International Journal of Research in Engineering and Management Sciences (IJREMS)
E-ISSN: 3107-7439

Research Article
VOL.: 2, Issue: 2, February 2026

Deep Learning–Driven Integration of Radiological and Genomic Data for Precision Cancer Care

[1]Dr. Jayanthi Kanaka Ram, [2]Dr. Arvind R. Menon, [3]Dr. Neha S. Kulkarni, [4]Dr. Raghav K. Sharma, [5]Mr. Shatrughna U. Nagrik

*[*1]Consultant ENT, elova hospitals , Bengaluru, india, jayanthikr33@gmail.com*
[2]MD, DM (Medical Oncology) Professor, Medical Oncology Institute of Cancer Sciences Bengaluru, India arvind.menon@ics.edu.in
[3]MD, DM (Medical Oncology) Senior Consultant and Professor of Oncology Centre for Advanced Cancer Care Pune, India n.kulkarni@cacc.ac.in
[4]MD, DM (Medical Oncology) Director, Precision Oncology Program National Cancer Research Institute New Delhi, India rk.sharma@ncri.org.in
[5]satyjeet college of pharmacy, Mehkar (MS), India, rahulnagrik96@gmail.com

ABSTRACT:

Precision oncology aims to deliver personalized cancer diagnosis, prognosis, and treatment by integrating the molecular and phenotypic characteristics of individual tumors. Rapid advances in high-throughput genomic sequencing and medical imaging technologies have generated vast volumes of heterogeneous data, offering unprecedented opportunities to characterize tumor biology comprehensively. Radiological imaging provides non-invasive insights into tumor morphology, spatial heterogeneity, and therapeutic response, while genomic profiling reveals the molecular alterations underlying tumor initiation, progression, metastasis, and treatment resistance. However, conventional analytical methods are often limited in their ability to capture the complex, nonlinear relationships that exist across these complementary data modalities.

Multimodal deep learning has emerged as a transformative computational framework for integrating radiological and genomic information, enabling more comprehensive and accurate clinical decision-making in precision oncology. Leveraging advanced neural network architectures—including convolutional neural networks, transformers, graph neural networks, and multimodal fusion models—these approaches can identify latent associations between imaging phenotypes and genomic signatures that are not readily discernible through traditional analyses. Recent studies have demonstrated that multimodal integration consistently outperforms unimodal models across a range of applications, including cancer detection, molecular subtyping, prognostic prediction, treatment response assessment, biomarker discovery, and patient risk stratification.

Despite these promising advances, several challenges continue to hinder widespread clinical adoption. Limited availability of large, well-annotated multimodal datasets, data heterogeneity across institutions, model interpretability, privacy and security concerns, computational complexity, and regulatory considerations remain significant barriers to implementation. At the same time, the emergence of foundation models, self-supervised learning, and large-scale multimodal pretraining is reshaping the field by enabling robust representation learning from extensive unlabeled medical datasets and improving model generalizability across diverse clinical settings.

This review provides a comprehensive overview of recent advances in multimodal deep learning for integrating radiological imaging and genomic data in precision oncology. It discusses key methodological developments, radiogenomic integration strategies, multimodal fusion architectures, clinical applications, current limitations, and emerging trends, with particular emphasis on foundation models and next-generation artificial intelligence frameworks that have the potential to accelerate the translation of precision oncology into routine clinical practice.

Keywords: Precision oncology; Multimodal deep learning; Radiogenomics; Artificial intelligence; Medical imaging; Genomics; Multimodal data fusion; Foundation models.

Received Date: 7 February 2026**; Accepted Date:** 17 February 2026**; Published Date:** 24 February 2026.

1. Introduction

Cancer continues to be one of the foremost causes of morbidity and mortality worldwide despite significant progress in diagnostic technologies and therapeutic interventions. Increasing recognition of both interpatient and intratumoral heterogeneity has accelerated the shift from conventional one-size-fits-all treatment strategies toward precision oncology, an approach that tailors diagnosis and therapy according to the unique molecular and phenotypic characteristics of individual tumors [1,2]. This paradigm shift relies heavily on the integration of diverse clinical and biological data sources that collectively provide a comprehensive understanding of cancer complexity.

Medical imaging has become an essential pillar of contemporary oncology practice. Imaging modalities, including computed tomography (CT), magnetic resonance imaging (MRI), positron emission tomography (PET), and ultrasound, offer non-invasive visualization of tumor anatomy, physiology, metabolic activity, and therapeutic response. These imaging techniques play a central role in cancer detection, staging, treatment planning, and longitudinal monitoring of disease progression [3]. Concurrently, advances in next-generation sequencing technologies have revolutionized genomic profiling by enabling comprehensive characterization of tumors, including actionable genetic mutations, gene expression signatures, epigenetic alterations, and molecular pathways associated with tumor progression, metastasis, and therapeutic response [4].

Although radiological imaging and genomic profiling independently provide valuable clinical information, each modality captures distinct yet complementary dimensions of tumor biology. Imaging reflects the macroscopic phenotypic manifestations of underlying biological processes, whereas genomic analyses reveal the genetic and epigenetic mechanisms responsible for tumor initiation, evolution, and treatment resistance. Integrating these complementary sources of information enables a more comprehensive understanding of cancer biology than either modality can achieve independently [5].

The emerging discipline of radiogenomics seeks to bridge this gap by exploring associations between quantitative imaging phenotypes and underlying genomic characteristics. Early radiogenomic investigations identified meaningful relationships between imaging features and molecular alterations across multiple malignancies, including glioblastoma, breast cancer, lung cancer, and other tumor types [6]. Nevertheless, conventional statistical approaches and traditional machine learning algorithms frequently struggle to model the high-dimensional, nonlinear relationships that characterize complex multimodal oncology datasets.

Recent advances in artificial intelligence, particularly deep learning, have fundamentally transformed biomedical data analysis. Deep learning algorithms are capable of automatically learning hierarchical feature representations from large-scale heterogeneous datasets, facilitating the effective integration of imaging, genomic, pathological, and clinical information. Extending these capabilities, multimodal deep learning simultaneously processes multiple data modalities to learn unified feature representations that capture intricate biological interactions and improve predictive performance across diverse clinical tasks [7,8].

The expanding availability of multimodal oncology datasets, combined with substantial advances in computational infrastructure and neural network architectures, has accelerated research in multimodal precision oncology. Current applications encompass cancer detection, molecular subtyping, survival prediction, therapeutic decision-making, immunotherapy response evaluation, and biomarker discovery. Furthermore, the emergence of foundation models and self-supervised learning has introduced new opportunities for multimodal integration by enabling scalable representation learning from vast collections of unlabeled biomedical data, thereby improving model robustness and generalizability across clinical settings [9].

This review presents a comprehensive overview of multimodal deep learning approaches for integrating radiological imaging and genomic data in precision oncology. It summarizes the fundamental concepts, methodological frameworks, and recent developments that underpin multimodal learning while critically evaluating their clinical applications,

current limitations, and future research directions. Ultimately, this review aims to provide researchers and clinicians with an in-depth understanding of how multimodal artificial intelligence can accelerate the development and clinical implementation of personalized cancer care.

2. Fundamentals of Multimodal Deep Learning in Oncology

Multimodal deep learning encompasses artificial intelligence methodologies designed to simultaneously process and integrate information from multiple heterogeneous data sources. Within oncology, these modalities commonly include radiological imaging, genomic sequencing, digital pathology, electronic health records, laboratory investigations, and molecular biomarkers. The primary objective of multimodal learning is to leverage complementary information across diverse data modalities to improve predictive accuracy and generate a more comprehensive understanding of tumor biology than can be achieved using individual data sources alone [7,10].

Radiological imaging represents one of the most critical modalities in contemporary cancer management. Computed tomography (CT) remains the cornerstone imaging technique for detecting, staging, and monitoring numerous solid malignancies because of its widespread availability, rapid acquisition, and high spatial resolution. Magnetic resonance imaging (MRI) provides exceptional soft-tissue contrast together with advanced functional imaging capabilities, making it particularly valuable for evaluating brain, prostate, liver, and breast cancers. Positron emission tomography (PET) complements anatomical imaging by providing metabolic and molecular information that facilitates disease characterization, whereas ultrasound serves as a cost-effective and readily accessible modality for diagnosis, image-guided interventions, and treatment monitoring across a wide range of clinical applications. In parallel, digital pathology has emerged as an increasingly important imaging modality capable of capturing high-resolution microscopic information regarding tumor architecture, cellular morphology, and tissue microenvironment [11].

Genomic data provide detailed molecular characterization of cancer at multiple biological levels. Whole-genome sequencing enables comprehensive identification of genomic alterations across both coding and non-coding regions of the genome. Whole-exome sequencing focuses specifically on protein-coding regions and has become an established approach for detecting clinically actionable mutations. Transcriptomic analyses, particularly RNA sequencing, characterize gene expression profiles associated with tumor progression, prognosis, and therapeutic response. Epigenomic profiling further elucidates DNA methylation patterns and chromatin modifications, while multi-omics strategies integrate genomic, transcriptomic, proteomic, and metabolomic information to generate a systems-level understanding of cancer biology and its underlying molecular mechanisms [12].

Deep learning provides the computational foundation that enables effective multimodal data integration. Convolutional neural networks (CNNs) remain the predominant architecture for medical image analysis because of their ability to automatically learn hierarchical spatial features directly from pixel-level information. Across diverse imaging modalities, CNNs have demonstrated remarkable performance in tumor detection, segmentation, classification, and quantitative image analysis, making them indispensable tools in modern computational oncology [13].

Recurrent neural networks (RNNs), particularly long short-term memory (LSTM) networks, were originally developed to model sequential and temporal data and have been widely applied to longitudinal clinical records, time-series biomedical signals, and serial imaging studies. However, transformer-based architectures have increasingly replaced conventional RNNs because of their superior capacity to capture long-range dependencies, process large-scale datasets in parallel, and learn complex contextual relationships with greater computational efficiency [14].

Vision Transformers (ViTs) have recently emerged as powerful alternatives to convolutional neural networks for medical image analysis. By employing self-attention mechanisms, ViTs effectively capture global contextual information across entire images and have achieved state-of-the-art performance in numerous diagnostic and predictive imaging tasks. In parallel, graph neural networks (GNNs) are becoming increasingly important for modeling complex biological interactions among genes, proteins, signaling pathways, and patient populations. Their ability to represent relational structures makes them particularly well suited for integrating heterogeneous biomedical data and

uncovering biologically meaningful associations that may not be readily identified using conventional deep learning architectures [15].

The integration of these complementary deep learning architectures within multimodal frameworks enables the extraction of modality-specific representations followed by their fusion into unified latent feature spaces. These shared representations capture intricate relationships between radiological imaging phenotypes and molecular characteristics, thereby supporting more accurate predictive modeling while providing deeper biological insights into tumor behavior. As oncology datasets continue to expand in both scale and complexity, multimodal deep learning is rapidly becoming an indispensable component of precision oncology, facilitating the transformation of heterogeneous biomedical data into clinically actionable knowledge that supports personalized cancer diagnosis, prognosis, and treatment.

3. Radiogenomics: Connecting Imaging Phenotypes with Genomic Signatures

Radiogenomics is an interdisciplinary research field that seeks to identify associations between quantitative imaging characteristics and the underlying genomic, transcriptomic, and molecular alterations present within tumors [16]. The central premise of radiogenomics is that medical images reflect the macroscopic manifestations of microscopic biological processes. Consequently, imaging phenotypes can function as non-invasive surrogate biomarkers for molecular characteristics, enabling repeated assessment of tumor biology throughout the disease course without requiring invasive tissue biopsies [17].

The biological foundation of radiogenomics stems from the intrinsic heterogeneity of cancer. Individual tumors comprise multiple cellular populations with distinct genetic, epigenetic, and metabolic profiles that evolve over time. Conventional biopsy-based molecular profiling often samples only a limited region of the tumor and may therefore fail to capture its full biological diversity. In contrast, radiological imaging provides comprehensive spatial visualization of the entire tumor together with its surrounding microenvironment, offering a more holistic representation of tumor heterogeneity [18]. The application of deep learning has further advanced radiogenomic research by enabling automated extraction of high-dimensional imaging features and uncovering complex nonlinear relationships between radiological phenotypes and molecular biomarkers.

Glioblastoma represents one of the most extensively investigated malignancies in radiogenomic research. Numerous studies have demonstrated significant associations between MRI-derived imaging characteristics and key molecular biomarkers, including isocitrate dehydrogenase (IDH) mutations, epidermal growth factor receptor (EGFR) amplification, O6-methylguanine-DNA methyltransferase (MGMT) promoter methylation, and telomerase reverse transcriptase (TERT) mutations [19,20]. Deep learning–based radiogenomic models have achieved high predictive performance for identifying these molecular alterations, thereby offering the potential to reduce dependence on invasive neurosurgical procedures while supporting personalized treatment planning and prognostic evaluation [21].

Breast cancer constitutes another major area of radiogenomic investigation. MRI-based radiogenomic analyses have demonstrated associations between imaging phenotypes and intrinsic molecular subtypes, including luminal A, luminal B, HER2-enriched, and triple-negative breast cancer [22]. More recently, multimodal deep learning frameworks integrating dynamic contrast-enhanced MRI with transcriptomic data have shown superior performance in predicting treatment response and disease recurrence compared with unimodal approaches [23]. These findings underscore the growing potential of radiogenomics to facilitate patient risk stratification and support individualized therapeutic decision-making.

In lung cancer, radiogenomic research has primarily focused on non-small cell lung cancer (NSCLC). Investigators have identified imaging signatures associated with clinically actionable molecular alterations, including EGFR mutations, ALK rearrangements, KRAS mutations, and programmed death-ligand 1 (PD-L1) expression [24]. The capability to predict these biomarkers directly from routine CT examinations has the potential to improve patient selection for targeted therapies and immunotherapy while minimizing diagnostic delays and reducing the need for repeated invasive procedures [25].

Prostate cancer has likewise benefited from significant advances in radiogenomic methodologies. The integration of multiparametric MRI with comprehensive genomic profiling has

enhanced the characterization of tumor aggressiveness and improved the prediction of clinically significant disease [26]. Deep learning models that combine imaging and molecular information have demonstrated superior performance in identifying high-risk tumors, optimizing risk stratification, and guiding personalized treatment selection [27].

Similarly, radiogenomic studies in colorectal cancer have investigated associations between imaging characteristics and important molecular biomarkers, including microsatellite instability (MSI), KRAS mutations, and consensus molecular subtypes [28]. Multiple investigations have reported that multimodal integration of imaging and genomic information provides greater predictive accuracy for molecular classification and prognostic assessment than conventional radiomics-based approaches alone [29].

Despite these encouraging developments, several challenges continue to limit the broader clinical implementation of radiogenomics. Many published studies rely on relatively small patient cohorts, thereby restricting model robustness and generalizability. In addition, variability arising from differences in imaging acquisition protocols, sequencing technologies, data preprocessing methods, and patient demographics complicates external validation across institutions. Nevertheless, the continued convergence of advanced medical imaging, comprehensive genomic profiling, and multimodal deep learning is strengthening the role of radiogenomics as a foundational component of precision oncology and paving the way toward more personalized, data-driven cancer care [30].

4. Multimodal Data Fusion Strategies

The effectiveness of multimodal deep learning largely depends on how information from different data sources is integrated. Data fusion strategies determine the extent to which complementary information can be exploited and significantly influence predictive performance. Contemporary multimodal oncology systems generally employ early fusion, intermediate fusion, late fusion, or hybrid fusion architectures [31].

Early fusion, also known as feature-level fusion, combines data from different modalities before model training. In this approach, imaging features, genomic profiles, and clinical variables are concatenated into a single feature representation that serves as input to a neural network [32]. Early fusion enables direct modeling of cross-modal interactions and may capture synergistic relationships between radiological and molecular characteristics. However, the approach is sensitive to differences in feature dimensionality and data quality. High-dimensional genomic datasets can dominate learning processes, potentially reducing the contribution of imaging features [33].

Intermediate fusion, often referred to as representation-level fusion, has become increasingly popular in multimodal oncology. Separate neural networks first extract modality-specific representations from imaging and genomic data. These learned embeddings are subsequently integrated through shared layers that model cross-modal interactions [34]. This strategy preserves modality-specific information while enabling the discovery of biologically meaningful relationships. Transformer-based multimodal architectures frequently employ intermediate fusion because self-attention mechanisms effectively capture complex dependencies among heterogeneous data types [35].

Late fusion, or decision-level fusion, integrates predictions generated independently by separate modality-specific models. The outputs are combined using methods such as weighted averaging, voting schemes, or meta-learning algorithms [36]. Late fusion offers greater flexibility because individual models can be developed and optimized independently. It also provides robustness against missing data, as predictions can still be generated when one modality is unavailable. However, late fusion may fail to capture deeper biological interactions between imaging and genomic information [37].

Hybrid fusion approaches combine multiple fusion strategies to leverage their respective advantages. For example, imaging and genomic representations may first undergo intermediate fusion, followed by late-stage integration with clinical variables. Such architectures have demonstrated superior performance in several cancer prediction tasks because they preserve both modality-specific and shared information [38].

Deep learning architectures play a central role in multimodal fusion. Convolutional neural networks remain the primary choice for extracting imaging features, whereas fully connected networks, autoencoders, and transformer encoders are commonly used for genomic data [39]. Recently, graph neural networks have gained attention because biological systems naturally exhibit graph

structures involving genes, proteins, pathways, and cellular interactions [40]. By incorporating biological network information into multimodal frameworks, GNNs can improve interpretability and predictive performance.

Transformer-based architectures have emerged as particularly promising for multimodal oncology. Their self-attention mechanisms enable flexible integration of heterogeneous modalities while capturing long-range relationships within and across data types [41]. Several studies have reported that transformer-based multimodal models outperform traditional CNN-based fusion methods in cancer prognosis prediction and molecular subtype classification [42].

Despite significant advances, selecting an optimal fusion strategy remains challenging. The best approach often depends on dataset characteristics, modality availability, and clinical objectives. Future research will likely focus on adaptive fusion mechanisms capable of dynamically weighting modalities according to their relevance for specific prediction tasks [43].

5. Clinical Applications in Precision Oncology

The integration of radiological and genomic information through multimodal deep learning has produced transformative advances across multiple areas of precision oncology. These applications extend from early cancer detection to treatment optimization and long-term disease management [44].

Cancer diagnosis represents one of the most mature applications of multimodal artificial intelligence. Deep learning models that combine imaging and molecular information consistently outperform unimodal systems in distinguishing benign from malignant lesions and identifying early-stage cancers [45]. In breast cancer, multimodal approaches integrating MRI features with gene expression profiles have demonstrated improved diagnostic sensitivity and specificity compared with imaging-based assessments alone [46].

Tumor classification and molecular subtyping constitute another major area of clinical impact. Accurate classification is essential because therapeutic strategies increasingly depend on molecular characteristics rather than solely histopathological findings. Multimodal models have shown remarkable success in classifying gliomas according to IDH mutation status, breast cancers according to intrinsic molecular subtype, and lung cancers according to actionable genomic alterations [47,48]. By incorporating both phenotypic and molecular information, these systems provide more comprehensive characterization of tumor biology.

Prognosis prediction and survival analysis have emerged as particularly valuable applications. Traditional prognostic models often rely on clinical and pathological variables, which may not fully capture disease complexity. Multimodal deep learning frameworks integrating imaging biomarkers and genomic signatures have achieved superior performance in predicting overall survival, progression-free survival, and disease-specific outcomes across numerous cancer types [49]. Studies involving glioblastoma, hepatocellular carcinoma, and NSCLC have reported significant improvements in risk stratification compared with conventional approaches [50].

Treatment response prediction is increasingly important as oncology transitions toward personalized therapeutic strategies. Multimodal models can identify patients likely to benefit from specific interventions by analyzing relationships between imaging phenotypes and molecular determinants of drug sensitivity [51]. In breast cancer, combined MRI-transcriptomic models have demonstrated enhanced prediction of neoadjuvant chemotherapy response. Similarly, multimodal approaches have improved prediction of radiotherapy outcomes and targeted therapy efficacy in lung and prostate cancers [52].

Immunotherapy response prediction has become a rapidly expanding research area. Immune checkpoint inhibitors have revolutionized cancer treatment, but only a subset of patients derive substantial benefit. Deep learning systems integrating radiological features with genomic biomarkers such as tumor mutational burden, PD-L1 expression, and immune-related gene signatures have shown promising performance in predicting immunotherapy response [53]. Such models may help optimize patient selection and reduce unnecessary exposure to potentially toxic treatments [54].

Recurrence prediction and disease monitoring represent additional applications where multimodal integration offers substantial advantages. Imaging captures dynamic changes in tumor morphology and treatment response, whereas genomic profiling provides insight into residual disease and emerging resistance mechanisms. Combining these data sources enables earlier detection of recurrence and

more accurate assessment of disease progression [55].

Clinical decision support systems increasingly incorporate multimodal artificial intelligence to assist oncologists in complex treatment planning. These systems integrate radiological, genomic, pathological, and clinical data to generate personalized recommendations regarding diagnosis, prognosis, and therapeutic options [56]. Several studies have demonstrated that multimodal decision-support frameworks can improve diagnostic consistency and treatment selection while reducing clinician workload [57].

Despite encouraging results, widespread clinical adoption remains limited. Most studies remain retrospective and involve relatively small datasets. External validation across diverse populations is often lacking, and regulatory requirements for clinical implementation remain substantial [58]. Nevertheless, the growing body of evidence suggests that multimodal deep learning will play an increasingly important role in precision oncology by enabling more accurate, individualized, and data-driven cancer care.

6. Foundation Models and Large Multimodal Models in Oncology

Recent advances in artificial intelligence have led to the emergence of foundation models, a new paradigm that is rapidly transforming precision oncology. Unlike traditional deep learning systems that are developed for specific tasks using limited labeled datasets, foundation models are trained on massive and diverse datasets using self-supervised learning and can subsequently be adapted to a wide range of downstream applications with minimal additional training [59]. These models have the potential to address several longstanding challenges in oncology, including data scarcity, limited generalizability, and fragmented multimodal information integration.

Self-supervised learning has become a key enabling technology for foundation models. Instead of relying on manually annotated datasets, models learn meaningful representations from unlabeled data through pretext tasks. In medical imaging, self-supervised approaches allow neural networks to extract biologically relevant features from millions of radiological images, while in genomics, transformer-based architectures can learn patterns from large-scale sequencing data [60]. These representations often outperform conventional supervised models when transferred to cancer-specific tasks.

Contrastive learning further enhances multimodal integration by encouraging representations from related modalities to occupy similar positions within a shared latent space. For example, imaging features extracted from MRI or CT scans can be aligned with corresponding genomic signatures, enabling models to identify biologically meaningful associations between phenotype and genotype [61]. Such approaches have demonstrated improved performance in tumor classification, survival prediction, and biomarker discovery.

Transformer architectures have accelerated the development of large multimodal models capable of simultaneously processing images, genomic sequences, pathology slides, clinical notes, and electronic health records. These systems employ attention mechanisms to model relationships across heterogeneous data modalities, thereby generating comprehensive patient-level representations [62]. In oncology, transformer-based multimodal models have shown considerable promise for integrating radiological and genomic information while maintaining scalability across diverse cancer types.

Vision-language models represent another important development. Inspired by large multimodal systems such as CLIP and GPT-derived architectures, these models learn joint representations of images and textual information. In oncology, they can integrate radiology reports, pathology descriptions, molecular findings, and imaging data to support diagnosis and clinical decision-making [63]. Emerging evidence suggests that multimodal foundation models may facilitate automated report generation, treatment recommendation, and precision risk assessment.

Transfer learning has become particularly valuable in radiogenomics because many cancer datasets remain relatively small. Foundation models pretrained on large biomedical datasets can be fine-tuned for specialized oncology applications, substantially reducing data requirements while improving predictive performance [64]. Recent studies indicate that foundation models can capture generalized biological knowledge that transfers effectively across tumor types and clinical settings.

Despite their promise, foundation models remain at an early stage of clinical development. Questions regarding interpretability, fairness, computational requirements, and regulatory approval remain unresolved. Nevertheless, their ability to learn from heterogeneous multimodal datasets positions them

as a central component of future precision oncology systems.

7. Challenges and Limitations

Although multimodal deep learning has demonstrated considerable promise, several challenges continue to impede its widespread clinical implementation.

One of the most significant barriers is the limited availability of large-scale multimodal datasets. Successful deep learning models typically require thousands of patients with matched radiological, genomic, pathological, and clinical information. However, such comprehensive datasets remain scarce due to logistical, financial, and regulatory constraints [65]. Many published studies rely on relatively small cohorts collected from single institutions, increasing the risk of overfitting and reducing external validity.

Data harmonization represents another major challenge. Variability in imaging protocols, scanner manufacturers, sequencing technologies, and preprocessing pipelines can introduce substantial technical bias into multimodal datasets [66]. These inconsistencies often degrade model performance when systems are applied across institutions or populations. Standardized acquisition and preprocessing frameworks are therefore essential for achieving robust and reproducible results.

Missing data are particularly problematic in precision oncology. Not all patients undergo comprehensive genomic profiling, and imaging examinations may vary according to clinical circumstances. Conventional deep learning models frequently struggle when one or more modalities are unavailable [67]. Developing flexible architectures capable of accommodating incomplete multimodal data remains an active area of research.

Batch effects pose additional challenges, especially in genomic and transcriptomic datasets. Technical variation arising from differences in laboratory procedures can obscure biologically meaningful signals and compromise predictive performance [68]. Advanced normalization strategies and domain adaptation techniques are increasingly being incorporated into multimodal frameworks to address this issue.

Interpretability and explainability remain critical concerns for clinical adoption. Deep learning models often function as "black boxes," making it difficult for clinicians to understand the rationale behind specific predictions [69]. This lack of transparency can undermine trust and limit integration into clinical workflows. Explainable artificial intelligence (XAI) techniques, including attention visualization, saliency mapping, and graph-based explanations, are being developed to improve model interpretability.

Generalizability represents another important limitation. Many multimodal models demonstrate excellent performance during internal validation but experience substantial performance degradation when evaluated in independent cohorts [70]. Differences in patient demographics, disease prevalence, healthcare systems, and data acquisition methods contribute to these challenges.

Algorithmic bias is increasingly recognized as a significant concern. Underrepresentation of certain demographic groups in training datasets may lead to disparities in model performance and potentially exacerbate healthcare inequities [71]. Ensuring fairness and inclusivity in model development is therefore essential.

Privacy and data security also present substantial obstacles. Genomic information is inherently identifiable, and sharing multimodal datasets across institutions raises important ethical and legal concerns [72]. Regulatory frameworks such as GDPR and HIPAA impose strict requirements regarding patient data protection, which can complicate collaborative research initiatives.

Finally, regulatory approval and clinical validation remain major hurdles. Most multimodal AI systems have not yet undergone prospective multicenter evaluation, and regulatory agencies continue to develop frameworks for assessing the safety and effectiveness of adaptive machine learning algorithms [73]. Addressing these challenges will be crucial for translating research advances into routine clinical practice.

8. Future Perspectives

The future of multimodal deep learning in precision oncology will likely be shaped by several emerging technological and clinical developments.

Federated learning is expected to play a transformative role by enabling institutions to collaboratively train AI models without directly sharing sensitive patient data [74]. This approach can increase dataset diversity while maintaining privacy and regulatory compliance. Early studies suggest that federated learning may improve model robustness and generalizability across geographically distributed healthcare systems.

Explainable artificial intelligence will become increasingly important as multimodal models transition toward clinical deployment. Future systems are expected to provide transparent reasoning pathways that connect radiological findings, genomic alterations, and clinical outcomes. Such capabilities will improve clinician trust and facilitate regulatory approval [75].

Digital twins represent another promising direction. By integrating longitudinal imaging, genomic, clinical, and treatment data, digital twins may create dynamic computational representations of individual patients. These virtual models could enable simulation of disease progression and prediction of therapeutic responses before treatment initiation [76].

Foundation models are likely to become the dominant computational paradigm in oncology. As larger multimodal datasets become available, these models will increasingly learn generalized representations that can be adapted to numerous downstream tasks. Integration of radiology, pathology, genomics, and electronic health records within unified foundation models may substantially enhance predictive accuracy and clinical utility.

Multicenter prospective validation studies will be essential for establishing clinical credibility. Future investigations must demonstrate reproducibility across diverse populations and healthcare environments while evaluating real-world clinical impact [77].

Real-world evidence generated through routine clinical practice will also contribute to model refinement and validation. Continuous learning systems capable of incorporating new data may facilitate adaptive precision oncology strategies that evolve alongside emerging evidence [78].

Ultimately, successful integration of multimodal deep learning into clinical workflows will require close collaboration among oncologists, radiologists, bioinformaticians, data scientists, regulators, and healthcare administrators. Such interdisciplinary efforts are essential for ensuring that artificial intelligence technologies improve patient outcomes while maintaining safety, equity, and transparency.

9. Conclusion

Multimodal deep learning has emerged as a powerful framework for integrating radiological and genomic data in precision oncology. By combining complementary information from imaging and molecular profiling, these approaches provide a more comprehensive characterization of tumor biology than traditional unimodal methods. Advances in radiogenomics, multimodal fusion architectures, and transformer-based learning have significantly improved cancer diagnosis, molecular subtyping, prognosis prediction, treatment response assessment, and biomarker discovery.

The recent emergence of foundation models and large multimodal systems represents a major step toward scalable and generalizable artificial intelligence in oncology. These technologies have the potential to overcome limitations associated with small datasets and fragmented information sources while enabling more personalized and data-driven clinical decision-making.

Despite encouraging progress, substantial challenges remain. Data scarcity, heterogeneity, missing modalities, limited interpretability, privacy concerns, algorithmic bias, and regulatory barriers continue to restrict widespread clinical implementation. Addressing these limitations will require standardized datasets, explainable AI methodologies, multicenter validation studies, and robust governance frameworks.

As federated learning, foundation models, digital twins, and real-world evidence infrastructures continue to evolve, multimodal deep learning is expected to become an integral component of future precision oncology ecosystems. Continued interdisciplinary collaboration will be essential for translating these innovations into clinically actionable tools that improve cancer outcomes and advance personalized medicine

References

1. Hanahan D. Hallmarks of cancer: new dimensions. Cancer Discov. 2022;12(1):31–46. Doi:10.1158/2159-8290.CD-21-1059.
2. Rajendran LKK. Hematological Malignancy Identification via K-means based ROI Extraction. International Journal of Clinical Research in Medical Sciences. 2026;1(2):1-10. Doi:10.67231/kt1w3e73.
3. Kumar RMH. Pan-System Cancer Intelligence: Integrating Blood, Immune, Microbiome, and Tumor Microenvironment Data Using Foundation Models. Power System Protection and Control. 2023;51(4):92-100. Doi:10.46121/pspc.51.4.8.
4. Rajendran LKK. Identifying Determinants of Outcome in Post-Radiotherapy Cervical Carcinoma Requiring Adjuvant Surgery.

International Journal of Clinical Research in Medical Sciences. 2026;1(2):1-10. Doi:10.67231/3acej759.
5. Maradi Hemanth Kumar R. AI-Driven Liquid Biopsy Systems for Early Cancer Detection and Personalized Oncology. Power System Protection and Control. 2023;51(4):66-83. Doi:10.46121/pspc.51.4.7.
6. Rajendran LKK. Machine Learning–Driven Symptom-Based Cancer Risk Stratification:A Systematic Review of Clinical Prediction Models and Methodological Rigor. Int J Drug Deliv Technol. 2026;16(40s):242-253.Doi:10.25258/ijddt.16.40s.26.
7. Rajendran OK. Bias, Fairness, and Ethical Challenges in Artificial Intelligence: A Comprehensive Review of Causes, Impacts, and Mitigation Strategies. Scientific Culture. 2026;12(2.1):13001-13010. Doi:10.5281/zenodo.20374091.
8. Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. Nat Med. 2019;25(1):44–56. Doi:10.1038/s41591-018-0300-7.
9. Rajendran OK. Clinical Translation of Artificial Intelligence in Oncology: Real-World Validation, Workflow Integration, and Precision Medicine Applications. Int J Drug Deliv Technol. 2026;16(49s):956-964. Doi:10.25258/ijddt.16.49s.110.
10. 10.Rajendran LKK. Interpretable Machine Learning for Early Mortality Prediction in Acute Myeloid Leukemia: A Decision Tree–Based Retrospective Cohort Study. Int J Drug Deliv Technol. 2026;16(40s):231-241.Doi:10.25258/ijddt.16.40s.25.
11. 11.Esteva A, Robicquet A, Ramsundar B, et al. A guide to deep learning in healthcare. Nat Med. 2019;25(1):24–29. Doi:10.1038/s41591-018-0316-z.
12. 12.Rajendran OK. Generative AI for Synthetic Medical Image Generation in Oncology: Addressing Data Scarcity in AI-Driven Cancer Diagnosis. Int J Drug Deliv Technol. 2026;16(49s):1010-1016. Doi:10.25258/ijddt.16.49s.117.
13. 13.Rajendran LKK. Integrated Prognostic Modeling of Tumor Stage, Multimodal Therapy, and Functional Status in Lung Cancer Survival: A Real-World Cohort Study. Scientific Culture. 2026;12(5):567-576. Doi:10.5281/zenodo.1250046.
14. 14.Bommasani R, Hudson DA, Adeli E, et al. On the opportunities and risks of foundation models. arXiv. 2021. Doi:10.48550/arXiv.2108.07258.
15. 15.Rajendran LKK. Integrative Pharmacogenomic Analysis of Drug Response Heterogeneity Across Cancer Cell Lines: Insights From Large-Scale GDSC Data. Scientific Culture. 2026;12(4):7537-7546. Doi:10.5281/zenodo.12426762.
16. 16.Acs B, Rantalainen M, Hartman J. Artificial intelligence as the next step towards precision pathology. J Intern Med. 2020;288(1):62–81. Doi:10.1111/joim.13030.
17. 17.Rajendran OK. Tumor Microenvironment Interaction-Guided Graph Neural Networks for Survival Prediction from Whole-Slide Pathology Images. Int J Drug Deliv Technol. 2026;16(49s):481-488. Doi:10.25258/ijddt.16.49s.50.
18. 18.Rajendran LKK. Evaluating the Association of Cancer-Related Risk Factors With Multisystem Health: Insights Into Fertility, Cardiovascular, and Renal Indicators. Scientific Culture. 2026;12(4):7520-7527.Doi:10.5281/zenodo.12426760.
19. 19.Rajendran LKK. From Prediction to Precision: An Externally Validated Deep Learning–Based Survival and Adjuvant Therapy Recommendation System for Resected Stage III Non–Small Cell Lung Cancer. Int J Drug Deliv Technol. 2026;16(30s):430-438.doi:10.25258/ijddt.16.30s.41.
20. 20.Chen RJ, Lu MY, Wang J, et al. Pathomic fusion: an integrated framework for fusing histopathology and genomic features for cancer diagnosis and prognosis. Nat Mach Intell. 2022;4:179–193. Doi:10.1038/s42256-022-00466-x.
21. 21.Rajendran LKK. From Prediction to Practice: A Machine Learning–Based Clinical Decision Support Tool for Bevacizumab Risk Stratification in Oncology. Int J Drug Deliv Technol. 2026;16(30s):414-429. Doi:10.25258/ijddt.16.30s.40.
22. 22.Rajendran OK. Self-supervised multimodal Learning for early cancer detection across Imaging and genomics. Power System Protection and Control. 2024;52(4):167-178. Doi:10.46121/pspc.52.4.14.
23. 23.Rajendran OK. Explainable AI-Driven Clinical. Decision Support Systems in Precision

Oncology:Interpretable Models for Multimodal Cancer Care. Scientific Culture. 2026;12(2.1):12359-12369. Doi:10.5281/zenodo.20328194.
24. 24.Rajendran LKK. Impact of Treatment Modalities on Fertility, Sexual Function, and Psychological Outcomes in Testicular Cancer Survivors: A Comprehensive Review. Int J Drug Deliv Technol. 2026;16(30s):447-453. Doi:10.25258/ijddt.16.30s.43.
25. 25.Rajendran LKK. Intelligent Omics-Driven Patient Stratification for Cancer Therapeutic Re-profiling. International Journal of Clinical Research in Medical Sciences. 2026;1(1):1-11. Doi:10.67231/gv5hck05.
26. 26.Rajendran LKK. Cancer nanomedicine: utilizing the enhanced permeability and retention (EPR) effect to deliver high payloads of chemotherapeutic agents directly to tumor sites. Power System Protection and Control. 2024;52(2):123-129. Doi:10.46121/pspc.52.2.12.
27. 27.Kather JN, Calderaro J. Development of AI in digital pathology. Nat Rev Clin Oncol. 2020;17(10):591–595. Doi:10.1038/s41571-020-00431-0.
28. 28.Rajendran OK. AI-based radiogenomic Models for predicting immunotherapy response In solid tumors. Power System Protection and Control. 2023;51(4):24-37. Doi:10.46121/pspc.51.4.4.
29. 29.Rajendran LKK. Enhanced Predictive Analytics for Early Malignancy Discovery in Routine Screening. International Journal of Clinical Research in Medical Sciences. 2026;1(1):1-10. Doi:10.67231/grams870.
30. 30.Wan JCM, Massie C, Garcia-Corbacho J, et al. Liquid biopsies come of age: towards implementation of circulating tumour DNA. Nat Rev Cancer. 2017;17(4):223–238. Doi:10.1038/nrc.2017.7.
31. 31.Rajendran OK. Machine Learning-Based Prediction of Chemotherapy Toxicity in Colorectal Cancer: A Personalized Risk Stratification Approach. Scientific Culture. 2026;12(5.1):942-952. Doi:10.5281/zenodo.12511075.
32. 32.Rajendran OK. Federated radiology AI Models for multi-institutional cancer diagnosis Without data sharing. Power System Protection And Control. 2023;51(4):38-54. Doi:10.46121/pspc.51.4.5.
33. 33.Bera K, Schalper KA, Rimm DL, et al. Artificial intelligence in digital pathology — new tools for diagnosis and precision oncology. Nat Rev Clin Oncol. 2019;16(11):703–715. Doi:10.1038/s41571-019-0252-y.
34. 34.Rajendran OK. Deep Reinforcement Learning in Oncology: Advances in Cancer Imaging, Radiotherapy, and Personalized Treatment. Scientific Culture. 2026;12(5):597-606. Doi:10.5281/zenodo.1250048.
35. 35.Rajendran OK. DEEP LEARNING FOR CROSS-MODALITY MAPPING BETWEEN HISTOPATHOLOGY AND RADIOLOGICAL IMAGING. Power System Protection and Control. 2025;53(3):313-328. Doi:10.46121/pspc.53.3.21.
36. 36. Lu MY, Chen TY, Williamson DFK, et al. AI-based pathology predicts origins for cancers of unknown primary. Nature. 2021;594(7861):106–110. Doi:10.1038/s41586-021-03512-4.
37. 37.Rajendran OK. Artificial Intelligence in Oncologic Imaging: Deep Learning, Radiomics, and Clinical Integration for Precision Cancer Diagnosis. Int J Drug Deliv Technol. 2026;16(50s):871-880. Doi:10.25258/ijddt.16.50s.92.
38. 38.Bilal M, Raza SEA, Azam A, et al. Development and validation of a weakly supervised deep learning framework to predict the risk of colorectal cancer recurrence from histology images. Lancet Oncol. 2021;22(11):153–163. Doi:10.1016/S1470-2045(21)00430-5.
39. 39.Rajendran OK. DIGITAL TWIN FRAMEWORKS FOR PERSONALIZED CANCER PROGRESSION MODELING USING LONGITUDINAL DATA. Power System Protection and Control. 2025;53(4):486-501. Doi:10.46121/pspc.53.4.33.
40. 40.Rajendran LKK. Genomic profiling: utilizing Multi-omics data to identify potential Therapeutic targets and resistance markers. Power System Protection and Control. 2024;52(4):159-166. Doi:10.46121/pspc.52.4.13.
41. 41.Rajendran OK. Artificial Intelligence–Driven Multimodal Imaging for Cancer During Pregnancy: Advances in Maternal–Fetal Diagnostics and Precision Oncology. Int J Drug Deliv Technol. 2026;16(50s):862-870. Doi:10.25258/ijddt.16.50s.91.

42. 42.Rajendran LKK. Immunotherapy and cell Therapy: developing CAR-T cell therapies and Other immune-based treatments for cancer and Autoimmune diseases. Power System Protection and Control. 2023;51(2):64-77. Doi:10.46121/pspc.51.2.7.
43. 43.Rajendran OK. FOUNDATION MODEL–DRIVEN PRECISION ONCOLOGY: INTEGRATING MULTI-OMICS, RADIOLOGY, AND CLINICAL DATA FOR PREDICTIVE CANCER CARE. Power System Protection and Control. 2024;52(2):154-163. Doi:10.46121/pspc.52.2.14.
44. 44.Rajendran LKK. Theranostics: integrating Diagnostic imaging agents and therapeutic Drugs into a single multifunctional nano-Platform for real-time monitoring of treatment. Power System Protection and Control.2025;53(2):376-386. Doi:10.46121/pspc.53.2.31.
45. 45.Rajendran LKK. Mechanisms driving Immunotherapy resistance in colorectal cancer Liver metastases. Power System Protection and Control. 2024;52(1):29-37. Doi:10.46121/pspc.52.1.5.
46. 46.Ching T, Himmelstein DS, Beaulieu-Jones BK, et al. Opportunities and obstacles for deep learning in biology and medicine. J R Soc Interface. 2018;15(141):20170387. Doi:10.1098/rsif.2017.0387.
47. 47. Litjens G, Kooi T, Bejnordi BE, et al. A survey on deep learning in medical image analysis. Med Image Anal. 2017;42:60–88. Doi:10.1016/j.media.2017.07.005.
48. 48.Hemanth Kumar RM. Integrated Transcriptomic and 3 Learning Framework Identifies a Blood-Based Biomarker Signature for Anthracycline-Induced Cardiotoxicity in Juvenile Cancer Survivors. Int J Drug Deliv Technol. 2026;16(40s):219-230.
49. Doi:10.25258/ijddt.16.40s.24.
50. 49.Mobadersany P, Yousefi S, Amgad M, et al. Predicting cancer outcomes from histology and genomics using convolutional networks. Proc Natl Acad Sci USA. 2018;115(13):E2970–E2979. Doi:10.1073/pnas.1717139115.
51. 50. Lambin P, Leijenaar RTH, Deist TM, et al. Radiomics: the bridge between medical imaging and personalized medicine. Nat Rev Clin Oncol. 2017;14(12):749–762. Doi:10.1038/nrclinonc.2017.141.
52. 51.Azizi S, Mustafa B, Ryan F, et al. Big self-supervised models advance medical image classification. Nature. 2021;594(7864):104–110. Doi:10.1038/s41586-021-03476-6.
53. 52.Dosovitskiy A, Beyer L, Kolesnikov A, et al. An image is worth 16×16 words: transformers for image recognition at scale. arXiv. 2020. Doi:10.48550/arXiv.2010.11929.
54. 53.Rajendran OK. DeepDRA: A Deep Learning Framework for Drug Repurposing and Cancer Drug Response Prediction Using Multi-Omics Data. Scientific Culture. 2026;12(3):68-77. Doi:10.5281/zenodo.12326001.
55. 54.Xu H, Usuyama N, Bagga J, et al. A whole-slide foundation model for digital pathology from real-world data. Nature. 2024;630(8015):181–188. Doi:10.1038/s41586-024-07441-w.
56. 55.Singhal K, Azizi S, Tu T, et al. Large language models encode clinical knowledge. Nature.2023;620(7972):172–180. Doi:10.1038/s41586-023-06291-2.
57. 56.Moor M, Banerjee O, Abad ZSH, et al. Foundation models for generalist medical artificial intelligence. Nature. 2023;616(7956):259–265. Doi:10.1038/s41586-023-05881-4.
58. 57.Chen RJ, Ding T, Lu MY, et al. Towards a general-purpose foundation model for computational pathology. Nat Med. 2024;30(3):850–862. Doi:10.1038/s41591-024-02857-3.
59. 58. Dr. Isabella Moore, MultimodalArtificial Intelligence in Oncology: IntegratingRadiomics, Pathomics, and Genomics, Int. J. of Pharm.Sci., 2026, Vol 4, Issue 5, 6745-6760.https://doi.org/10.5281/zenodo.20391855
59. Dr. Benjamin Walker, Dr. EleanorHayes, Dr. Christopher Nolan, Foundation Models inCancer Medicine: Revolutionizing Precision Diagnosticsand Clinical Oncology, Int. J. of Pharm. Sci., 2026, Vol4, Issue 5, 6733-6744.https://doi.org/10.5281/zenodo.20391588