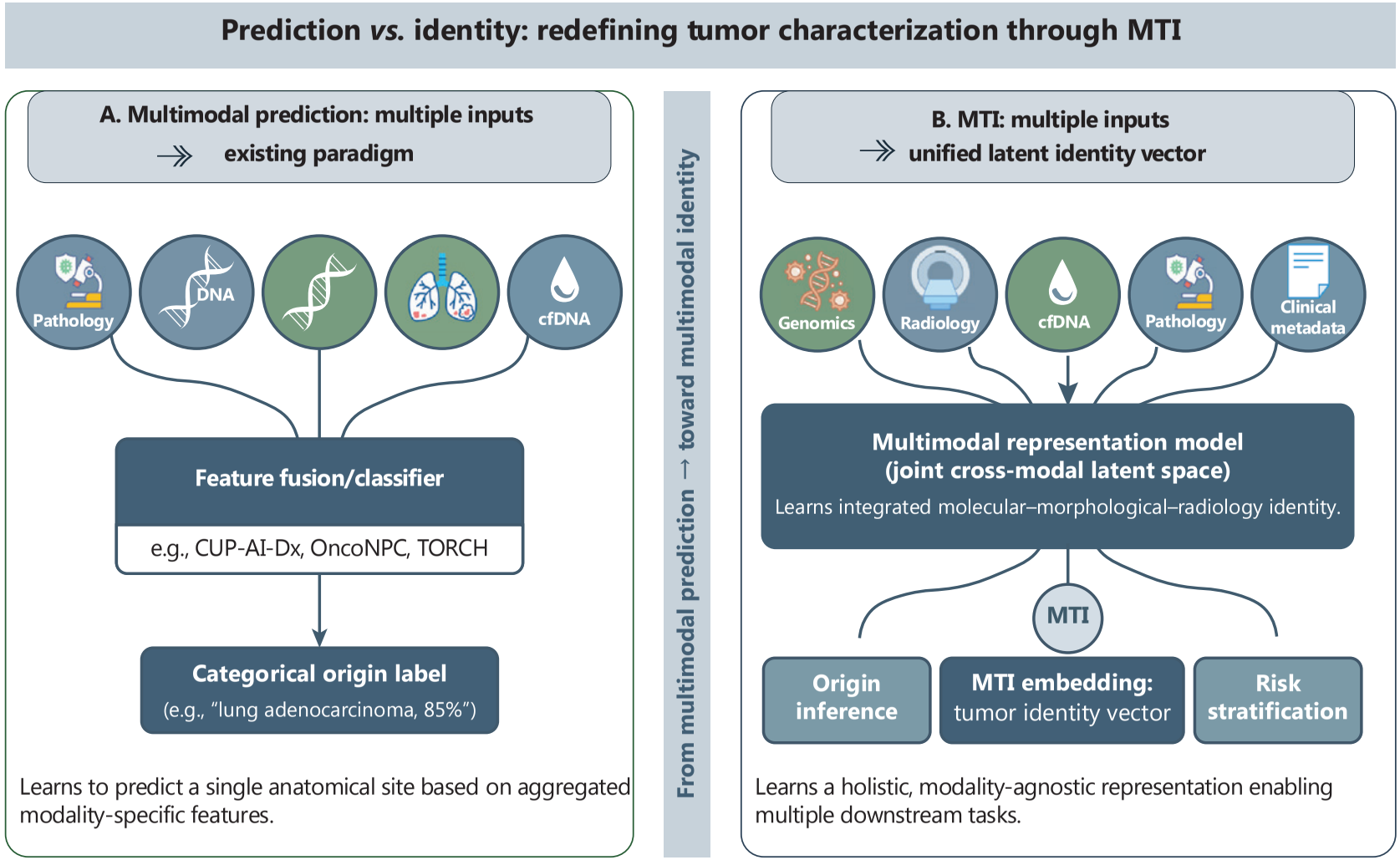
These models represent a fundamental leap beyond their unimodal predecessors. Architecturally, this advancement has been achieved by projecting disparate data types, such as genomic signatures and pathology image embeddings, into a shared latent space, thus allowing the model to learn a common, cross-modal language of cancer biology. Instead of merely classifying, these models synthesize data. A multimodal system for CUP would not just process a pathology image but interpret that image in the context of the patient's genomic alterations and radiological findings. For example, it could learn that a specific pattern of poorly differentiated adenocarcinoma on a liver biopsy, when combined with a KRAS G12D mutation and a characteristic pattern of fluorodeoxyglucose (FDG) uptake on a PET scan, confers a 95% probability of pancreatic primary cancer. The system would view genomic alterations not as isolated events but as features whose diagnostic significance is modulated by the complete clinical picture. This integrated analysis enables a far more robust and nuanced probabilistic scoring for the tumor's origin than any single modality could ever achieve. For example, our TORCH model[13] provides direct clinical evidence for CUP by combining cytological image representations with clinical and molecular features *via* a transformer-based multimodal attention architecture. In a cohort of 391 patients with CUP, those who received treatment concordant with their TORCH-predicted tissue of origin achieved significantly longer overall survival than discordant cases [median 27 *vs.* 17 months; hazard ratio (HR) ≈ 0.53]. These findings demonstrated that multimodal synthesis can inform real-world therapeutic decision-making and translate into measurable patient benefit.

The ultimate evolution of this paradigm is the AI agent, an autonomous system that not only integrates data but also actively performs reasoning regarding the data. Powered by large language models and equipped with specialized tools, these agents can navigate complex diagnostic pathways, plan workflows, and even generate novel hypotheses[22,23]. An oncologist could engage in a multi-turn dialogue with such an agent, by querying the integrated data to refine diagnostic hypotheses in real time. For example, one could ask, "Given the ctDNA methylation profile and the absence of a lesion on the chest CT, what is the likelihood of a gastro-esophageal primary *vs.* a biliary primary cancer?" The agent could then access specialized radiology and pathology foundation models such as RadFM or PathChat to provide an evidence-based answer and cite the specific features in each data stream that support its

Table 1 Evolution of diagnostic paradigms for CUP

Characteristic	Traditional era*	Current era: unimodal AI*	Future era: multimodal synthesis*
Guiding principle	A cancer's identity is its anatomical origin.	Cancer origin can be inferred from molecular/morphological signatures.	Cancer identity transcends anatomical origin; it is an emergent biological state.
Diagnostic goal	Identify primary tumors' anatomical sites through exhaustive investigation	Predict TOO with high probability from a single data modality	Define a comprehensive, actionable biological portrait to guide precision therapy
Primary data sources	Exhaustive clinical workup, IHC antibody panels, and anatomical imaging (CT/MRI)	– Deep genomics (NGS/GEP) or – Digital pathology (WSI) or – Radiomics (CT/PET) or – Liquid biopsy (ctDNA methylation/ fragmentation)	Integrated genomics, pathology, radiology, liquid biopsy, and clinical metadata from a single patient in a unified framework
Analytical methods	Human histopathological interpretation and clinical reasoning; pattern recognition based on experience	Siloed, unimodal deep learning models operating independently (e.g., OncoNPC or TORCH)	Multimodal foundation models and AI agents that learn cross-modal relationships (e.g., Med-Gemini, PathChat, or RadFM)
Diagnostic output	Negative diagnosis: what the tumor is not; broad differential diagnosis; diagnosis of exclusion	Categorical prediction: "85% probability of lung adenocarcinoma" with confidence intervals from a single modality	Quantitative biological identity: MTI score integrating probability, therapeutic vulnerabilities, and actionable molecular targets
Therapeutic strategies	Broad-spectrum empirical chemotherapy (e.g., taxane-carboplatin regimens); limited access to targeted therapies	Molecularly guided therapy according to predicted origin or actionable mutations (e.g., CUPISCO trial and Fudan CUP-001)	MTI-driven therapy selection based on an integrated biological profile, potentially superseding origin as the primary decision criterion
Clinical impact	Median OS: 2–12 months; diagnosis of exclusion; minimal trial eligibility; patient uncertainty and distress	Improved TOO prediction accuracy (70%–90%); gradual increase in targeted therapy access; progression-free survival benefits	Personalized, MTI-guided precision therapy; potential transformation of CUP from terminal diagnosis to manageable disease; enhanced trial matching
Key limitations	Low diagnostic accuracy; limited therapeutic options; prolonged diagnostic journey with high patient burden	Disconnected data streams yield contradictory results; each modality inherits intrinsic neglected areas: – Genomics: lacks spatial/architectural context – Pathology: fails with dedifferentiated tumors – Radiomics: lacks molecular depth – Liquid biopsy: low signal-to-noise ratio, no anatomical localization	Requires multi-institutional federated datasets with linked multi-modal data; demands model explainability (XAI) for clinical trust; risk of algorithmic bias across diverse populations; computational infrastructure needs

*__This table illustrates the evolution of the CUP diagnosis paradigm across 3 eras.__ The **traditional era** relied on anatomical origin as the defining principle, by using exhaustive but often futile clinical workups. The **current era** leverages AI to predict tumor origin from individual data modalities (genomics, pathology, imaging, or liquid biopsy), each of which has high accuracy but inherent limitations when used in isolation. The proposed **future era** represents a fundamental shift toward multimodal AI synthesis, wherein integrated biological data define a tumor's identity through a MTI score. This approach moves beyond simple tissue-of-origin classification by creating a comprehensive biological portrait that can directly guide precision therapy. Key enhancements include (1) shorter time to actionable diagnosis through real-time data integration, (2) superior clinical outcomes by addressing each modality's neglected areas through cross-modal learning, and (3) transformation of CUP from a diagnosis of exclusion to a model for integrated precision oncology. AI, artificial intelligence; ctDNA, circulating tumor DNA; CUP, cancer of unknown primary; CT, computed tomography; GEP, gene expression profiling; IHC, immunohistochemistry; MTI, multimodal tumor identity; NGS, next-generation sequencing; OS, overall survival; PET, positron emission tomography; TOO, tissue of origin; WSI, whole slide imaging; XAI, explainable artificial intelligence.



MTI represents a paradigm shift: from predicting where a tumor originated to characterizing what it fundamentally is.

This unified representation enables tasks that transcend traditional unimodal and multimodal prediction frameworks.

Figure 2 Conceptual distinction between multimodal prediction and MTI. (A) Existing multimodal prediction frameworks integrate features from one or several data types, such as pathology, genomics, radiology, or cfDNA, to generate a single categorical tissue-of-origin label. These systems operate within a supervised classification paradigm and are optimized for origin inference alone. (B) The proposed MTI framework synthesizes pathology, genomic, radiologic, liquid biopsy, and clinical inputs into a unified latent identity representation. Rather than producing only an origin label, this representation is designed to support multiple downstream clinical tasks, including origin inference, therapeutic prioritization, and risk stratification. MTI therefore represents a shift from classification toward integrated tumor identity modeling. AI, artificial intelligence; cfDNA, cell-free DNA; CUP, cancer of unknown primary; MTI, multimodal tumor identity.

conclusion[15,24,25]. This collaborative human-AI interaction moves beyond static predictions to a dynamic process of diagnostic discovery, by augmenting clinicians' expertise and surpassing the limits of unassisted human cognition.

A new grand challenge: MTI

Translating this transformative potential into clinical practice requires a concerted effort to overcome major hurdles, most notably data scarcity. Training robust multimodal models requires large, high-quality, and ethically sourced datasets that link genomic, pathologic, radiologic, and clinical outcomes data from the same patient. The rarity of CUP makes assembling such cohorts challenging and necessitates multi-institutional and international collaboration. Furthermore, the "black box" nature of complex deep learning models poses a barrier to clinical trust. Further development of Explainable AI (XAI) is necessary to ensure transparency. Finally, the risks of algorithmic bias must be proactively addressed to ensure that these tools are equitable across diverse patient populations.

However, these challenges should not be considered roadblocks but engineering problems to be solved on the path to a much larger vision. The goal should not be merely to build a better tool for predicting a tumor's origin but to question the primacy of origin itself. We propose that the field should focus on a new grand challenge: the creation and validation of a MTI score. Although multimodal prediction frameworks already exist, such as CUP-AI-Dx, OncoNPC, and more general biomedical models such as Med-Gemini, these systems are ultimately designed to produce a single tissue-of-origin label. In contrast, MTI is conceptualized as a unified latent representation that jointly encodes molecular, morphological, radiologic, cfDNA, and clinical characteristics. Rather than providing only an origin classification, MTI defines a modality-agnostic tumor identity space capable of supporting multiple downstream clinical tasks, including origin inference, therapeutic prioritization, and risk stratification (**Figure 2**). This AI-generated signature, derived from the synthesis of all available biological data, would represent a holistic, quantitative measure of a tumor's biological state.

For CUP, and eventually other cancers, this MTI could entirely supersede the traditional tissue-of-origin classification. Instead of forcing a tumor into an "anatomical box" that it might no longer resemble, the MTI would define it according to integrated molecular, morphological, and microenvironmental properties. This perspective does not erase the fundamental importance of developmental lineage but instead reframes it as one of many features in a more holistic and clinically actionable biological identity. This concept aligns with emerging clinical evidence from trials such as CUPISCO, wherein histology-agnostic, molecularly guided therapy was found to be superior to standard care, thereby suggesting that a tumor's actionable biology is more clinically relevant than its origin[10,26]. The MTI would serve as a fundamental new classification system to directly guide therapy according to a comprehensive biological profile rather than an inferred anatomical history. Although this long-term vision moves beyond origin-based frameworks, early clinical use of MTI will necessarily be transitional. MTI can function as a secondary classifier that refines origin assessment and highlights cross-modal discordance, and its integrated representations may support therapy prioritization by identifying actionable biological programs even when the primary site remains uncertain. These transitional roles position MTI as an augmentative layer within current CUP workflows that enables clinically compatible, incremental adoption. To operationalize this framework, we outline a concise implementation roadmap including data linkage, model architecture, validation strategy, and clinical deployment (**Figure 3**). Simultaneously, MTI faces practical limitations including the risk of cross-modal overfitting, biases introduced during data harmonization, and the challenge of ensuring that latent identity representations align with clinically actionable endpoints. Recognizing these constraints will be essential as MTI progresses toward broader clinical integration.

Developing MTI systems across institutions will require privacy-preserving data infrastructures. Federated learning offers a feasible approach by enabling joint model training on distributed datasets while keeping raw patient data local, thereby facilitating multi-center collaboration without compromising confidentiality. Ensuring robustness will also require systematic cross-population generalizability testing, evaluating performance across demographic groups, geographic regions,

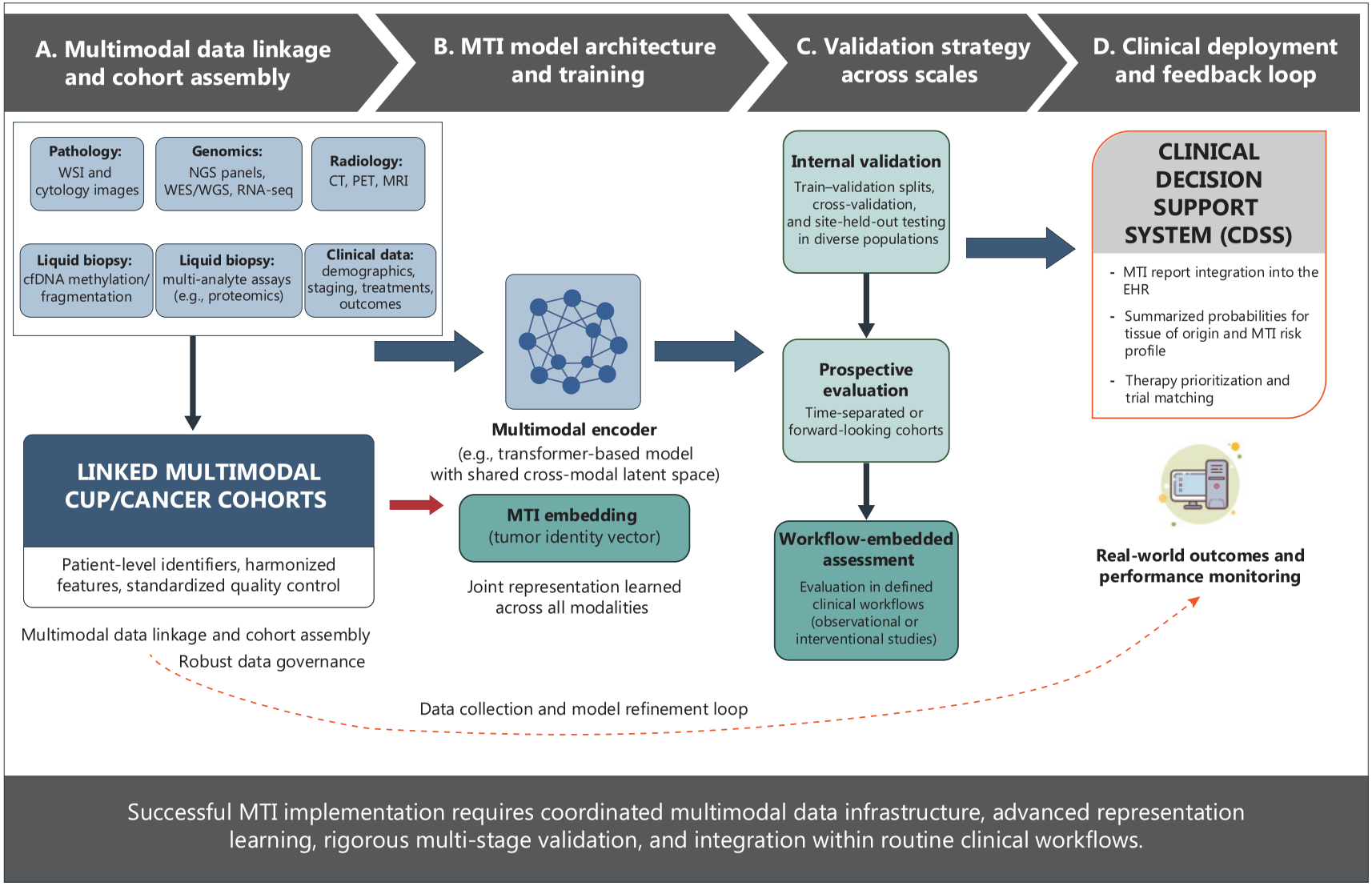

A. Multimodal data linkage and cohort assembly
B. MTI model architecture and training
C. Validation strategy across scales
D. Clinical deployment and feedback loop
Pathology: WSI and cytology images
Genomics: NGS panels, WES/WGS, RNA-seq
Radiology: CT, PET, MRI
Liquid biopsy: cfDNA methylation/ fragmentation
Liquid biopsy: multi-analyte assays (e.g., proteomics)
Clinical data: demographics, staging, treatments, outcomes
LINKED MULTIMODAL CUP/CANCER COHORTS
Patient-level identifiers, harmonized features, standardized quality control
Multimodal data linkage and cohort assembly
Robust data governance
Multimodal encoder
(e.g., transformer-based model with shared cross-modal latent space)
MTI embedding
(tumor identity vector)
Joint representation learned across all modalities
Internal validation
Train–validation splits, cross-validation, and site-held-out testing in diverse populations
Prospective evaluation
Time-separated or forward-looking cohorts
Workflow-embedded assessment
Evaluation in defined clinical workflows (observational or interventional studies)
CLINICAL DECISION SUPPORT SYSTEM (CDSS)
- MTI report integration into the EHR
- Summarized probabilities for tissue of origin and MTI risk profile
- Therapy prioritization and trial matching
Real-world outcomes and performance monitoring
Data collection and model refinement loop
Successful MTI implementation requires coordinated multimodal data infrastructure, advanced representation learning, rigorous multi-stage validation, and integration within routine clinical workflows.

Figure 3 End-to-end workflow for the development, validation, and clinical deployment of a MTI system. (A) Multimodal data linkage and cohort assembly. Pathology images, genomic sequencing data, radiologic imaging, liquid biopsy assays, and structured clinical variables are linked at the patient level under harmonized metadata standards and quality-control procedures, thus forming integrated CUP/cancer cohorts suitable for model training. (B) MTI model architecture and training. A multimodal encoder (e.g., transformer-based) learns a shared cross-modal latent space and generates an MTI embedding that captures the joint morphological, molecular, radiologic, and clinical characteristics of each tumor. (C) Validation strategy across scales. Model performance is assessed through internal validation (training-validation splits, cross-validation, and site-held-out testing), followed by prospective evaluation and workflow-embedded assessment within defined clinical settings. (D) Clinical deployment and feedback loop. MTI outputs, such as probabilities for tissue of origin, risk profiles, and therapy prioritization, are integrated into a clinical decision support system. Real-world outcomes and performance monitoring enable iterative model refinement and ensure safe, reliable clinical implementation. cfDNA, cell-free DNA; CT, computed tomography; CUP, cancer of unknown primary; EHR, electronic health record; MRI, magnetic resonance imaging; MTI, multimodal tumor identity; NGS, next-generation sequencing; PET, positron emission tomography; WSI, whole slide imaging.

and institutional workflows to identify and mitigate potential sources of algorithmic bias. Finally, clinical deployment must align with emerging regulatory pathways. Current FDA and EMA guidance for AI/ML-based medical devices emphasizes transparency, reproducibility, and post-deployment monitoring. These considerations will be essential for MTI systems intended to support oncology decision-making.

To achieve this vision, we call for a "CUP Moonshot": a global, collaborative effort to build the federated datasets and foundational models required to establish and validate the MTI as a new clinical standard.

CUP has long been emblematic of limitations in oncology and a diagnostic dead end. However, by forcing us to look beyond anatomical origin and embrace a more holistic, data-driven definition of cancer, it may paradoxically reveal a path forward. The multimodal AI synthesis required to solve the CUP enigma provides a blueprint for the future of cancer diagnostics, by moving from a series of disconnected tests to a single, unified understanding of a patient's disease. In resolving the identity of this one enigmatic cancer, tools will be developed to better understand all cancers. Ultimately, clinical encounters will be transformed by replacing diagnoses of exclusion and uncertainty with clear, actionable biological identities that offer patients not only targeted therapies but also a concrete understanding of their diseases.

Grant support

This work was supported by the National Natural Science Foundation of China (Grant Nos. 32270688, 31801117, and 82430107 to X.L., and 32500589 to H.S.), the China Postdoctoral Science Foundation (Grant Nos. BX20240253 and 2024M762384 to H.S.), the Natural Science Foundation of Tianjin (Grant No. 24JCQNJC01280 to H.S.), and Tianjin Key Medical Discipline (Specialty) Construction Project (Grant No. TJYXZDXK-3-003A).

Author contributions

Conceived and designed the analysis: Xiangchun Li, Hongru Shen.
Wrote the paper: Xiangchun Li, Hongru Shen.

Conflict of interest statement

No potential conflicts of interest are disclosed.

References

1. Varadhachary GR, Raber MN. Cancer of unknown primary site. N Engl J Med. 2014; 371: 757-65.
2. Kramer A, Bochtler T, Pauli C, Baciarello G, Delorme S, Hemminki K, et al. Cancer of unknown primary: ESMO clinical practice guideline for diagnosis, treatment and follow-up. Ann Oncol. 2023; 34: 228-46.
3. Pavlidis N, Pentheroudakis G. Cancer of unknown primary site. Lancet. 2012; 379: 1428-35.
4. Pouyiourou M, Bochtler T, Pauli C, Moch H, Brobeil A, Pantel K, et al. Rethinking cancer of unknown primary: from diagnostic challenge to targeted treatment. Nat Rev Clin Oncol. 2025; 22: 781-99.
5. Pavlidis N, Briasoulis E, Hainsworth J, Greco FA. Diagnostic and therapeutic management of cancer of an unknown primary. Eur J Cancer. 2003; 39: 1990-2005.
6. Zhao Y, Pan Z, Namburi S, Pattison A, Posner A, Balachander S, et al. CUP-AI-Dx: a tool for inferring cancer tissue of origin and molecular subtype using RNA gene-expression data and artificial intelligence. EBioMedicine. 2020; 61: 103030.
7. Moon I, LoPiccolo J, Baca SC, Sholl LM, Kehl KL, Hassett MJ, et al. Machine learning for genetics-based classification and treatment response prediction in cancer of unknown primary. Nat Med. 2023; 29: 2057-67.
8. Darmofal M, Suman S, Atwal G, Toomey M, Chen JF, Chang JC, et al. Deep-learning model for tumor-type prediction using

targeted clinical genomic sequencing data. Cancer Discov. 2024; 14: 1064-81.
9. Liu J, Yang M, Bi Y, Zhang J, Yang Y, Li Y, et al. Large language models enable tumor-type classification and localization of cancers of unknown primary from genomic data. Cell Rep Med. 2025; 6: 102332.
10. Kramer A, Bochtler T, Pauli C, Shiu KK, Cook N, de Menezes JJ, et al. Molecularly guided therapy versus chemotherapy after disease control in unfavourable cancer of unknown primary (CUPISCO): an open-label, randomised, phase 2 study. Lancet. 2024; 404: 527-39.
11. Liu X, Zhang X, Jiang S, Mo M, Wang Q, Wang Y, et al. Site-specific therapy guided by a 90-gene expression assay versus empirical chemotherapy in patients with cancer of unknown primary (Fudan CUP-001): a randomised controlled trial. Lancet Oncol. 2024; 25: 1092-102.
12. Lu MY, Chen TY, Williamson DFK, Zhao M, Shady M, Lipkova J, et al. AI-based pathology predicts origins for cancers of unknown primary. Nature. 2021; 594: 106-10.
13. Tian F, Liu D, Wei N, Fu Q, Sun L, Liu W, et al. Prediction of tumor origin in cancers of unknown primary origin with cytology-based deep learning. Nat Med. 2024; 30: 1309-19.
14. Lu MY, Chen B, Williamson DFK, Chen RJ, Liang I, Ding T, et al. A visual-language foundation model for computational pathology. Nat Med. 2024; 30: 863-74.
15. Lu MY, Chen B, Williamson DFK, Chen RJ, Zhao M, Chow AK, et al. A multimodal generative AI copilot for human pathology. Nature. 2024; 634: 466-73.
16. Conway AM, Pearce SP, Clipson A, Hill SM, Chemi F, Slane-Tan D, et al. A cfDNA methylation-based tissue-of-origin classifier for cancers of unknown primary. Nat Commun. 2024; 15: 3292.
17. Cristiano S, Leal A, Phallen J, Fiksel J, Adleff V, Bruhm DC, et al. Genome-wide cell-free DNA fragmentation in patients with cancer. Nature. 2019; 570: 385-9.
18. Cohen JD, Li L, Wang Y, Thoburn C, Afsari B, Danilova L, et al. Detection and localization of surgically resectable cancers with a multi-analyte blood test. Science. 2018; 359: 926-30.
19. Shen H, Liu J, Chen K, Li X. Language model enables end-to-end accurate detection of cancer from cell-free DNA. Brief Bioinform. 2024; 25: bbae053.
20. Radford A, Kim JW, Hallacy C, Ramesh A, Goh G, Agarwal S, et al. Learning transferable visual models from natural language supervision. In: Proceedings of the 38th International Conference on Machine Learning, PMLR 139, 18-24 July; 2021: 8748-63.
21. Tu T, Azizi S, Driess D, Schaekermann M, Amin M, Chang P-C, et al. Towards generalist biomedical AI. NEJM AI. 2024; 1.
22. Gao S, Fang A, Huang Y, Giunchiglia V, Noori A, Schwarz JR, et al. Empowering biomedical discovery with AI agents. Cell. 2024; 187: 6125-51.
23. Gottweis J, Weng W-H, Daryin A, Tu T, Palepu A, Sirkovic P, et al. Towards an AI co-scientist. arXiv preprint arXiv:2502.18864. 2025. https://arxiv.org/abs/2502.18864.
24. Wu C, Zhang X, Zhang Y, Hui H, Wang Y, Xie W. Towards generalist foundation model for radiology by leveraging web-scale 2D&3D medical data. Nat Commun. 2025; 16: 7866.
25. Ferber D, El Nahhas OSM, Wolflein G, Wiest IC, Clusmann J, Lessmann ME, et al. Development and validation of an autonomous artificial intelligence agent for clinical decision-making in oncology. Nat Cancer. 2025; 6: 1337-49.
26. Kato S, Alsafar A, Walavalkar V, Hainsworth J, Kurzrock R. Cancer of unknown primary in the molecular era. Trends Cancer. 2021; 7: 465-77.

Cite this article as: Shen H, Li X. Beyond origin: multimodal AI synthesis to resolve cancers of unknown primary. Cancer Biol Med. 2026; 23: 21-29. doi: 10.20892/j.issn.2095-3941.2025.0636