Intelligent Oncology 1 (2025) 204–215

Contents lists available at ScienceDirect

Intelligent Oncology

journal homepage: www.keaipublishing.com/en/intonc/

Review

Deep learning-based multimodal data fusion in bone tumor management: Advances in clinical decision support

Tongtong Huo[a,#], Wei Wu[b,#], Xiaoliang Chen[b,#], Mingdi Xue[b], Pengran Liu[b], Jiayao Zhang[b], Yi Xie[b], Honglin Wang[b], Hong Zhou[b], Zineng Yan[b], Songxiang Liu[b], Lin Lu[b], Jiaming Yang[b], Jin Liu[a,*], Zhewei Ye[b,*]

[a] School of Electronic Information, Wuhan University of Science and Technology, Wuhan Hubei 430081, China
[b] Union Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan Hubei 430022, China

ARTICLE INFO

Keywords:
Bone tumors
Multimodal data fusion
Artificial intelligence
Clinical decision support systems
Deep learning

ABSTRACT

Bone tumors (BTs)—including osteosarcoma, Ewing sarcoma, and chondrosarcoma—are rare but biologically complex malignancies characterized by pronounced heterogeneity in anatomical location, histological subtype, and molecular alterations. Recent advances in artificial intelligence (AI), particularly deep learning, have enabled the integration of diverse clinical data modalities to support diagnosis, treatment planning, and prognostication in bone oncology. This review provides a comprehensive synthesis of AI-driven multimodal fusion strategies that incorporate radiological imaging, digital pathology, multi-omics profiling, and electronic health records. We conducted a structured review of peer-reviewed literature published between 2015 and early 2025, focusing on the development, validation, and clinical applicability of AI models for BT diagnosis, subtyping, treatment response prediction, and recurrence monitoring. Although multimodal models have demonstrated advantages over unimodal approaches, especially in handling missing data and improving generalizability, most remain constrained by single-center study designs, small sample sizes, and limited prospective or external validation. Persistent technical and translational challenges include semantic misalignment across modalities, incomplete datasets, limited model interpretability, and regulatory and infrastructural barriers to clinical integration. To address these limitations, we highlight emerging directions such as contrastive representation learning, generative data augmentation, transformer-based fusion architectures, and privacy-preserving federated learning. We also discuss the evolving role of foundation models and workflow-integrated AI agents in enhancing scalability and clinical usability. In summary, multimodal AI represents a promising paradigm for advancing precision care in BTs. Realizing its full clinical potential will require methodologically rigorous, biologically informed, and system-level approaches that bridge algorithmic innovation with real-world healthcare delivery.

1. Introduction

Bone tumors (BTs) —including osteosarcoma (OS), Ewing sarcoma, and chondrosarcoma—are rare but biologically diverse neoplasms that originate from bone or adjacent soft tissues.[1] These tumors exhibit substantial heterogeneity in growth patterns, histopathological features, and molecular alterations,[2] leading to variable clinical courses and often suboptimal outcomes in advanced stages.[3] This complexity presents significant challenges for accurate diagnosis, risk stratification, and treatment planning, and complicates the design of artificial intelligence (AI)-driven clinical decision support systems.[4]

BTs are broadly classified as primary (e.g., OS, Ewing sarcoma, chondrosarcoma) or secondary (metastatic), with the latter often originating from breast, lung, or prostate malignancies.[5] Effective management of both types requires multidisciplinary coordination and intelligent systems capable of integrating fragmented, heterogeneous, and patient-specific data across the clinical continuum.[6,7]

In routine clinical workflows, BTs evaluation involves synthesizing diverse data modalities, including radiological imaging,[8] digital pathology,[9] high-throughput multi-omics profiles, and both structured and unstructured clinical records.[10] Each modality contributes distinct information—ranging from anatomical localization to cellular morphology and molecular signatures—yet their asynchronous acquisition,

* Corresponding authors.
E-mail addresses: liujin@wust.edu.cn (J. Liu), yezhewei@hust.edu.cn (Z. Ye).
# These authors contributed equally to this work.

https://doi.org/10.1016/j.intonc.2025.06.005
Received 16 April 2025; Received in revised form 15 June 2025; Accepted 16 June 2025
Available online 3 July 2025
2950-2616/© 2025 The Author(s). Publishing services by Elsevier B.V. on behalf of KeAi Communications Co. Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

incomplete availability, and inconsistent formats pose major obstacles to both clinical interpretation and computational modeling.[11]

Recent advances in AI, particularly deep learning (DL), have shown promise in integrating multimodal data into unified, task-driven models.[12] Over the past decade, unimodal AI systems have demonstrated strong performance in tasks such as lesion classification, histological grading, and mutation detection.[13–18] However, the truly transformative potential lies in multimodal approaches that combine structural, functional, molecular, and clinical features to generate temporally contextualized and patient-specific predictions (Figure 1).[19,20]

Such multimodal AI systems offer the ability to model the dynamic trajectories of individual tumors, capturing cross-modal relationships and compensating for missing or incomplete data. When successfully deployed, these systems can enhance early diagnosis, guide individualized treatment planning, and improve outcome prediction through mechanism-aware modeling that reflects the full biological complexity of bone sarcomas.[21]

Despite promising developments in other cancer types, such as lung, breast, and brain, applications of multimodal AI in BTs remain nascent.[14] This is largely due to disease rarity, data fragmentation, inconsistent annotations, and limited access to large, multi-institutional datasets for training and validation.[22]

This review aims to provide a comprehensive synthesis of how AI, particularly DL, has been applied to multimodal data fusion in BTs care. We examine the integration of imaging, pathology, omics, and clinical data for diagnostic, therapeutic, and prognostic purposes, with a focus on recent technical advances, clinical applications, and translational challenges. To ensure methodological rigor, we conducted a structured literature review of English-language publications from 2015 to early 2025, using PubMed and ScienceDirect with keywords such as “bone tumors,” “multimodal data fusion,” “clinical decision support,” and “artificial intelligence.”

Finally, this review highlights persistent barriers—including model interpretability, external validation, and workflow integration—and proposes future directions for building biomedically informed, clinically deployable, and ethically sound AI systems. Our goal is to provide a roadmap toward intelligent decision support tools that are both biomedically informed and practically deployable in bone oncology.

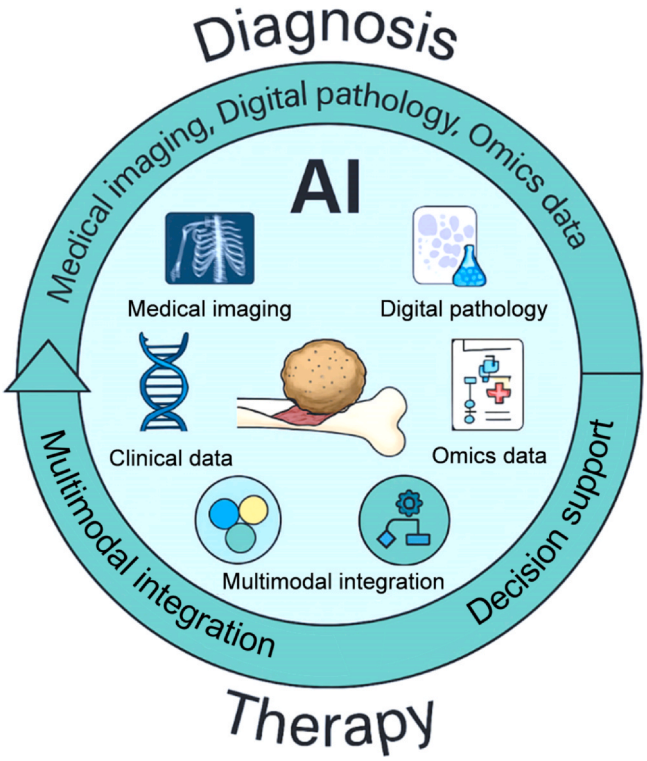


Figure 1. AI-powered multimodal integration in bone tumor care.

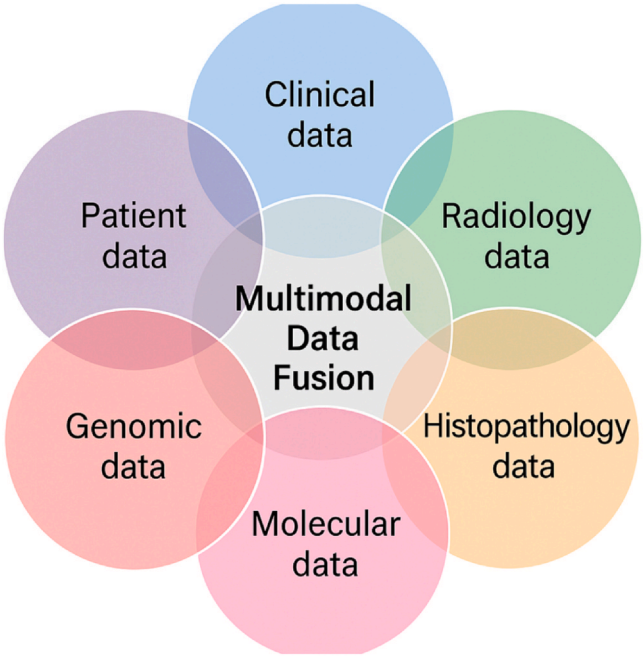


Figure 2. Core data types involved in multimodal fusion for bone tumor analysis.

2. Multimodal data in bone tumor management

The clinical management of BTs is inherently complex, requiring insights from a broad spectrum of data sources across diagnostic, therapeutic, and longitudinal follow-up stages.[3] This data ecosystem is increasingly defined by multimodal inputs—including medical imaging, digital pathology, multi-omics profiling, and clinical records—each reflecting a distinct layer of tumor biology[23]: anatomical, cellular, or molecular (Figure 2).

Although complementary, these modalities are often acquired asynchronously, stored in siloed systems, and vary widely in format and completeness.[24] This heterogeneity presents major challenges for both clinical reasoning and computational modeling, limiting the potential of conventional data interpretation frameworks and underscoring the urgent need for robust, AI-driven multimodal fusion (MF) strategies (Table 1).[25]

2.1. *Medical imaging: structural and functional insights*

Medical imaging constitutes a cornerstone of BTs evaluation, offering critical insights into lesion morphology, spatial extent, and biological activity.[26] Common modalities include radiographs, computed tomography (CT), magnetic resonance imaging (MRI), and positron emission tomography–computed tomography (PET-CT), each contributing distinct spatial and functional characteristics.[27] Radiographs are typically employed for initial assessment, providing a cost-effective means to detect calcifications and osteolytic changes.[28] CT offers high-resolution visualization of cortical bone integrity, mineralization patterns, and anatomical landmarks critical for surgical planning.[29] MRI excels at delineating soft tissue contrast, marrow involvement, and neurovascular structures.[30] PET-CT incorporates metabolic imaging to assist in disease staging and early detection of metastatic spread.[31]

Despite their strengths, these imaging modalities vary in acquisition protocols, spatial resolution, and contrast mechanisms, thereby complicating cross-modal alignment and standardization efforts. These discrepancies pose significant obstacles for AI-driven fusion models, particularly in tasks requiring temporal tracking or image-to-image registration.

Table 1
Core data modalities, information attributes, and AI-relevant challenges in bone tumor management.

Modality	Data source	Encoded information	Clinical value	AI-specific challenges	Clinical applications
Radiologic imaging (X-ray, CT, MRI, PET)	Radiology departments; imaging systems	Anatomical structures, tumor boundaries, functional/metabolic activity	Non-invasive visualization, staging, surgical planning	Multi-vendor heterogeneity, resolution mismatch, registration difficulty	Diagnosis, tumor characterization, treatment response assessment
Digital pathology (H&E, IHC, WSI)	Pathology labs; whole-slide scanners	Cell morphology, tissue architecture, necrosis, mitotic rate	Histologic subtyping, grading, outcome prediction	Large file size, staining variability, annotation burden	Definitive diagnosis, recurrence prediction, survival modeling
Multi-omics profiling (WES, RNA-Seq, Methylation)	Genomic/transcriptomic labs; TCGA, GEO	Mutations, gene expression, fusion events, epigenetic marks	Therapeutic target identification, risk stratification	High dimensionality, sparsity, cross-platform inconsistency	Subtype classification, personalized therapy selection
Clinical records (Structured & Free-text)	EHR systems	Demographics, lab values, medication history, narrative notes	Longitudinal insight into patient status and response	Terminology inconsistency, unstructured text, missing data	Risk prediction, care pathway optimization, clinical decision support

Abbreviations: 1. AI: artificial intelligence 2. CT: computed tomography 3. MRI: magnetic resonance imaging 3. PET: positron emission tomography 4. H&E: hematoxylin and eosin staining 5. IHC: immunohistochemistry 6. WSI: whole-slide imaging 7. WES: whole exome sequencing 8. RNA-Seq: RNA sequencing 9. TCGA: The Cancer Genome Atlas 10. GEO: Gene Expression Omnibus 11. EHR: electronic health record.

Radiomics and DL have expanded analytical capabilities—enabling automated segmentation, malignancy grading, and treatment response prediction—yet their clinical deployment remains limited by interoperability issues and the absence of standardized preprocessing pipelines.[32–35]

Robust AI integration requires imaging workflows to incorporate harmonization techniques, including resolution normalization, artifact mitigation, and modality-aware domain adaptation. Architectures should support variable input formats while maintaining semantic fidelity across institutions and scanner platforms.[36,37]

2.2. Digital pathology: microstructural resolution

Histopathological analysis remains the definitive diagnostic gold standard for BTs classification and grading.[9] Digitizing whole-slide images (WSIs) at sub-micron resolution enables computational interrogation of key morphological features, including nuclear atypia, necrosis, mitotic activity, and extracellular matrix composition.[38]

However, digital pathology in BTs is still in its early stages compared to epithelial malignancies.[39] Challenges include the vast size of WSIs, variability in staining protocols, and limited availability of high-quality, annotated datasets.[40] Bone-specific complexities, such as calcification artifacts and heterogeneous tissue architecture, further hinder the development and generalization of computational models.

Emerging approaches, such as weakly supervised learning, multiple instance learning, and transformer-based attention mechanisms, have shown promise in reducing annotation burdens and improving model robustness.[41,42] Nonetheless, the absence of standardized preprocessing workflows and limited cross-modal interoperability with imaging or omics data continue to hinder their translational applicability.[20]

2.3. Multi-omics profiling: molecular signatures

Genomic and epigenomic alterations play a central role in BT pathogenesis. Common examples include tumor protein p53 and retinoblastoma 1 mutations in OS, EWS RNA binding protein 1–friend leukemia virus integration 1 gene fusions in Ewing sarcoma, and isocitrate dehydrogenase 1 and 2 mutations in chondrosarcoma.[43] These molecular aberrations serve as both diagnostic biomarkers and therapeutic targets, providing high-value inputs for AI models in precision oncology.[44]

High-throughput platforms such as whole exome sequencing, RNA sequencing, and DNA methylation profiling allow for comprehensive tumor characterization.[45,46] However, omics data are often high-dimensional, sparse, and heterogeneous across institutions.[47] These properties elevate the risk of overfitting and complicate integration into AI pipelines, particularly for rare tumors with inherently limited cohort sizes.

Recent efforts to fuse transcriptomic features with imaging phenotypes (radiogenomics) have demonstrated improved performance in recurrence prediction and treatment stratification.[48] Nonetheless, persistent challenges remain, including semantic misalignment between omics and imaging modalities and the lack of biologically grounded cross-modal mappings.[49]

Future research should prioritize representation learning techniques that support cross-modal embedding, noise tolerance, and semantic consistency, especially under conditions of missing or incomplete omics data.[50,51]

2.4. Clinical data: capturing patient background and dynamic trajectories

Electronic health records offer longitudinal insights into patient health trajectories, capturing both structured data (e.g., demographics, lab results, medication history) and unstructured narratives (e.g., physician notes, operative reports).[52] Structured biomarkers such as alkaline phosphatase, lactate dehydrogenase, and C-reactive protein are frequently utilized for risk assessment and prognostic modeling.

Unstructured clinical narratives capture nuanced descriptions of presenting symptoms, therapeutic interventions, and longitudinal disease progression.[53]

Advancements in natural language processing have enabled the extraction of clinically relevant features from free-text data through named entity recognition, temporal mapping, and semantic embedding.[54] However, electronic health records data are frequently characterized by inconsistent terminology, incomplete entries, and institution-specific documentation styles, which undermine generalizability and model reproducibility.[55]

To fully leverage clinical records, AI models must incorporate robust preprocessing strategies, ontology mapping, and techniques for handling noisy or incomplete input streams. Systems should be capable of capturing longitudinal context and accommodating fragmented data typical of real-world settings.[56]

2.5. Summary

BTs care is underpinned by a highly heterogeneous and temporally distributed data ecosystem encompassing structural imaging, microscopic pathology, molecular profiling, and evolving clinical records. While this data richness enables integrative modeling opportunities, it also introduces substantial challenges in data harmonization, semantic integration, and handling of incomplete inputs.

To address these challenges, AI models must be designed with flexible architectures, standardized preprocessing pipelines, and adaptive learning mechanisms. Building on the diverse yet fragmented multimodal landscape discussed above, the next section explores how AI-driven fusion strategies can integrate these modalities into cohesive, clinically useful models.

3. MF strategies in AI-assisted bone tumor care

Multimodal data fusion serves as a cornerstone for AI-driven clinical decision support in bone oncology.[57] Owing to the anatomical complexity, histopathological diversity, and molecular heterogeneity of BTs, unimodal approaches—while useful for specific tasks—often lack generalizability and robustness.[58,59] In contrast, MF frameworks aim to integrate heterogeneous data types, such as imaging, pathology, genomics, and clinical records, into unified representations that enable context-aware and personalized inference (Figure 3).[60]

By capturing inter-modal relationships and aligning semantic content across data sources, these frameworks enhance predictive performance, model interpretability, and resilience to missing or noisy inputs.[61] This section outlines the evolution of MF strategies—from classical paradigms to state-of-the-art DL architectures—highlighting their design principles, advantages, limitations, and clinical relevance in BTs applications.

3.1. Classical fusion paradigms: early, intermediate, and late fusion

Foundational MF strategies are commonly categorized into three paradigms—early, intermediate, and late fusion—based on the stage in the modeling pipeline at which data integration is performed.[57]

Early fusion involves combining raw or low-level features from multiple modalities at the input level. This approach is straightforward and effective when input formats are inherently compatible, such as structured clinical data and gene expression vectors.[62,63] However, early fusion often encounters difficulties when integrating heterogeneous modalities, such as three-dimensional imaging and histopathology slides, due to scale mismatches and insufficient semantic abstraction.

Intermediate fusion involves extracting modality-specific features through dedicated encoders, which are subsequently merged into a shared latent representation.[64,65] This approach facilitates semantic alignment and cross-modal reasoning, making it especially well-suited for tasks such as subtype classification and treatment response prediction. Intermediate fusion is currently the dominant strategy in bone oncology AI due to its balance between interpretability and flexibility.

Late fusion combines independent predictions from each modality using ensemble techniques such as majority voting, weighted averaging, or stacking.[61,66] This modular strategy is notably robust in real-world settings where data availability is inconsistent—for example, between community hospitals and tertiary referral centers. However, it lacks cross-modal interactions and is incapable of modeling synergistic relationships among modalities. Table 2 provides a comparative summary of the core characteristics and trade-offs of these fusion strategies within the context of BTs care.

3.2. Modern fusion architectures

Contemporary multimodal AI systems in BTs advance beyond classical fusion paradigms by enabling adaptive information flow and context-sensitive feature interactions. Two key architectural innovations—multi-input neural networks and cross-modal attention mechanisms—form the foundation of many state-of-the-art fusion models.[67]

Multi-input networks process each data modality through a dedicated subnetwork—for instance, convolutional neural networks for radiologic images, multilayer perceptron for omics data, and Transformer-based encoders for clinical text.[68] The extracted modality-specific embeddings are subsequently fused into a unified latent representation that facilitates joint inference. This modular approach supports flexible training and scalable deployment, especially in clinical environments where certain modalities may be missing or inconsistently collected.[69]

For instance, Zheng et al. [70] proposed a fusion radiomics model to predict osteosarcoma response to neoadjuvant chemotherapy. This model integrates MRI-derived radiomics features (T2WI and T1CE) with

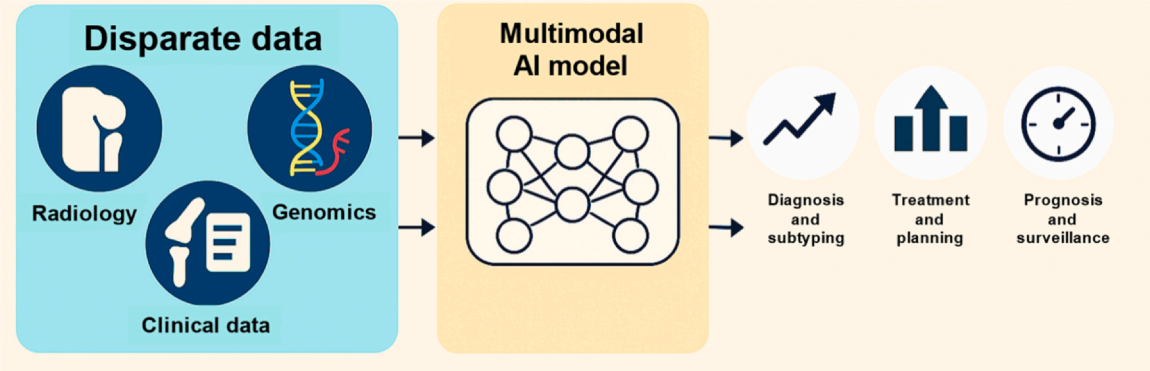


Figure 3. Overview of AI-driven multimodal data fusion for bone tumor modeling.

Table 2
Overview of classic multimodal fusion strategies in bone tumor AI applications.

Fusion strategy	Integration level	Representative architecture	Strengths	Limitations	Applicable tasks in bone tumor care
Early fusion	Input level (raw/low-level features)	Direct concatenation, joint input embedding	Simple to implement; captures early co-dependencies	Sensitive to scale mismatch; poor semantic abstraction	Structured tabular + omics integration in survival prediction
Intermediate fusion	Feature level (after modality-specific encoding)	Parallel CNN/MLP encoders + attention or tensor fusion	Captures modality-specific semantics; supports flexible integration	Computationally complex; requiring modality-specific design	Subtype classification, radio pathological grading, treatment response prediction
Late fusion	Decision level (prediction output)	Model ensembles, stacking, weighted voting	Modular; robust to missing modalities	Lacks cross-modal interaction; no joint reasoning	Cross-site deployment where full modality access is inconsistent

Abbreviations: 1. AI: artificial intelligence 2. CNN: support vector machine 3. MLP: multilayer perceptron.

clinical variables. In a separate study, Fu et al.[35] introduced a deep model with Siamese network (DS-Net) was proposed to address the limitations of small-sample learning in histological classification of osteosarcoma. DS-Net integrates an auxiliary supervision network with a classification branch to extract robust features from paired hematoxylin and eosin staining histology images.

While feature concatenation remains a common fusion technique, it often fails to capture the intricate dependencies across modalities. Cross-modal attention mechanisms mitigate this limitation by dynamically reweighting each modality's contribution according to data quality and task-specific relevance. Transformer-based attention—via Query–Key–Value operations—has become a leading strategy for enabling fine-grained cross-modal interaction.[71] In BTs research, attention-enabled frameworks such as Pathomic Fusion have demonstrated strong performance in malignancy classification and survival prediction by adaptively weighting inputs from histopathology, genomic data, and imaging modalities. Extensions like the weak-label attention mechanism proposed by Lu et al.[9] further enhance model robustness by sustaining performance despite noisy or incomplete labels—a frequent limitation in rare cancer datasets.

3.3. Shared representations and generative alignment

Beyond architectural-level fusion, effective multimodal learning also depends on aligning heterogeneous inputs within a shared semantic space. This alignment enables structurally and semantically distinct modalities, such as CT images and gene expression profiles, to be jointly interpreted in a biologically coherent manner. Several methodological strategies have been proposed to achieve this goal.

A widely adopted approach is joint embedding learning, where each modality is mapped to a latent feature space optimized for semantic alignment. This process is typically guided by loss functions such as cosine similarity or maximum mean discrepancy. For instance, Zheng et al.[72] introduced a heterogeneous graph-based framework that integrates histopathological and transcriptomic features using attention mechanisms to classify OS subtypes. This approach exemplifies how structural and biological relationships across modalities can be captured within a unified embedding space.

Contrastive learning offers another powerful strategy, particularly valuable in data-scarce or domain-shifted settings. By encouraging the alignment of semantically matched modality pairs (e.g., MRI and RNA-Seq from the same patient) and the separation of unmatched pairs, contrastive learning facilitates the acquisition of discriminative and generalizable representations.[73] Hsieh et al.[74] proposed a Twin Swin Transformer architecture, which achieved efficient OS cell segmentation by leveraging multi-scale feature fusion.

In scenarios involving incomplete data, generative modeling techniques become relevant in particular. Techniques such as multimodal variational autoencoders,[75] conditional generative adversarial networks,[76] and diffusion-based models can be employed to impute missing modalities or generate surrogate representations. For instance, Nguyen et al.[77] introduced FDDM, a novel framework bridging patch-level classification with region-based segmentation, which set a new benchmark in OS assessment. This framework highlights the potential of foundation models and diffusion-based refinements in complex imaging tasks.

Collectively, these strategies extend multimodal AI beyond passive data fusion, enabling active representation alignment and joint reasoning under noisy, sparse, or incomplete input conditions—critical capabilities for real-world application in bone oncology.

3.4. Fusion under incomplete and missing modalities

In real-world clinical environments, multimodal data collection is often hindered by logistical, economic, and institutional constraints. Consequently, radiologic images, pathology slides, omics data, and

clinical records are often partially or entirely unavailable, particularly in retrospective cohorts or low-resource settings.[78] To ensure clinical viability, AI systems for BTs care must demonstrate robust performance even when one or more modalities are absent.

3.4.1. Patterns of missingness and modeling implications

Missing modalities in clinical datasets generally follow two distinct patterns. Random missingness occurs sporadically, such as when patients cannot undergo MRI due to contraindications or scheduling issues.[79] In contrast, structured missingness reflects systemic deficiencies—for example, the routine unavailability of omics data in low-resource hospitals or missing follow-up documentation in archival cohorts. Such incomplete data scenarios can substantially degrade model performance, especially in rare tumors like OS or Ewing sarcoma, where datasets are often small and heterogeneous. [78]

Naïve imputation methods, such as zero-filling or mean substitution, often distort statistical properties and exacerbate bias.[80] To address this, recent approaches focus on learning modality-invariant representations or leveraging available modalities to infer missing data, while preserving semantic consistency across sources.[60]

3.4.2. Modality reconstruction and modality-adaptive fusion

To enhance robustness under incomplete input scenarios, recent MF models increasingly incorporate architectures explicitly designed to handle missing modalities. One such technique is modality dropout, which involves randomly omitting entire modalities during training to encourage learning of complementary and modality-agnostic representations.

Wu et al.[81] proposed DECIDE, an architecture comprising three specialized modules designed to achieve accurate OS segmentation by integrating complementary information from multi-sequence MRI. Zeng et al.[82] introduced an attention-based fusion and feature-sharing module that aggregates complementary features from available modalities to generate semantically aligned shared representations. These shared features are subsequently used to approximate missing modalities, mitigating completion bias and enhancing BTs classification accuracy.

3.4.3. Self-supervised learning for robustness

Beyond architectural innovations, self-supervised learning (SSL) presents a powerful training paradigm to improve model generalization in scenarios with sparse annotations and incomplete inputs. SSL utilizes label-free pretext tasks, such as masked reconstruction, intra-modal prediction, and cross-modal alignment, to learn semantically rich representations without full supervision.

In bone oncology, where annotated multimodal datasets are scarce and expensive to obtain, SSL offers a scalable and annotation-efficient alternative. For example, Wang et al.[83] proposed a DL segmentation model—MSAPN—that integrates multi-scale attention with pixel-level reconstruction. They demonstrated that radiomic features extracted from MSAPN-segmented lesions could effectively support the classification of musculoskeletal tumors as benign or malignant.

3.5. Summary

MF serves as a foundational pillar in the development of clinically meaningful AI systems for BTs management.[84] In contrast to unimodal approaches, fusion models facilitate integrated reasoning across spatial, morphological, molecular, and clinical dimensions—effectively addressing the intrinsic complexity of BTs, which are frequently marked by histologic heterogeneity, anatomically challenging locations, and aggressive metastatic potential.

This section has outlined the evolution from classical paradigms—namely early, intermediate, and late fusion—to contemporary architectures that leverage multi-input neural networks, cross-modal attention mechanisms, shared embedding spaces, and generative or contrastive learning techniques. These strategies are not mutually exclusive and may be flexibly integrated depending on task-specific requirements, modality availability, and institutional implementation contexts.

Despite notable progress, several critical challenges remain: (1) Missing modalities—both random and systematic—continue to impair model completeness and generalizability; (2) Cross-modal misalignment and feature redundancy introduce noise and hinder interpretability; (3) The lack of prospective, multicenter validation impedes real-world deployment; and (4) Limited transparency undermines clinician trust and integration into decision-making workflows.[85,86]

Overcoming these barriers will necessitate sustained innovation across multiple fronts: (1) Developing topology-aware architectures (e.g., graph neural networks, diffusion models, transformers) that better capture spatial and temporal tumor dynamics; (2) Employing SSL and few-shot learning approaches to alleviate annotation scarcity; (3) Advancing causality-aware and interpretable fusion strategies that align with biomedical knowledge; and (4) Designing edge–cloud–device coordination frameworks to enable scalable and context-aware clinical deployment.[87–89]

Ultimately, the most impactful AI systems will achieve a balance between predictive accuracy, computational efficiency, and clinical usability, offering robust, transparent, and actionable solutions tailored to the multifaceted demands of bone oncology.

4. Clinical applications of multimodal AI in bone tumor decision support

The clinical management of BTs spans a complex continuum, encompassing early detection, histopathological subtyping, individualized treatment planning, and long-term surveillance.[90] This process is challenged by substantial inter-patient variability in tumor morphology, molecular profiles, and therapeutic responsiveness.[91] Traditional workflows, often reliant on fragmented data sources and subjective clinical interpretation, frequently fail to deliver consistent and personalized care amidst such biological and clinical heterogeneity.[92]

Multimodal AI systems present a transformative alternative by integrating radiological imaging, digital pathology, omics data, and clinical records into unified, biologically coherent predictive frameworks.[93] These systems enable characterization of tumor features across spatial and biological scales, thereby supporting precise clinical decision-making that incorporates both anatomical context and molecular dynamics.[94] By augmenting and standardizing expert judgment, AI-powered multimodal models hold the potential to enhance diagnostic accuracy, enable risk stratification, optimize treatment planning, and predict disease progression (Figure 4).

This section reviews the application of advanced MF approaches across key stages of BTs management, including diagnosis, subtype classification, treatment planning, and recurrence prediction. Particular attention is given to the design of task-specific AI models, the challenges of interpretability and generalization, and the feasibility of clinical integration.

We categorize clinical applications into three core domains: diagnosis and subtyping, treatment planning, and recurrence prediction—each posing distinct requirements for data integration, interpretability, and clinical utility.

4.1. Diagnosis and subtyping with AI-based multimodal fusion

Accurate diagnosis and subtype classification of BTs remain challenging due to overlapping morphological and histological characteristics among benign, intermediate, and malignant lesions. A well-known example is the diagnostic ambiguity in differentiating enchondroma from atypical cartilaginous tumor or low-grade chondrosarcoma. This task is often associated with poor inter-rater agreement, even among experienced musculoskeletal radiologists and pathologists.

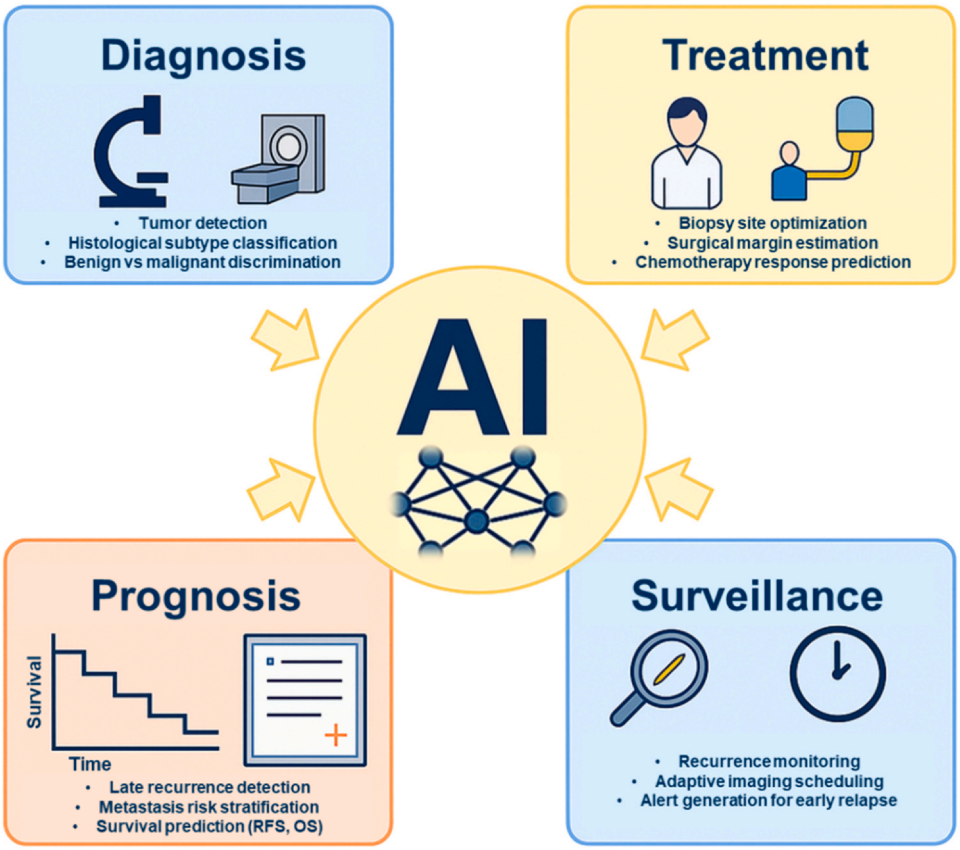


Figure 4. Clinical applications of multimodal AI in bone tumor decision support. Abbreviations: 1. AI: artificial intelligence 2. RFS: recurrence-free survival 3. OS: overall survival.

Multimodal AI models address this diagnostic ambiguity by integrating complementary data streams, including radiomic features from CT or MRI, morphometric descriptors from digital pathology, and structured clinical variables such as patient age, symptom duration, and relevant biomarkers. Such models often employ intermediate fusion architectures, wherein each modality is encoded through a dedicated subnetwork and subsequently integrated using attention-based or tensor-based aggregation mechanisms. This design preserves modality-specific information while enabling joint representation learning, thereby improving discriminative performance in challenging diagnostic tasks.

For example, Gassert et al.[95] developed a DL model based on CT-derived radiomic features to classify cartilaginous tumors. The model outperformed senior radiologists in differentiating atypical cartilaginous tumor from borderline or low-grade lesions. Similarly, several studies have combined MRI-derived features with clinical variables to predict tumor aggressiveness and metastatic potential, demonstrating improved subtype stratification and informing more targeted biopsy and surveillance strategies.[96]

Despite promising performance in retrospective analyses, most existing models are trained on limited, single-center datasets and lack robust external validation. To enable routine clinical deployment, these systems must undergo prospective, multicenter validation and be recalibrated across diverse patient populations, imaging protocols, and scanner types. In addition, they must adhere to evolving regulatory standards, including the U.S. FDA's Software as a Medical Device framework[97] and the European Union AI Act[98], both of which mandate stringent requirements for transparency, accountability, and human oversight.

4.2. Personalized treatment planning

The therapeutic management of BTs typically involves a combination of surgical resection, chemotherapy, radiotherapy, and, increasingly, molecularly targeted or immunotherapeutic strategies.[99] However, treatment responses vary considerably depending on tumor subtype, anatomical location, and individual patient characteristics, including age, immune function, and comorbidities. Integrating these factors into a personalized treatment plan remains a substantial challenge, especially for rare and histologically heterogeneous malignancies such as OS and Ewing sarcoma.[100]

Multimodal AI systems present a promising solution by synthesizing data from multiple modalities, such as preoperative imaging, molecular profiling, and clinical trajectories, to guide therapeutic decision-making. For example, Yin et al.[101] introduced an AI-enhanced 3D reconstruction pipeline that fused contrast-enhanced CT and MRI data to delineate tumor margins in pelvic OS. This approach enhanced the visualization of surgical margins and supported optimized preoperative planning.

In image-guided interventions, AI-enabled technologies have demonstrated utility in optimizing biopsy procedures and specimen-handling workflows. Witkowska et al.[102] conducted a retrospective analysis of 39 patients undergoing 40 image-guided bone biopsies using a robotic-assisted platform with steerable needle pathways. The system achieved a high technical success rate, diagnostic adequacy, and a favorable safety profile—even in anatomically complex regions. These tools exemplify how AI integration can bridge radiologic imaging, procedural precision, and pathological validation—critical components for the personalization of BTs therapy.

Nonetheless, the real-world deployment of such AI systems is hindered by multiple barriers, including variability in imaging protocols and scanner settings, lack of data standardization, and poor interoperability with clinical information systems.[103] To enable scalable adoption, future solutions must prioritize interoperability, regulatory compliance, and seamless integration into established surgical and diagnostic workflows, especially in resource-constrained or infrastructure-fragmented clinical environments.

Table 3
Representative AI applications across multimodal bone tumor datasets.

Authors	Clinical task	AI techniques	Modalities
Liu et al.[13]	Diagnosis and subtyping	DL + ML	MRI + Clinical
Noguchi et al.[15]	Diagnosis and subtyping	DL	CT
von. et al.[16]	Diagnosis and subtyping	Multitask DL	Radiographs
Zhang et al.[17]	Recurrence prediction	Multitask DL	Radiomics + Path omics
Hsieh et al.[74]	Diagnosis and subtyping	Modality dropout	X-ray, CT, MRI, Clinical
Liu et al.[83]	Diagnosis and subtyping	M3AE	Multimodal MRI
Gassert et al.[95]	Diagnosis and subtyping	CNN	CT
Yin et al.[101]	Recurrence prediction	ResUNet +3D ResNet	T2WI + CE-T1 MRI
Chen et al.[111]	Recurrence prediction	CNN + Pathomic Fusion	WSI, Genomics
Lu et al.[112]	Diagnosis and subtyping	Cross-modal attention	WSI Pathology
Zhang et al.[113]	Recurrence prediction	Graph learning	WSI, RNA-Seq
Yin et al.[114]	Diagnosis and subtyping	Radiomics + Modality dropout	Radiographs, Clinical
Wang et al.[115]	Treatment planning	CNN	CT
Mahootiha et al.[116]	Recurrence prediction	Multitask DL	MRI, Clinical, Genomics
Wen et al.[117]	Treatment planning	Radiomics-based model	CT, Clinicopathology

Abbreviations: 1. AI: artificial intelligence 2. CT: computed tomography 3. MRI: magnetic resonance imaging 4. DL: deep learning 5. ML: machine learning 6. CNN: convolutional neural network 7.WSI: whole-slide imaging 8. RNA-Seq: RNA sequencing 9. T2WI: T2-weighted imaging 10. CE-T1 MRI: contrast-enhanced T1-weighted magnetic resonance imaging 11. M3AE: multimodal masked autoencoder.

4.3. Prognosis and recurrence prediction

Accurate prediction of recurrence and long-term survival is critical for informing adjuvant therapy decisions, tailoring follow-up intervals, and supporting patient-centered communication.[104] However, conventional prognostic systems—relying on histologic grade, tumor size and margin status—are often inadequate for capturing the biological complexity of BTs. This limitation is especially pronounced in high-grade OS and Ewing sarcoma, where relapse patterns are unpredictable and metastasis to the lungs or bones can emerge months or even years after initial remission.[105]

Multimodal AI frameworks offer an integrated alternative by combining post-treatment imaging, longitudinal clinical records, histopathologic data, and molecular profiles to generate patient-specific predictions of recurrence, metastasis, and survival.[106] Such models are particularly valuable in BTs, where the interplay between phenotype and genotype is intricate and inadequately addressed by unimodal predictors.[107] By harnessing temporal continuity and cross-modal synergy, AI systems can uncover latent risk signatures that remain undetected by clinicians or conventional statistical methods.

For instance, He et al.[108] developed a composite scoring system to predict response to neoadjuvant chemotherapy in high-grade primary bone sarcomas, incorporating multi-timepoint clinical parameters (e.g., pain scale, lab values), multimodal imaging (PET-CT, contrast-enhanced MRI), and histopathological markers. The resulting model outperformed conventional evaluation metrics and enabled individualized surgical readiness assessments, underscoring the importance of temporally informed and modality-aware fusion.

At a broader scale, Jiang et al.[109] proposed Meta Net, a multimodal DL framework that integrates genomic, transcriptomic, and clinical data from over 32,000 cancer cases to stratify metastatic potential. Although initially developed for prostate cancer, its ability to differentiate bone-seeking from liver-seeking metastatic phenotypes illustrates its potential to inform organ-specific surveillance in BTs—where metastatic patterns remain poorly understood and underrepresented in current grading systems.

These developments demonstrate that AI-enhanced prognostic models can go beyond static baseline risk scores to support adaptive, risk-guided follow-up and post-treatment care. Future work should incorporate time-to-event modeling, dynamic risk recalibration, and clinically interpretable metrics, such as recurrence-free survival, concordance index, and progression hazard ratios, rather than relying solely on cross-sectional accuracy or area under the curve (AUC). Furthermore, rigorous external validation across multi-institutional and diverse patient cohorts, especially in rare BTs subtypes—is essential to ensure model robustness, reliability, and translational applicability.

4.4. Summary

Multimodal AI systems are increasingly demonstrating clinical utility across the entire continuum of BTs care, including diagnosis, histologic subtyping, treatment planning, and recurrence surveillance. By integrating radiologic, pathologic, molecular, and clinical data, these systems emulate the multidisciplinary reasoning foundational to bone oncology and facilitate scalable, patient-specific decision support. Their capacity to capture macroscopic, microscopic, and temporal patterns positions them as compelling tools for bridging biological complexity with practical clinical decision-making.

Although metrics such as AUC and classification accuracy are frequently reported, their translation into clinical utility, such as influence on treatment decisions, surgical planning, or patient outcomes, remains largely unassessed. This evidentiary gap underscores the need to transition from performance-centered benchmarking toward clinically meaningful, outcome-driven evaluation.[110] Table 3 outlines representative AI models developed for BTs, detailing data modalities used, clinical endpoints targeted, and validation status.

5. Challenges in multimodal AI for bone tumor care

Despite significant advances in algorithmic design and promising proof-of-concept studies, the clinical translation of AI-based MF systems for BTs remains limited.[118] These limitations arise from a confluence of factors, including the rarity and biological complexity of BTs, technical challenges in model development, and broader concerns related to data quality, system integration, and ethical oversight.[119]

This section provides a structured overview of the primary challenges across four dimensions: (1) disease-specific modeling constraints, (2) data fragmentation and workflow gaps, (3) deployment barriers and clinical–technical disconnects, and (4) ethical and societal risks.

5.1. Disease-specific modeling challenges

BTs such as OS, Ewing sarcoma, and chondrosarcoma exhibit substantial biological and anatomical heterogeneity, manifesting across imaging, histopathologic, and molecular dimensions.[120,121] These tumors frequently demonstrate irregular growth, multi-tissue infiltration, and ill-defined margins, especially in anatomically complex regions such as the pelvis and spine, posing significant challenges for accurate delineation and segmentation.[122] Such variability complicates the development of robust spatial models and undermines the reproducibility of AI predictions across heterogeneous anatomical regions.

From a pathological perspective, biopsy specimens often sample only partial or non-representative regions of the tumor. When coupled with inter-observer grading variability, this causes substantial annotation noise.[123] Molecular data are frequently limited by poor annotation quality, and tumor subtypes may be inconsistently defined across cohorts. Collectively, these issues undermine the reliability of SSL approaches and hinder the transferability of models across institutions.

Addressing these challenges necessitates the development of modeling frameworks that are semantically aligned across modalities, robust to noisy or incomplete labels, and capable of learning from anatomically complex or partially annotated datasets.[124] Emerging techniques such as contrastive representation learning, spatially aware neural architectures, and semi-supervised training paradigms may prove critical in overcoming these disease-specific obstacles.[125]

5.2. Data fragmentation and workflow gaps

The rarity of BTs directly contributes to the limited availability of high-quality, multimodal datasets. Public repositories such as The Cancer Genomic Atlas-OS (https://portal.gdc.cancer.gov/)and The Cancer Imaging Archive (https://www.cancerimagingarchive.net/) are valuable but often constrained by small cohort sizes, missing modalities, and inconsistent annotation standards. Moreover, most well-curated datasets are derived from academic medical centers, while patients from low-resource or rural settings are significantly underrepresented.[126] This data imbalance raises critical concerns regarding the generalizability, fairness, and equity of AI systems when applied to diverse clinical populations.[127]

Compounding this challenge is the longitudinal nature of BTs management, which encompasses diagnosis, treatment, surveillance, and recurrence monitoring. Most current AI models rely on single time-point data, failing to capture the temporal dynamics of disease progression or therapeutic response.[128] Consequently, these models are restricted to isolated clinical tasks and fall short in supporting longitudinal continuity of care.

To address these gaps, future research should prioritize the development of time-resolved, multimodal datasets that incorporate standardized labeling protocols and demographically diverse patient populations. In parallel, the adoption of federated learning and secure data-sharing protocols can facilitate multi-institutional collaboration while preserving privacy. Additionally, longitudinal modeling techniques, such as spatiotemporal graph neural networks and transformer-based architectures, will be essential for capturing temporal dependencies and enabling dynamic decision-making throughout the patient care continuum.[129]

5.3. Deployment barriers and clinical–technical disconnect

Despite promising technical performance, many AI models fail to achieve clinical deployment due to a persistent disconnect between algorithmic outputs and clinician needs. Model outputs are often limited to probability scores or risk stratifications which, though statistically informative, lack operational specificity and clinical applicability. Clinicians require models that provide actionable insights, such as individualized resection margins, optimal implant trajectories, or tailored surveillance intervals, which can be directly interpreted and applied in clinical decision-making.[130]

This disconnect is exacerbated by the limited involvement of clinicians in the early stages of model development. Many AI systems are developed in technical silos, optimized for metrics such as AUC or Dice scores, which may not align with real-world clinical priorities, such as limb preservation, postoperative function, or patient-reported outcomes. Consequently, even technically robust models may offer limited practical utility within real-world clinical workflows.

Bridging this divide requires co-design methodologies, wherein clinicians and AI developers collaborate from the outset to define clinically meaningful endpoints, interpretability constraints, and validation standards.[131] In addition, incorporating multi-objective optimization strategies can help balance statistical accuracy with practical clinical relevance, yielding models that are not only performant but also interpretable, trustworthy, and responsive to complex care environments.

5.4. Ethical and social risks in bone oncology

The deployment of multimodal AI systems in bone oncology raises critical ethical and social concerns, especially considering the sensitive and longitudinal nature of the data involved. Adolescents, who constitute a large proportion of primary BTs patients, often contribute longitudinal, multimodal datasets spanning radiology, pathology, genomics, and clinical narratives. Without stringent safeguards, such comprehensive records elevate the risk of reidentification, privacy violations, and unauthorized use, particularly when data are shared inter-institutionally or repurposed beyond their original scope.[132] Furthermore, population-level AI models may overlook individual patient preferences, such as values related to limb preservation, cosmetic outcomes, or quality of life, resulting in recommendations that may conflict with personalized care goals.[133] The opaque, black-box nature of many models compounds these risks by limiting patient understanding and undermining trust, autonomy, and informed consent.

To mitigate these challenges, AI systems should be designed with intrinsic safeguards for data privacy, mechanisms for aligning outputs with patient preferences, and transparent, interpretable architectures. Techniques such as differential privacy and data traceability can help reduce the likelihood of unauthorized exposure or misuse. Simultaneously, embedding value-sensitive design principles ensures that models reflect not only clinical efficacy but also patient-centered ethical and personal values. Finally, transparent model architectures, augmented by explainable AI techniques and clinician-oriented interpretability tools, are essential for supporting shared decision-making, enhancing user confidence, and ensuring accountability in AI-assisted care.

5.5. Summary

The integration of multimodal AI into clinical BTs care continues to face persistent challenges, spanning scientific, technical, systemic, and ethical domains. While advances in fusion algorithms and predictive modeling continue, successful real-world implementation requires a holistic strategy, one that incorporates clinical domain knowledge, respects existing workflows, and upholds patient-centered safeguards.

Addressing these multifaceted challenges is critical to realizing the transformative potential of multimodal AI, which is not merely as a computational tool, but as a trustworthy, integrative partner in the future of precision bone oncology.

6. Discussion

DL–based MF has emerged as a powerful paradigm for managing BTs, allowing the integration of diverse data types, including radiologic images, pathology slides, omics profiles, and clinical records, into unified predictive models. Despite rapid algorithmic advances, the clinical translation of these models remains constrained by key limitations in data quality, validation rigor, and deployment feasibility.

Most published models are trained on retrospective, single-institution datasets with modest sample sizes and incomplete modality coverage. External validation is rarely performed, and models that appear promising in internal testing often fail to generalize across institutions, imaging vendors, or population demographics. Moreover, some unimodal approaches, such as those using only WSI or genomics, achieve competitive results in tightly controlled settings, questioning whether fusion necessarily confers superior performance in all scenarios. These

concerns underscore the need for prospective, multicenter studies, task-specific evaluation benchmarks, and open validation platforms.

From a technical standpoint, advances such as modality dropout, cross-modal attention, and generative alignment have expanded model robustness to missing inputs and heterogeneous data sources. However, these techniques often require substantial engineering, and their clinical scalability remains uncertain. Furthermore, contrastive, and SSL approaches, though powerful, demand careful calibration of feature granularity, augmentation logic, and negative sampling schemes, which may not align with real-world data acquisition patterns.

Looking ahead, the deployment of AI-assisted decision systems in bone oncology will require alignment with regulatory, infrastructural, and ethical frameworks. The U.S. FDA's Software as a Medical Device guidance, the European Union AI Act, and local institutional review boards pose nontrivial approval hurdles. Edge–cloud collaboration, essential for deploying AI in resource-constrained environments, faces technical barriers including bandwidth limits, data privacy regulations, and workflow interoperability with PACS/RIS systems. These factors must be incorporated into model development pipelines through design-for-deployment principles.

Emerging paradigms may help bridge these gaps. For example, multimodal foundation models such as BioMedCLIP and PathChat have shown the potential of pretraining across massive, cross-domain datasets, enabling few-shot adaptation to rare tumor types. Medical large language models and agentic systems could serve as central orchestrators, interpreting textual radiology/pathology reports, integrating patient history, and generating contextualized recommendations in an explainable, dialogue-based manner. These architectures may offer more human-AI collaboration than model-centric prediction alone

Finally, future directions must explicitly map to the challenges identified in this review. To address data fragmentation, we advocate for federated learning, harmonized annotation standards, and open-source benchmarks. For model interpretability and equity, we emphasize explainable AI, uncertainty quantification, and audits of performance stratified by patient subgroups. For ethical and legal governance, principles such as value-sensitive design, differential privacy, and cross-border compliance auditing should guide clinical-grade AI deployment.

In summary, while multimodal AI offers a promising framework for personalized BTs care, its success depends not only on algorithmic novelty but also on reproducibility, integration, and responsible innovation across the clinical ecosystem.

7. Conclusions

Multimodal AI holds strong promise to reshape BTs care by enabling more precise, data-driven clinical decision-making. Through the integration of imaging, pathology, omics, and clinical records, these systems can assist across key tasks, including diagnosis, treatment planning, and outcome prediction, offering a holistic view of tumor biology and patient context.

Yet realizing this promise requires moving beyond technical development alone. For AI to be clinically impactful, it must align with the complexity of real-world workflows, accommodate diverse and often incomplete data, and comply with rigorous standards of safety, interpretability, and accountability. Success will depend on bridging innovation with implementation, ensuring that models are not only powerful, but usable, trusted, and grounded in biomedical relevance.

Looking ahead, efforts should prioritize: improving generalizability through multicenter validation; embedding AI into routine clinical pathways; and fostering systems that are not only predictive, but explainable and ethically deployable. When guided by interdisciplinary collaboration and aligned with patient and clinician needs, multimodal AI is well-positioned to evolve from research concept into a clinically credible, practice-changing tool for bone oncology.

CRediT authorship contribution statement

Hong Zhou: Writing – review & editing, Validation. **Zineng Yan:** Investigation, Conceptualization. **Songxiang Liu:** Writing – review & editing, Methodology. **Lin Lu:** Investigation, Formal analysis. **Tongtong Huo:** Writing – original draft, Validation, Methodology, Conceptualization. **Jiaming Yang:** Writing – review & editing, Formal analysis. **Wei Wu:** Writing – original draft, Validation, Supervision. **Jin Liu:** Writing – original draft, Validation, Methodology. **Xiaoliang Chen:** Writing – review & editing, Methodology, Formal analysis. **Zhewei Ye:** Writing – review & editing, Funding acquisition. **Mingdi Xue:** Writing – review & editing, Project administration. **Pengran Liu:** Writing – review & editing, Formal analysis. **Jiayao Zhang:** Writing – review & editing, Supervision. **Yi Xie:** Investigation, Formal analysis. **Honglin Wang:** Methodology, Conceptualization.

Ethical statement

The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are Authors' contributions appropriately investigated and resolved.

Funding

This work was supported by the National Natural Science Foundation of China (82172524) and the Natural Science Foundation of Hubei Province (2025AFB240).

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was supported by the National Natural Science Foundation of China [Grant No.: 82172524], and the Natural Science Foundation of Hubei Province [Grant No.: 2025AFB240].

References

1. Ward E, DeSantis C, Robbins A, Kohler B, Jemal A. Childhood and adolescent cancer statistics, 2014. *CA Cancer J Clin.* 2014;64(2):83–103.
2. Jiang Y, Wang J, Sun M, et al. Multi-omics analysis identifies osteosarcoma subtypes with distinct prognosis indicating stratified treatment. *Nat Commun.* 2022;13(1):7207.
3. Bădilă AE, Rădulescu DM, Niculescu AG, Grumezescu AM, Rădulescu M, Rădulescu AR. Recent advances in the treatment of bone metastases and primary bone tumors: an up-to-date review. *Cancers (Basel).* 2021;13(16):4229.
4. Jiao W, Atwal G, Polak P, et al. A deep learning system accurately classifies primary and metastatic cancers using passenger mutation patterns. *Nat Commun.* 2020;11(1):728.
5. Anderson WJ, Doyle LA. Updates from the 2020 World Health Organization classification of soft tissue and bone tumours. *Histopathology.* 2021;78(5):644–657.
6. Do BH, Langlotz C, Beaulieu CF. Bone tumor diagnosis using a naïve bayesian model of demographic and radiographic features. *J Digit Imaging.* 2017;30(5):640–647.
7. Jacobson D, Cadieux B, Higano CS, et al. Risk factors associated with skeletal-related events following discontinuation of denosumab treatment among patients with bone metastases from solid tumors: a real-world machine learning approach. *J Bone Oncol.* 2022;34:100423.
8. Rashidi A, Baratto L, Theruvath AJ, et al. Improved Detection of Bone Metastases in Children and Young Adults with Ferumoxytol-enhanced MRI. *Radiol Imaging Cancer.* 2023;5(2):220080.
9. Lu MY, Williamson DFK, Chen TY, Chen RJ, Barbieri M, Mahmood F. Data-efficient and weakly supervised computational pathology on whole-slide images. *Nat Biomed Eng.* 2021;5(6):555–570.
10. Zeng J, Banerjee I, Henry AS, et al. Natural language processing to identify cancer treatments with electronic medical records. *JCO Clin Cancer Inf.* 2021(5):379–393.
11. Cook GJ, Goh V. Functional and Hybrid Imaging of Bone Metastases. *J Bone Miner Res.* 2018;33(6):961–972.

12. LeCun Y, Bengio Y, Hinton G. Deep learning. *Nature.* 2015;521(7553):436–444.
13. Liu R, Pan D, Xu Y, et al. A deep learning-machine learning fusion approach for the classification of benign, malignant, and intermediate bone tumors. *Eur Radiol.* 2022;32(2):1371–1383.
14. Vogrin M, Trojner T, Kelc R. Artificial intelligence in musculoskeletal oncological radiology. *Radiol Oncol.* 2020;55(1):1–6.
15. Noguchi S, Nishio M, Sakamoto R, et al. Deep learning-based algorithm improved radiologists' performance in bone metastases detection on CT. *Eur Radiol.* 2022;32(11):7976–7987.
16. von Schacky CE, Wilhelm NJ, Schäfer VS, et al. Multitask deep learning for segmentation and classification of primary bone tumors on radiographs. *Radiology.* 2021;301(2):398–406.
17. Zhang YF, Zhou C, Guo S, et al. Deep learning algorithm-based multimodal MRI radiomics and pathomics data improve prediction of bone metastases in primary prostate cancer. *J Cancer Res Clin Oncol.* 2024;150(2):78.
18. Coudray N, Ocampo PS, Sakellaropoulos T, et al. Classification and mutation prediction from non-small cell lung cancer histopathology images using deep learning. *Nat Med.* 2018;24(10):1559–1567.
19. 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.
20. Lo Bianco G, Lanza E, Provenzano S, et al. A Multimodal Clinical Approach for the Treatment of Bone Metastases in Solid Tumors. *Anesth Pain Med.* 2022;12(3):126333.
21. Moon JB, Yoo SW, Lee C, Kim DY, Pyo A, Kwon SY. Multimodal Imaging-Based Potential Visualization of the Tumor Microenvironment in Bone Metastasis. *Cells.* 2021;10(11):2877.
22. Masoudi S, Mehralivand S, Harmon SA, et al. Deep learning based staging of bone lesions from computed tomography scans. *IEEE Access.* 2021;9:87531–87542.
23. Hinterwimmer F, Guenther M, Consalvo S, et al. Impact of metadata in multimodal classification of bone tumours. *BMC Musculoskelet Disord.* 2024;25(1):822.
24. Kline A, Wang H, Li Y, et al. Multimodal machine learning in precision health: a scoping review. *NPJ Digit Med.* 2022;5(1):171.
25. Liang PP, Zadeh A, Morency LP. Foundations & trends in multimodal machine learning: Principles, challenges, and open questions. *ACM Comput Surv.* 2024;56(10):1–42.
26. Wu JS, Hochman MG. *Bone Tumors: A Practical Guide to Imaging.* Springer Science & Business Media; 2012.
27. Gaume M, Chevret S, Campagna R, Larousserie F, Biau D. The appropriate and sequential value of standard radiograph, computed tomography and magnetic resonance imaging to characterize a bone tumor. *Sci Rep.* 2022;12(1):6196.
28. He Y, Pan I, Bao B, et al. Deep learning-based classification of primary bone tumors on radiographs: a preliminary study. *EBioMedicine.* 2020;62:103121.
29. Zhang R, Huang L, Xia W, Zhang B, Qiu B, Gao X. Multiple supervised residual network for osteosarcoma segmentation in CT images. *Comput Med Imaging Graph.* 2018;63:1–8.
30. Leporq B, Bouhamama A, Pilleul F, et al. MRI-based radiomics to predict lipomatous soft tissue tumors malignancy: a pilot study. *Cancer Imaging.* 2020;20(1):78.
31. Lindgren Belal S, Sadik M, Kaboteh R, et al. Deep learning for segmentation of 49 selected bones in CT scans: first step in automated PET/CT-based 3D quantification of skeletal metastases. *Eur J Radiol.* 2019;113:89–95.
32. Faiella E, Santucci D, Calabrese A, et al. Artificial intelligence in bone metastases: an MRI and CT imaging review. *Int J Environ Res Public Health.* 2022;19(3):1880.
33. Chang CY, Buckless C, Yeh KJ, Torriani M. Automated detection and segmentation of sclerotic spinal lesions on body CTs using a deep convolutional neural network. *Skelet Radiol.* 2022;51(2):391–399.
34. Ren M, Yang H, Lai Q, et al. MRI-based radiomics analysis for predicting the EGFR mutation based on thoracic spinal metastases in lung adenocarcinoma patients. *Med Phys.* 2021;48(9):5142–5151.
35. Fu Y, Xue P, Ji H, Cui W, Dong E. Deep model with Siamese network for viable and necrotic tumor regions assessment in osteosarcoma. *Med Phys.* 2020;47(10):4895–4905.
36. Bao N, Zhang J, Li Z, et al. CT-less whole-body bone segmentation of PET images using a multimodal deep learning network. *IEEE J Biomed Health Inf.* 2025;29(2):1151–1164.
37. Xiao Z, Zhang Y, Deng Z, Liu F. Light3DHS: A lightweight 3D hippocampus segmentation method using multiscale convolution attention and vision transformer. *Neuroimage.* 2024;292:120608.
38. Wang X, Zhao J, Marostica E, et al. A pathology foundation model for cancer diagnosis and prognosis prediction. *Nature.* 2024;634(8035):970–978.
39. Bera K, Schalper KA, Rimm DL, Velcheti V, Madabhushi A. Artificial intelligence in digital pathology - new tools for diagnosis and precision oncology. *Nat Rev Clin Oncol.* 2019;16(11):703–715.
40. Bhalla S, Laganà A. Artificial intelligence for precision oncology. *Adv Exp Med Biol.* 2022;1361:249–268.
41. Yang Z, Wei T, Liang Y, et al. A foundation model for generalizable cancer diagnosis and survival prediction from histopathological images. *Nat Commun.* 2025;16(1):2366.
42. Schmidt A, Morales-Alvarez P, Molina R. Probabilistic attention based on Gaussian processes for deep multiple instance learning. *IEEE Trans Neural Netw Learn Syst.* 2024;35(8):10909–10922.
43. Behjati S, Tarpey PS, Presneau N, et al. Distinct H3F3A and H3F3B driver mutations define chondroblastoma and giant cell tumor of bone. *Nat Genet.* 2013;45(12):1479–1482.
44. Rajan S, Zaccaria S, Cannon MV, et al. Structurally complex osteosarcoma genomes exhibit limited heterogeneity within individual tumors and across evolutionary time. *Cancer Res Commun.* 2023;3(4):564–575.
45. Sayles LC, Breese MR, Koehne AL, et al. Genome-informed targeted therapy for osteosarcoma. *Cancer Discov.* 2019;9(1):46–63.
46. Thway K, Fisher C. Mesenchymal tumors with EWSR1 gene rearrangements. *Surg Pathol Clin.* 2019;12(1):165–190.
47. Cai Z, Poulos RC, Liu J, et al. Machine learning for multi-omics data integration in cancer. *iScience.* 2022;25(2):103798.
48. Jamal-Hanjani M, Wilson GA, McGranahan N, et al. Tracking the evolution of non-small-cell lung cancer. *N Engl J Med.* 2017;376(22):2109–2121.
49. Kickingereder P, Bonekamp D, Nowosielski M, et al. Radiogenomics of glioblastoma: machine learning-based classification of molecular characteristics by using multiparametric and multiregional MR imaging features. *Radiology.* 2016;281(3):907–918.
50. Zhang Y, Peng C, Wang Q, et al. Unified multi-modal image synthesis for missing modality imputation. *IEEE Trans Med Imaging.* 2025;44(1):4–18.
51. Wu R, Wang H, Chen HT, Carneiro G. Deep multimodal learning with missing modality: A survey. *arXiv preprint arXiv:240907825*. Published online 2024.
52. Wang Y, Wang L, Rastegar-Mojarad M, et al. Clinical information extraction applications: a literature review. *J Biomed Inf.* 2018;77:34–49.
53. Bielack SS, Kempf-Bielack B, Delling G, et al. Prognostic factors in high-grade osteosarcoma of theextremities or trunk: an analysis of 1,702 patients treatedon neoadjuvant cooperative osteosarcoma study GroupProtocols. *J Clin Oncol.* 2023;41(27):4323–4337.
54. Change Fang L, Chen Q, Wei CH, to Fang L, Chen Q, Wei CH. Bioformer: an efficient transformer language model for biomedical text mining. *arXiv preprint arXiv:2302.01588*. Published online 2023.
55. Landolsi MY, Hlaoua L, Romdhane LB. Extracting and structuring information from the electronic medical text: state of the art and trendy directions. *Multimed Tools Appl.* 2024;83(7):21229–21280.
56. Pellegrini C, Özsoy E, Bani-Harouni D. From EHRs to Patient Pathways: Scalable Modeling of Longitudinal Health Trajectories with LLMs. *arXiv preprint arXiv:2506.04831*. Published online 2025.
57. Baltrušaitis T, Ahuja C, Morency L-P. Multimodal machine learning: a survey and taxonomy. *IEEE Trans Pattern Anal Mach Intell.* 2018;41(2):423–443.
58. Ye Q, Yang H, Lin B, et al. Automatic detection, segmentation, and classification of primary bone tumors and bone infections using an ensemble multi-task deep learning framework on multi-parametric MRIs: a multi-center study. *Eur Radio.* 2024;34(7):4287–4299.
59. Yildiz Potter I, Yeritsyan D, Mahar S, et al. Automated bone tumor segmentation and classification as benign or malignant using computed tomographic imaging. *J Digit Imaging.* 2023;36(3):869–878.
60. Tsai YHH, Bai S, Liang PP, Kolter JZ, Morency LP, Salakhutdinov R. Multimodal transformer for unaligned multimodal language sequences. In: *Proceedings of the Conference. Association for Computational Linguistics. Meeting.* Vol 2019. ; 2019:6558.
61. Huang L, Denoeux T, Vera P, et al. Evidence fusion with contextual discounting for multi-modality medical image segmentation. *Int Conf Med Image Comput Comput-Assist Interv.* 2022:401–411.
62. Li Y, Wu FX, Ngom A. A review on machine learning principles for multi-view biological data integration. *Brief Bioinform.* 2018;19(2):325–340.
63. Ballard JL, Wang Z, Li W, Shen L, Long Q. Deep learning-based approaches for multi-omics data integration and analysis. *BioData Min.* 2024;17(1):38.
64. Zhuang Y, Liu H, Song E, Ma G, Xu X, Hung CC. APRNet: a 3D anisotropic pyramidal reversible network with multi-modal cross-dimension attention for brain tissue segmentation in MR images. *IEEE J Biomed Health Inf.* 2022;26(2):749–761.
65. Xu Y, Cai M, Lin L, et al. PA-ResSeg: a phase attention residual network for liver tumor segmentation from multiphase CT images. *Med Phys.* 2021;48(7):3752–3766.
66. Zhou T, Ruan S, Canu SJA. A review: a review: deep learning for medical image segmentation using multi-modality fusion. *Array.* 2019;3:100004.
67. Ramachandram D, Taylor GWJIspm. Deep multimodal learning: a survey on recent advances and trends. *IEEE Signal Process Mag.* 2017;34(6):96–108.
68. Wen T, Tong B, Fu Y, Li Y, Ling M, Chen X. A novel adjunctive diagnostic method for bone cancer: Osteosarcoma cell segmentation based on Twin Swin Transformer with multi-scale feature fusion. *J Bone Oncol.* 2024;49:100647.
69. Kang M, Ko E, Mersha TB. A roadmap for multi-omics data integration using deep learning. *Brief Bioinform.* 2022;23(1):bbab454.
70. Zheng F, Yin P, Liang K, et al. Fusion Radiomics-Based Prediction of Response to Neoadjuvant Chemotherapy for Osteosarcoma. *Acad Radiol.* 2024;31(6):2444–2455.
71. Pan L, Wang H, Wang L, et al. Noise-reducing attention cross fusion learning transformer for histological image classification of osteosarcoma. *Biomed Signal Process Control.* 2022;77:103824.
72. Zheng Y, Conrad RD, Green EJ, et al. Graph Attention-Based Fusion of Pathology Images and Gene Expression for Prediction of Cancer Survival. *IEEE Trans Med Imaging.* 2024;43(9):3085–3097.
73. Yuan X, Lin Z, Kuen J, et al. Multimodal contrastive training for visual representation learning. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition.*; 2021:6995-7004.
74. Hsieh ME, Chiou CY, Tsai HW, Chang YC, Chung PC. MSCS: Multi-Scale Consistency Supervision with CNN-Transformer Collaboration for Semi-Supervised Histopathology Image Semantic Segmentation. *IEEE Transactions on Artificial Intelligence.* Published online 2024.
75. Khattar D, Goud JS, Gupta M, Varma V. Mvae: Multimodal variational autoencoder for fake news detection. In: *The World Wide Web Conference*; 2019:2915-2921.
76. Sohn K, Lee H, Yan X. Learning structured output representation using deep conditional generative models. *Adv Neural Inf Process Syst.* 2015:28.
77. Nguyen MD, Nguyen DT, Nguyen TV, Yamada H, Pham HH, Le Nguyen P. Bridging Classification and Segmentation in Osteosarcoma Assessment via Foundation and Discrete Diffusion Models. In: *2025 IEEE 22nd International Symposium on Biomedical Imaging (ISBI)*. IEEE; 2025:1–5.

78. Liang PP, Lyu Y, Fan X, et al. Multibench: Multiscale benchmarks for multimodal representation learning. *Adv Neural Inf Process Syst.* 2021;2021(DB1):1.
79. Zhang C, Chu X, Ma L, et al. M3care: Learning with missing modalities in multimodal healthcare data. In: *Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining.* ; 2022:2418–2428.
80. Liew AW, Law NF, Yan H. Missing value imputation for gene expression data: computational techniques to recover missing data from available information. *Brief Bioinform.* 2011;12(5):498–513.
81. Wu Y, Li J, Wang X, et al. DECIDE: a decoupled semantic and boundary learning network for precise osteosarcoma segmentation by integrating multi-modality MRI. *Comput Biol Med.* 2024;174:108308.
82. Zeng Q, Yang Y. Incomplete Multimodal Bone Tumor Image Classification Based on Attention Fusion and Feature Sharing. In: *Proceedings of the 2024 5th International Conference on Intelligent Medicine and Health.*; 2024:32-39.
83. Wang S, Sun M, Sun J, et al. Advancing musculoskeletal tumor diagnosis: Automated segmentation and predictive classification using deep learning and radiomics. *Comput Biol Med.* 2024;175:108502.
84. Lu G, Tian R, Yang W, et al. Deep learning radiomics based on multimodal imaging for distinguishing benign and malignant breast tumours. *Front Med (Lausanne).* 2024;11:1402967.
85. Chen Z, Du Y, Hu J, et al. Mapping medical image-text to a joint space via masked modeling. *Med Image Anal.* 2024;91:103018.
86. Hou W, Lin C, Yu L, et al. Hybrid graph convolutional network with online masked autoencoder for robust multimodal cancer survival prediction. *IEEE Trans Med Imaging.* 2023;42(8):2462–2473.
87. Fu X, Patrick E, Yang JYH, et al. Deep multimodal graph-based network for survival prediction from highly multiplexed images and patient variables. *Comput Biol Med.* 2023;154:106576.
88. Kipf TN, Welling M. Semi-supervised classification with graph convolutional networks. *arXiv preprint arXiv:160902907*. Published online 2016.
89. Ho J, Jain A, Abbeel P. Denoising diffusion probabilistic models. *Adv Neural Inf Process Syst.* 2020;33:6840–6851.
90. Morii T, Asano N, Nakayama R, et al. Complications of surgery for giant cell tumor of bone in the extremities: Incidence, risk factors, management modality, and impact on functional and oncological outcomes. *J Orthop Sci.* 2022;27(3):681–688.
91. Gunanithi S, Ilavarasan S, Karthika RN. Revolutionizing Osteosarcoma Diagnosis: A Comparative Analysis of Deep Learning Models for Precise Bone Cancer Detection Using Multi-Modal Medical Imaging. In: *International Conference Summit on Artificial Intelligence*. Springer; 2024:197–207.
92. Papalia GF, Brigato P, Sisca L, et al. Artificial Intelligence in Detection, Management, and Prognosis of Bone Metastasis: A Systematic Review. *Cancers (Basel).* 2024;16(15):2700.
93. Li MD, Ahmed SR, Choy E, et al. Artificial intelligence applied to musculoskeletal oncology: a systematic review. *Skelet Radiol.* 2022;51(2):245–256.
94. Yang F, Yan D, Wang Z. Large-Scale assessment of ChatGPT's performance in benign and malignant bone tumors imaging report diagnosis and its potential for clinical applications. *J Bone Oncol.* 2024;44:100525.
95. Gassert FG, Lang D, Hesse N, et al. A deep learning model for classification of chondroid tumors on CT images. *BMC Cancer.* 2025;25(1):561.
96. Yin P, Mao N, Chen H, et al. Machine and deep learning based radiomics models for preoperative prediction of benign and malignant sacral tumors. *Front Oncol.* 2020;10:564725.
97. Benjamens S, Dhunnoo P, Meskó B. The state of artificial intelligence-based FDA-approved medical devices and algorithms: an online database. *NPJ Digit Med.* 2020;3:118.
98. Aboy M, Minssen T, Vayena E. Navigating the EU AI Act: implications for regulated digital medical products. *NPJ Digit Med.* 2024;7(1):237.
99. Rai V, Munazzam SW, Wazir NU, et al. Revolutionizing bone tumor management: cutting-edge breakthroughs in limb-saving treatments. *Eur J Orthop Surg Trauma.* 2024;34(4):1741–1748.
100. Misir A, Yuce A AI. in Orthopedic Research: A Comprehensive Review. *J Orthop Res.* 2025;43(8):1508–1527.
101. Yin P, Zhang X, Liu Y, et al. End-to-End Deep Learning Prediction of Neoadjuvant Chemotherapy Response in Osteosarcoma Patients Using Routine MRI. *J Imaging Inform Med.* Published online January 28, 2025.
102. Witkowska A, Petre EN, Moussa AM, et al. Feasibility and safety of percutaneous CT-guided bone biopsies in patients with cancer using a patient-mounted robotic system: a retrospective analysis of 40 consecutive biopsies. *J Vasc Inter Radio.* 2023;34(12):2174–2179.
103. Isaac A, Klontzas ME, Dalili D, et al. Revolutionising osseous biopsy: the impact of artificial intelligence in the era of personalized medicine. *Br J Radio.* 2025;98(1170):795–809.
104. Xin S, Wei G. Prognostic factors in osteosarcoma: A study level meta-analysis and systematic review of current practice. *J Bone Oncol.* 2020;21:100281.
105. Shulman DS, Whittle SB, Surdez D, et al. An international working group consensus report for the prioritization of molecular biomarkers for Ewing sarcoma. *NPJ Precis Oncol.* 2022;6(1):65.
106. Lipkova J, Chen RJ, Chen B, et al. Artificial intelligence for multimodal data integration in oncology. *Cancer Cell.* 2022;40(10):1095–1110.
107. Yu Y, Ren W, Mao L, et al. MRI-based multimodal AI model enables prediction of recurrence risk and adjuvant therapy in breast cancer. *Pharm Res.* 2025;216:107765.
108. He F, Xie L, Sun X, et al. A scoring system for predicting neoadjuvant chemotherapy response in primary high-grade bone sarcomas: a multicenter study. *Orthop Surg.* 2022;14(10):2499–2509.
109. Jiang B, Mu Q, Qiu F, et al. Machine learning of genomic features in organotropic metastases stratifies progression risk of primary tumors. *Nat Commun.* 2021;12(1):6692.
110. Huang Y, Li J, Li M, et al. Application of machine learning in predicting survival outcomes involving real-world data: a scoping review. *BMC Med Res Method.* 2023;23(1):268.
111. Chen RJ, Lu MY, Wang J, et al. Pathomic fusion: an integrated framework for fusing histopathology and genomic features for cancer diagnosis and prognosis. *IEEE Trans Med Imaging.* 2022;41(4):757–770.
112. Lu MY, Chen TY, Williamson DFK, et al. AI-based pathology predicts origins for cancers of unknown primary. *Nature.* 2021;594(7861):106–110.
113. Zhang Z, Zhao Y, Duan J, et al. Pathology-genomic fusion via biologically informed cross-modality graph learning for survival analysis. *arXiv preprint arXiv:2404.08023*. Published online 2024.
114. Yin P, Zhi X, Sun C, et al. Radiomics models for the preoperative prediction of pelvic and sacral tumor types: a single-center retrospective study of 795 cases. *Front Oncol.* 2021;11:709659.
115. Wang J, Zhao Z, Liang H, et al. Artificial intelligence assisted preoperative planning and 3D-printing guiding frame for percutaneous screw reconstruction in periacetabular metastatic cancer patients. *Front Bioeng Biotechnol.* 2024;12:1404937.
116. Mahootiha M, Tak D, Ye Z, et al. Multimodal deep learning improves recurrence risk prediction in pediatric low-grade gliomas. *Neuro Oncol.* 2025;27(1):277–290.
117. Wen Q, Yang Z, Dai H, et al. Radiomics Study for Predicting the Expression of PD-L1 and tumor mutation burden in non-small cell lung cancer based on CT images and clinicopathological features. *Front Oncol.* 2021;11:620246.
118. Zhou X, Wang H, Feng C, et al. Emerging applications of deep learning in bone tumors: current advances and challenges. *Front Oncol.* 2022;12:908873.
119. Koh DM, Papanikolaou N, Bick U, et al. Artificial intelligence and machine learning in cancer imaging. *Commun Med (Lond).* 2022;2:133.
120. Dagogo-Jack I, Shaw AT. Tumour heterogeneity and resistance to cancer therapies. *Nat Rev Clin Oncol.* 2018;15(2):81–94.
121. Rich JM, Bhardwaj LN, Shah A, et al. Deep learning image segmentation approaches for malignant bone lesions: a systematic review and meta-analysis. *Front Radiol.* 2023;3:1241651.
122. Elfarra FG, Calin MA, Parasca SV. Computer-aided detection of bone metastasis in bone scintigraphy images using parallelepiped classification method. *Ann Nucl Med.* 2019;33(11):866–874.
123. Homeyer A, Geißler C, Schwen LO, et al. Recommendations on compiling test datasets for evaluating artificial intelligence solutions in pathology [published correction appears in Mod Pathol. *Mod Pathol.* 2022;35(12):1759–1769.
124. Rayed MdE, Islam SMS, Niha SI, Jim JR, Kabir MM, Mridha MF. Deep learning for medical image segmentation: State-of-the-art advancements and challenges. *Inform Med Unlocked.* 2024;47:101504.
125. Gomariz A, Kikuchi Y, Li YY, et al. Joint semi-supervised and contrastive learning enables domain generalization and multi-domain segmentation. *Med Image Anal.* 2025;103:103575.
126. Pollack SM, He Q, Yearley JH, et al. T-cell infiltration and clonality correlate with programmed cell death protein 1 and programmed death-ligand 1 expression in patients with soft tissue sarcomas. *Cancer.* 2017;123(17):3291–3304.
127. Bi WL, Hosny A, Schabath MB, et al. Artificial intelligence in cancer imaging: clinical challenges and applications. *CA Cancer J Clin.* 2019;69(2):127–157.
128. Holste G, Lin M, Zhou R, et al. Harnessing the power of longitudinal medical imaging for eye disease prognosis using Transformer-based sequence modeling. *NPJ Digit Med.* 2024;7(1):216.
129. Sun Y, Li K, Chen D, Hu Y, Zhang S. LOMIA-T: A Transformer-based LOngitudinal Medical Image Analysis framework for predicting treatment response of esophageal cancer. In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer; 2024:426–436.
130. Korfiatis P, Kline TL, Meyer HM, et al. Implementing artificial intelligence algorithms in the radiology workflow: challenges and considerations. *Mayo Clin Proc Digit Health.* 2024;3(1):100188.
131. Guarrasi V, Soda P. Multi-objective optimization determines when, which and how to fuse deep networks: an application to predict COVID-19 outcomes. *Comput Biol Med.* 2023;154:106625.
132. Ziller A, Mueller TT, Stieger S, et al. Reconciling privacy and accuracy in AI for medical imaging. *Nat Mach Intell.* 2024;6(7):764–774.
133. Long Y, Novak L, Walsh CG. Searching for value sensitive design in applied health AI: a narrative review. *Yearb Med Inf.* 2024;33(1):75–82.