Exploration of Digital Health Technologies

Open Exploration

Open Access Systematic Review

Explainable AI for multi-omics in precision medicine: a systematic review

Muhammad Hasnain[1], Sidra Hameed[1], Komal Azim[1], Muhammad Waseem[2*]

[1]Department of Artificial Intelligence, Islamia University of Bahawalpur, Bahawalpur 63100, Pakistan
[2]Department of Zoology, Cholistan University of Veterinary and Animal Sciences, Bahawalpur 63100, Pakistan

***Correspondence:** Muhammad Waseem, Department of Zoology, Cholistan University of Veterinary and Animal Sciences, Bahawalpur 63100, Pakistan. dr.waseemcuvas@gmail.com
Academic Editor: James M. Flanagan, Imperial College London, UK
Received: April 16, 2026 **Accepted:** June 12, 2026 **Published:** August 25, 2026

Cite this article: Hasnain M, Hameed S, Azim K, Waseem M. Explainable AI for multi-omics in precision medicine: a systematic review. Explor Digit Health Technol. 2026;4:1011102. https://doi.org/10.37349/edht.2026.1011102

Abstract

Background: The convergence of multi-omics technologies and artificial intelligence (AI) has opened new frontiers in precision medicine; however, the complexity and opacity of advanced AI models remain a major barrier to clinical adoption. This systematic review aims to critically evaluate explainable AI (XAI) strategies for multi-omics integration and their role in bridging the translational gap between computational innovation and clinical utility.

Methods: A systematic literature search was conducted across PubMed/MEDLINE, Scopus, and Web of Science databases for studies published between 2020 and 2025, following PRISMA 2020 guidelines. Studies addressing multi-omics integration using explainable or interpretable AI methods in precision medicine were included. Data extraction and narrative synthesis were performed due to methodological heterogeneity.

Results: A total of 116 studies were included in the final analysis. Computational approaches ranged from classical machine learning and deep learning to graph-based and transformer architectures. XAI techniques, including SHAP (SHapley Additive exPlanations), attention mechanisms, and saliency maps, enabled interpretable predictions across gene, pathway, and network levels. Applications were most prominent in cancer subtyping, biomarker discovery, drug response prediction, and prognosis modeling. Despite promising performance, key challenges persist, including data heterogeneity, high dimensionality, batch effects, overfitting, limited reproducibility, and insufficient clinical validation.

Discussion: XAI enhances transparency, trust, and biological interpretability in multi-omics models, facilitating their integration into clinical workflows. Emerging directions such as federated learning, causal AI, foundation models, digital twins, and human-in-the-loop systems offer potential solutions to current limitations. Standardized evaluation frameworks and robust clinical validation are essential to advance real-world implementation. This review provides a comprehensive roadmap for developing reliable and clinically actionable XAI-driven multi-omics systems in precision medicine.

Keywords

explainable AI, multi-omics integration, precision medicine, machine learning, clinical translation

Introduction

Precision medicine and the need for multi-omics analysis

Precision medicine shifts healthcare from a one-size-fits-all model to interventions tailored to individual genetic, environmental, and lifestyle factors [1, 2]. This approach promises improved diagnostic accuracy, prognostic stratification, and therapeutic efficacy, particularly in oncology [3, 4]. Its realization depends on comprehensive molecular characterization, enabled by advances in high-throughput biotechnologies.

Growth and challenges of high-throughput omics data

The past two decades have seen an explosion in biomedical data from omics technologies (genomics, transcriptomics, proteomics, etc.) [5]. However, this data deluge presents significant analytical challenges: high dimensionality (often tens of thousands of features) relative to sample size, complex inter-feature correlations, and high noise levels. These factors contribute to the "curse of dimensionality," complicating statistical inference and machine learning (ML) model development [6–8].

Why multi-omics integration is essential and difficult

Single omics layers offer only a partial view of biological systems. A holistic understanding requires integrating multiple modalities to capture the interplay from DNA to metabolism [9–11]. Multi-omics integration improves disease subtyping, prognostic accuracy, and therapeutic target identification compared to single-omics analyses [12, 13]. However, integration is challenging due to heterogeneous data types, scale disparities, missing data, batch effects, and a lack of standardized preprocessing pipelines [14–17].

Limitations of conventional ML

Conventional ML models, often used for disease classification and biomarker discovery, frequently prioritize accuracy over interpretability, functioning as "black boxes" [18, 19]. This opacity is problematic in clinical settings where understanding model predictions is crucial for trust and safety [20]. Moreover, conventional feature selection often fails to capture complex, non-linear interactions across omics layers. Recent work also highlights that no single feature attribution method can satisfy all desirable properties for interpretability, leading to potential inconsistencies [21, 22].

Need for explainable artificial intelligence in clinical settings

Explainable artificial intelligence (XAI) aims to make AI models transparent, interpretable, and trustworthy [23, 24]. In precision medicine, XAI bridges predictive performance and clinical utility by elucidating the molecular drivers of model predictions [20, 25, 26]. XAI includes inherently interpretable models and post-hoc explanation methods [27, 28]. Applied to multi-omics, XAI can reveal cross-omics interactions and generate testable hypotheses [29]. For instance, attention mechanisms can highlight important features across omics layers, while graph neural networks (GNNs) can leverage biological pathways for interpretability [30]. In clinical settings, XAI fosters trust, facilitates regulatory compliance, and supports decision-making [31, 32]. However, rigorous evaluation is essential to avoid false assurances of understanding [19, 33].

Scope, objectives, and novelty of this review

While recent reviews have begun to explore connections between XAI and multi-omics integration, few provide a critically oriented examination that systematically addresses the translational gap from methodology to clinical deployment. Existing reviews have primarily focused either on technical integration methods, explainability techniques in isolation, or disease-specific applications, often without systematically addressing translational barriers such as clinical workflow integration, regulatory

requirements, explanation faithfulness, and deployment readiness. In contrast, the present review integrates computational, biological, and translational perspectives within a unified framework. This review provides a holistic, critically oriented synthesis of XAI strategies for multi-omics integration, with a dedicated focus on clinical translation. Our objectives are threefold: (1) to systematically categorize computational strategies for explainable multi-omics integration; (2) to critically assess challenges hindering clinical adoption; and (3) to identify future research directions for accelerating clinical translation. The novelty lies in its integrated focus on both computational methodologies and practical clinical deployment, providing a unique framework for developing trustworthy AI systems [28–30]. This review critically examines the translational gap, emphasizing the need for rigorous evaluation, prospective validation, and clinical workflow integration [31, 34–36], serving as a roadmap for advancing clinically impactful, interpretable multi-omics solutions.

While spatial omics technologies are rapidly advancing, this review focuses on molecular multi-omics data (genomics, transcriptomics, proteomics, metabolomics, epigenomics, microbiomics) and does not extensively cover spatial resolution or spatial omics integration. The principles discussed are extendable, but a dedicated treatment of spatial omics is beyond our scope. Figure 1 illustrates the overall workflow from multi-omics data acquisition to clinical decision support, highlighting where XAI methods intervene to enable interpretable predictions.

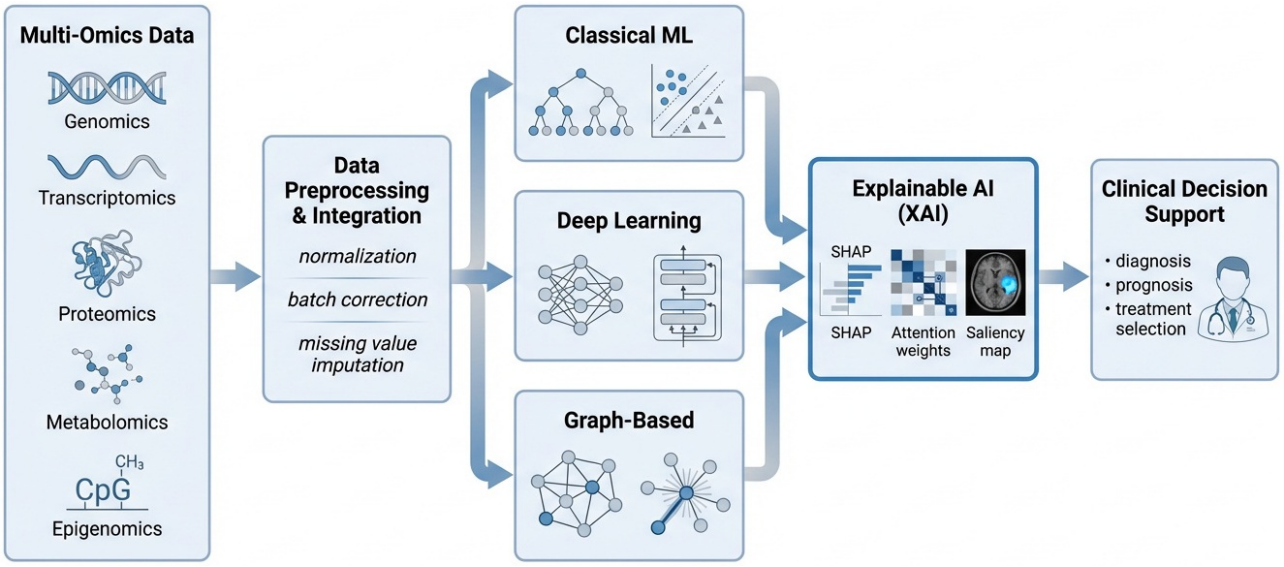

Figure 1. Workflow of multi-omics integration with XAI for precision medicine. ML: machine learning.

Conceptual background

Overview of omics layers

Biological complexity is captured across hierarchical layers by distinct omics technologies. Understanding their nature is essential for developing effective integration strategies.

Table 1 provides a summary of common omics layers, detailing their biological significance, key characteristics, and common measurement platforms. Each layer presents unique challenges—dimensionality, sparsity, noise—for ML integration. Advances in single-cell technologies now allow simultaneous measurement of multiple layers, adding further complexity [37, 38]. Collectively, these layers provide a multi-scale view where integration yields more robust and clinically actionable insights than any single layer alone [9, 10].

Genomics provides information on inherited and somatic genetic variation, including single nucleotide polymorphisms (SNPs), copy number variations (CNVs), and mutations associated with disease susceptibility and therapeutic response. Transcriptomics captures dynamic gene expression patterns and

Table 1. Common types of omics data and their biological significance.

Omics layer	Primary data type	Biological significance	Key characteristics	Common platforms
Genomics	DNA sequences, variants (SNPs: single nucleotide polymorphisms, CNVs: copy number variations, mutations)	Germline and somatic alterations; disease susceptibility; drug metabolism; heritable traits	Static across cell types; high dimensionality (millions of variants); discrete, sparse data	Whole-genome sequencing, whole-exome sequencing, SNP arrays
Transcriptomics	RNA abundance (mRNA, ncRNA)	Gene expression patterns; regulatory networks; cellular state; functional activity	Dynamic, cell-type specific; count-based, overdispersed; high dimensionality (thousands of transcripts)	RNA-seq, microarrays
Proteomics	Protein abundance, post-translational modifications	Functional effectors; signaling pathways; drug targets; direct phenotype mediators	Moderate dimensionality; continuous intensity data; partial correlation with transcriptomics	Mass spectrometry, antibody-based assays
Metabolomics	Metabolite concentrations (small molecules)	Metabolic state; physiological readout; closest to phenotype; biomarker-rich	Lower dimensionality (hundreds to thousands); high chemical diversity; dynamic range	NMR, mass spectrometry
Epigenomics	DNA methylation, histone modifications, chromatin accessibility	Gene regulation; environmental response; cell identity; tissue specificity	Tissue-specific; spatial dependencies; binary or continuous signals	Bisulfite sequencing, ChIP-seq, ATAC-seq
Microbiomics	Microbial taxonomic and functional profiles	Host-microbe interactions; immune modulation; drug metabolism; disease associations	Compositional data; high sparsity; cross-sectional or longitudinal	16S rRNA sequencing, metagenomic sequencing

Sources: Adapted from [14, 39, 40].

cellular activity, enabling characterization of disease states and regulatory networks. Proteomics reflects downstream functional protein abundance and signaling activity, often correlating more directly with phenotype than transcriptomic data alone. Metabolomics captures metabolic states and biochemical pathway activity, representing the closest molecular layer to phenotype. Epigenomics describes regulatory mechanisms such as DNA methylation and chromatin accessibility that influence gene expression without altering DNA sequence. Microbiomics characterizes host–microbe interactions that influence immunity, metabolism, and disease progression.

Characteristics of multi-omics data

Multi-omics datasets exhibit complex properties that profoundly impact modeling and interpretation.

- **High dimensionality:** The "$p \gg n$" problem (features ≫ samples) leads to overfitting, unreliable estimators, and unstable feature attributions [6–8, 28].
- **Heterogeneity:** Data types vary in statistical properties and scales (e.g., discrete genomic variants vs. continuous proteomic intensities). Integration requires careful normalization, appropriate feature processing, and model selection to prevent one modality from dominating learning [39, 41].
- **Missing values:** Ubiquitous due to technical or biological reasons. Missing data can bias estimates and complicate joint modeling, requiring sophisticated imputation or model architectures [14, 15, 40].
- **Batch effects:** Non-biological variation from sample processing can obscure true signals and lead to spurious findings if not corrected [14, 15, 17].
- **Sample imbalance:** Rare diseases or subtypes lead to biased models and explanations that may not capture clinically critical minority classes [5, 13].

These characteristics demand models that are robust to population imbalance and sampling variability and can learn structured, sparse representations. Importantly, explanations are only as reliable as the data; poor preprocessing can lead to false biological insights [19].

Computational strategies for multi-omics integration

Classical statistical and ML approaches

Before deep learning, classical statistical and ML methods formed the backbone of multi-omics analysis. These approaches remain widely used due to their interpretability, computational efficiency, and established theoretical foundations. They encompass dimensionality reduction, clustering, feature selection, and supervised algorithms such as random forests (RFs) and support vector machines (SVMs).

While classical models offer relative transparency—for example, RFs provide intrinsic feature importance measures and sparse canonical correlation analysis (sCCA) yields interpretable cross-omics relationships—their linear or shallow non-linear nature limits their capacity to capture complex, non-linear interactions across molecular layers. Moreover, their performance often plateaus when applied to high-dimensional, multi-modal data [6, 42].

Deep learning-based integration

Deep learning offers the capacity to learn hierarchical representations, capture non-linear interactions, and scale to large, heterogeneous datasets, automatically discovering complex patterns across multiple molecular layers. Architectures such as autoencoders (AEs), variational AEs (VAEs), deep neural networks (DNNs), and attention-based fusion mechanisms have been successfully applied.

These models can handle missing data, reduce dimensionality, and integrate modalities through early, intermediate, or late fusion strategies. However, their latent representations are often not directly interpretable, requiring post-hoc explanation methods. Attention mechanisms provide a degree of intrinsic interpretability, though their faithfulness is debated [21]. Emerging strategies build interpretability directly into architectures by incorporating biological prior knowledge (e.g., pathways) or using disentangled representation learning [43].

Graph-based and network-based methods

Biological systems are inherently structured as networks. Graph-based methods leverage this structure, incorporating prior knowledge and facilitating interpretability. Key approaches include GNNs such as graph convolutional networks (GCNs) and graph attention networks (GATs), as well as pathway-informed models that embed biological knowledge directly into the architecture.

The graph structure itself provides a transparent scaffold for explanations—important nodes and edges can be identified via attention weights or gradient-based attribution [44]. However, the choice of graph (e.g., protein–protein interaction networks, patient similarity graphs) significantly influences both model behavior and the resulting explanations. Additionally, aggregated messages from neighbors can complicate feature attribution, and scalability remains a challenge for very large graphs [45].

Transformer and attention-based architectures

Transformers capture long-range dependencies and model complex interactions across heterogeneous data, with intrinsic interpretability via attention mechanisms [46]. The following content highlights key transformer architectures used in multi-omics, including cross-modal attention models and hierarchical attention networks.

These models excel at modeling interactions between distant features, a key capability for biological systems. Attention weights provide a potential basis for interpretation, although they may not always faithfully represent true feature contributions. Approaches such as attention regularization can encourage more focused and interpretable patterns. The quadratic complexity of standard attention with respect to feature count remains a practical limitation for very high-dimensional omics data [46].

Hybrid and ensemble frameworks

Hybrid and ensemble frameworks combine multiple strategies to leverage complementary strengths, balancing accuracy with interpretability and robustness. Common designs include using deep learning for

feature embedding followed by an interpretable classifier (e.g., RFs), or combining predictions from models trained on different omics layers or algorithms.

These multi-stage pipelines decompose the integration task into sequential steps, simplifying development and facilitating clinical translation by allowing domain-specific constraints at each stage and incremental validation. Ensembles improve robustness and performance, and analyzing the contributions of different base models can provide insights into the relative importance of data sources. However, ensembles can be computationally expensive, and a large number of base models may reduce overall interpretability.

Data fusion strategies

The choice of fusion strategy profoundly influences model architecture, performance, and interpretability [47, 48]. The main strategies—early, intermediate, late, and hierarchical fusion—are compared in Table 2.

Table 2. Comparison of computational integration methods.

Method category	Key techniques	Strengths	Limitations	Typical applications	References
Classical statistical/ML	PCA, clustering, RF, SVM, sCCA	Interpretable; computationally efficient; established theoretical foundations; works well with limited samples	Linear or shallow non-linear; limited cross-omics interaction capture; manual feature engineering required	Disease subtyping; biomarker discovery; baseline benchmarks	[49–51]
AE-based	Denoising AEs, VAEs, stacked AEs	Unsupervised feature learning; handles missing data; dimensionality reduction; generative capabilities	Latent space may not be interpretable; requires large samples; computationally intensive	Single-cell integration; missing modality imputation; representation learning	[9]
Graph neural networks	GCNs, GATs, geometric GNNs	Incorporates biological priors; captures network structure; intrinsically interpretable via attention	Graph construction critical; scalability challenges; over-smoothing with deep layers	Pathway-informed modeling; patient similarity networks; biomarker discovery	[52–54]
Transformer/Attention	Self-attention, cross-modal attention, hierarchical attention	Captures long-range dependencies; intrinsic interpretability; flexible fusion strategies	Quadratic complexity with feature count; large data requirements; attention faithfulness concerns	Cancer classification; survival prediction; multimodal fusion	[55, 56]
Fusion strategies	Early, intermediate, late, hierarchical fusion	Balances integration depth and interpretability; modular design	No universal optimal strategy; trade-offs between cross-layer interaction capture and interpretability	All multi-omics tasks; clinical decision support	[13, 47]

ML: machine learning; PCA: principal component analysis; RF: random forest; SVM: support vector machine; sCCA: sparse canonical correlation analysis; AEs: autoencoders; VAEs: variational autoencoders; GCNs: graph convolutional networks; GATs: graph attention networks.

- Early fusion (input-level) is simple and captures cross-omics interactions but suffers from the curse of dimensionality and treats all features uniformly, complicating interpretability. Nevertheless, certain early fusion models retain partial interpretability through intrinsic feature importance measures, such as coefficient analysis in logistic regression or feature importance rankings in RFs.
- Intermediate fusion (feature-level) preserves layer-specific structure through modality-specific encoders and often uses attention to dynamically weigh contributions, balancing integration depth with interpretability.
- Late fusion (decision-level) respects distinct data characteristics and offers modality-level interpretability but cannot capture cross-omics interactions below the decision level.
- Hierarchical fusion combines multiple strategies in a nested manner, reflecting biological hierarchies (e.g., integrating features at the pathway level before modeling cross-pathway interactions).

No single fusion strategy is universally optimal; the choice depends on the biological question, data availability, and the desired trade-off between predictive accuracy and interpretability.

The four main fusion strategies—early, intermediate, late, and hierarchical—are compared schematically in Figure 2, which also indicates where interpretability components (e.g., attention) can be incorporated.

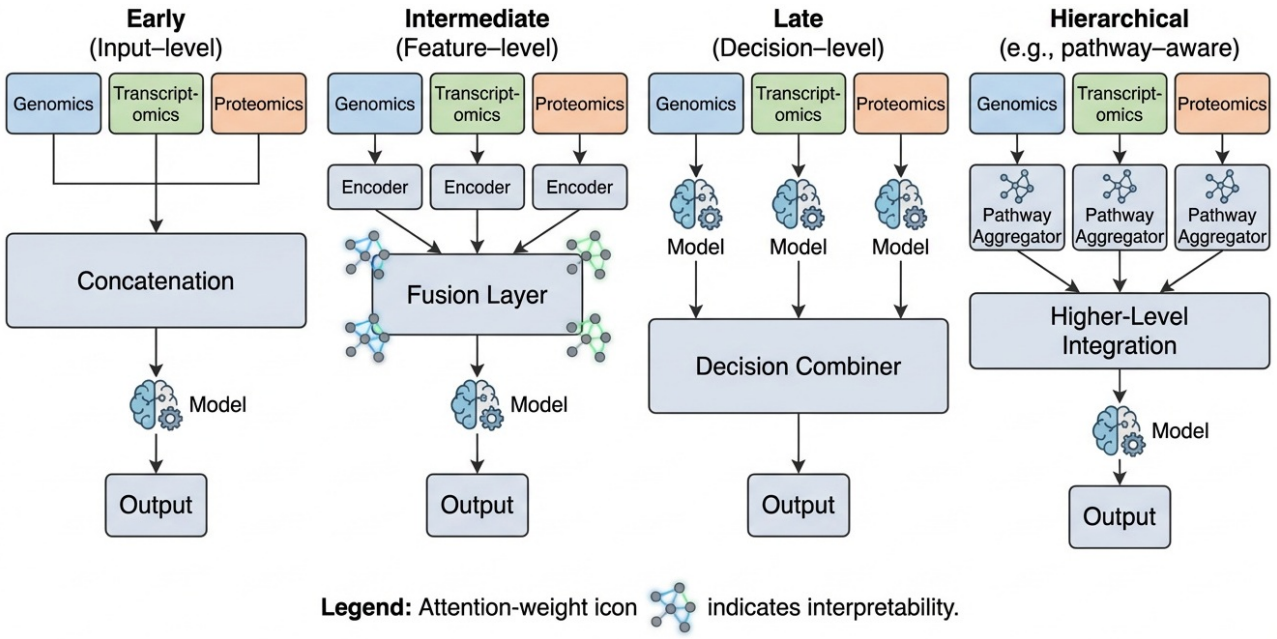

Figure 2. Comparison of data fusion strategies.

Materials and methods

Study design and reporting standards

This systematic review was conducted and reported in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA 2020) guidelines. The review methodology follows established frameworks for systematic mapping studies in computational biology and biomedical AI. This review was not registered in a prospective database such as PROSPERO.

Literature search strategy

A comprehensive literature search was performed across three major scientific databases: PubMed/MEDLINE, Scopus, and Web of Science. The search covered publications from 2020 to 2025. The search strategy was designed to capture the intersection of three conceptual domains: (1) multi-omics integration, (2) XAI, and (3) precision medicine applications. A combination of controlled vocabulary (e.g., MeSH terms) and free-text keywords was used. The search terms included:

1. Multi-omics domain: "multi-omics," "multiomics," "integrative omics," "genomics AND transcriptomics AND proteomics," "multi-modal omics," "multi-platform omics"
2. XAI domain: "explainable AI," "interpretable machine learning," "SHAP," "LIME (Local Interpretable Model-agnostic Explanations)," "attention mechanism," "saliency map," "feature attribution," "model interpretability," "transparent AI"
3. Clinical domain: "precision medicine," "personalized medicine," "clinical decision support," "cancer subtyping," "biomarker discovery," "drug response prediction," "survival analysis," "clinical translation"

These terms were combined using Boolean operators (AND, OR), with the core query structure:

("multi-omics" OR "integrative omics") AND ("explainable AI" OR "interpretable machine learning") AND ("precision medicine" OR "clinical application").

In addition, manual screening of reference lists from relevant review articles and citation tracking using Google Scholar were performed to identify additional studies. Records identified through these methods were included within the overall dataset but were not separately quantified.

Eligibility criteria

Studies were included if they met the following criteria:

(1) addressed multi-omics data integration involving at least two molecular layers;

(2) employed XAI methods or explicitly discussed model interpretability;

(3) were applied within precision medicine contexts (e.g., oncology, cardiovascular, neurological, or complex diseases);

(4) reported quantitative performance metrics or provided sufficient methodological detail;

(5) were published in peer-reviewed journals or conference proceedings;

(6) were written in English.

Studies were excluded if they:

(1) focused solely on single-omics analysis;

(2) used AI models without any interpretability component;

(3) lacked biomedical or clinical relevance;

(4) were editorials, opinion pieces, or abstracts without full methodological detail;

(5) relied exclusively on simulated data without validation on real datasets;

(6) focused primarily on spatial omics without molecular multi-omics integration.

Study selection process

The study selection process was conducted in accordance with PRISMA 2020 guidelines. Screening was performed in three stages.

First, titles and abstracts were screened independently by two reviewers (S.H. and M.H.) to exclude irrelevant studies. Disagreements were resolved through discussion, and when necessary, a third reviewer (M.W.) adjudicated. Second, full-text articles were assessed against the eligibility criteria. Third, a methodological quality assessment was performed to ensure inclusion of robust and clinically relevant studies. A total of 847 records were identified through database searches. After title and abstract screening, 412 records were excluded, leaving 435 articles for full-text assessment. Following full-text review and quality assessment, 116 studies were included in the final synthesis. The study selection process is illustrated in the PRISMA 2020 flow diagram (Figure 3).

Data extraction and synthesis

Data extraction was performed using a standardized template. The following information was collected from each included study:

(1) publication details (authors, year, journal);

(2) omics layers integrated;

(3) integration strategy (early, intermediate, late, hierarchical fusion);

(4) AI/ML methodology;

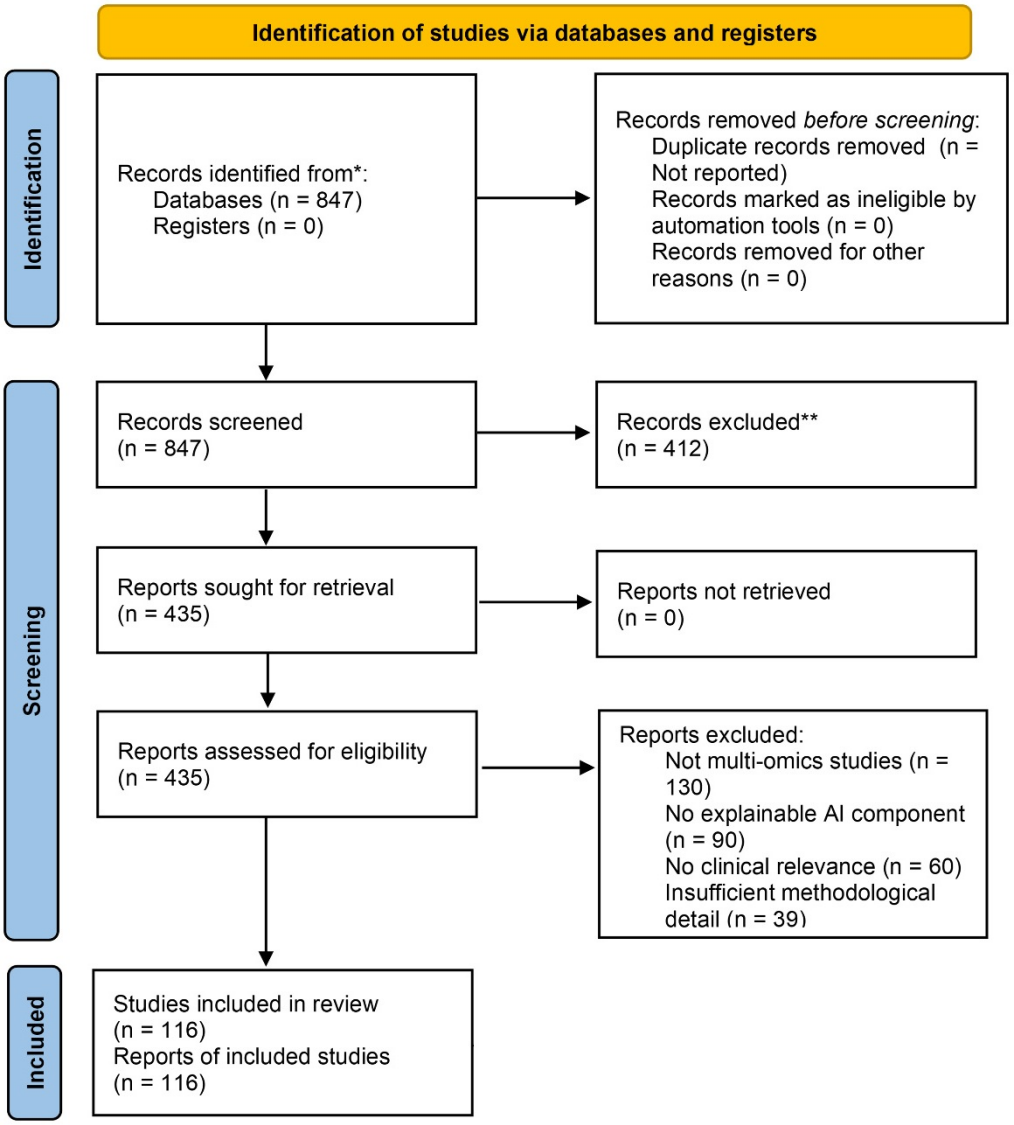

Figure 3. PRISMA 2020 flow diagram showing the study selection process. *: Records identified through manual searching and citation tracking. **: Records excluded after eligibility assessment. Adapted from https://www.prisma-statement.org/. Accessed May 20, 2026. © 2024-2026 the PRISMA Executive. Licensed under a CC BY 4.0.

(5) XAI methods employed;

(6) application domain;

(7) datasets used;

(8) evaluation metrics;

(9) validation approach;

(10) reported limitations and future directions.

Due to heterogeneity in datasets, study designs, and evaluation metrics, a quantitative meta-analysis was not feasible. Therefore, a narrative synthesis approach was adopted, organizing findings into thematic categories including computational strategies, XAI methodologies, clinical applications, and translational challenges.

Quality assessment and risk of bias

Risk of bias was assessed using an adapted version of the PROBAST (Prediction model Risk of Bias Assessment Tool) framework. The following domains were evaluated:

(1) participant selection and data representativeness;

(2) predictor assessment and preprocessing;

(3) outcome definition;

(4) model development and overfitting control;

(5) appropriateness and faithfulness of XAI methods;

(6) performance evaluation;

(7) validation strategy.

Each study was categorized as having low, moderate, or high risk of bias. A subset of studies (25%) was independently evaluated by two reviewers to ensure consistency, yielding substantial inter-rater agreement ($\kappa = 0.78$). Discrepancies were resolved through consensus.

Limitations of the review

Several limitations should be acknowledged. Despite a comprehensive search strategy, relevant studies may have been missed due to database coverage or non-standard terminology. Publication bias may have favored studies reporting positive results. Rapid methodological advancements may not be fully captured despite inclusion up to 2025. Heterogeneity across studies precluded quantitative meta-analysis. Additionally, assessment of XAI methods involves subjective interpretation due to the absence of standardized evaluation benchmarks.

Despite these limitations, this methodology provides a robust and reproducible synthesis of current research on XAI for multi-omics integration in precision medicine.

Results

Importance of explainability in biomedicine

The increasing complexity of AI models in biomedicine has created a fundamental tension between predictive accuracy and interpretability. This tension is a critical barrier to clinical adoption, regulatory approval, and scientific discovery [23, 24]. XAI addresses this by making model decisions transparent and trustworthy. Importantly, the required depth and form of explainability vary according to clinical application. Screening and risk stratification tasks may only require global feature importance or simplified explanations, whereas high-risk decisions such as treatment selection or intensive care interventions demand more rigorous, causally informed, and patient-specific explanations. Therefore, explainability should be considered context-dependent rather than universally standardized across all clinical tasks. The importance of explainability in biomedicine stems from several imperatives:

- **Clinical reasoning:** Physicians must understand why a model recommends a particular diagnosis or treatment to integrate it into their reasoning and communicate it to patients [32, 33].
- **Regulatory compliance:** Frameworks like the EU AI Act and FDA guidance increasingly mandate transparency and interpretability for high-risk medical AI systems [34, 57].
- **Safety and accountability:** Understanding model drivers enables clinicians to identify errors, assess alignment with clinical knowledge, and perform root cause analysis when adverse outcomes occur [19, 58].
- **Scientific discovery:** Explainable models reveal molecular features and interactions, generating testable hypotheses, identifying novel biomarkers, and uncovering disease mechanisms [27, 29, 59].
- **Fairness and bias:** XAI can reveal whether models rely on protected attributes or proxies for demographic characteristics, enabling bias mitigation before deployment [31, 32, 60].

In multi-omics integration, explainability is particularly critical. Models must not only identify important features but also contextualize them within biological pathways, distinguish signal from noise,

and communicate uncertainty—demanding XAI approaches that are both technically sophisticated and biologically meaningful [28].

Types of XAI

Explainability encompasses approaches that differ in scope, timing, and relationship to the underlying model.

Global vs. local explainability

Global explainability characterizes model behavior across the entire input space, revealing which features, omics layers, or pathways are most influential on average [20, 26]. This is valuable for hypothesis generation but can obscure subgroup-specific patterns. Local explainability focuses on individual predictions, providing insight into why a model made a specific decision for a particular patient—essential for clinical decision-making [23, 31].

Intrinsic interpretability vs. post-hoc explainability

Intrinsically interpretable models (e.g., linear models, decision trees, sparse models, attention mechanisms) are transparent by design, with explanations faithful to the model's computation [22, 24]. However, they may trade off performance for transparency. Post-hoc explainability methods [e.g., SHapley Additive exPlanations (SHAP), Local Interpretable Model-agnostic Explanations (LIME)] explain trained models without modifying architecture, offering flexibility but potentially sacrificing faithfulness [20, 28]. Attention mechanisms, while often used post-hoc, can be considered intrinsic when attention weights directly inform interpretation [46, 61].

Model-agnostic methods

Model-agnostic methods interpret any model by analyzing input-output relationships without accessing internal parameters. They are widely used in multi-omics for their flexibility.

While these methods enable the interpretation of any black-box model, they are not without drawbacks. SHAP, though theoretically grounded, can be computationally expensive, and its approximations may be unstable. LIME offers fast, local explanations but can be sensitive to perturbations. Permutation importance is efficient for global feature ranking but ignores interactions, and partial dependence plots assume feature independence—an assumption rarely met in multi-omics data. Given these limitations, practitioners are advised to use multiple methods and assess convergence of explanations [21, 45].

Model-specific interpretability methods

Model-specific methods leverage internal architecture components to generate explanations that are inherently faithful to the model's computation. The following content summarizes key approaches, including attention weights, saliency maps, gradient-based methods, and rule-based/sparse models.

Attention weights, prevalent in transformers and GATs, offer intrinsic interpretability and can highlight cross-modal interactions. However, their faithfulness is debated, as attention may not always reflect true feature importance [21]. Saliency and gradient-based methods provide fine-grained, computationally efficient attributions but can be noisy and unstable near sharp decision boundaries. Rule-based and sparse models (e.g., LASSO, decision trees) are intrinsically interpretable and produce transparent decision logic, though they may underperform on highly complex tasks. Hybrid approaches that combine deep feature extraction with sparse or rule-based final layers are increasingly used to bridge this performance-interpretability gap [62, 63]. Table 3 provides a comparative overview of the most common model-agnostic techniques—SHAP, LIME, permutation importance, and partial dependence analysis—detailing their core ideas, advantages, and limitations.

Table 3. Comparison of XAI methods with advantages and limitations.

Method category	Specific methods	Core idea	Advantages	Limitations	Use in multi-omics	References
Model-agnostic (Post-hoc)	SHAP	Shapley values for feature attribution	Theoretically grounded; consistent; global and local explanations	Computationally expensive; approximations may be biased	Identifying key genes, pathways; biomarker prioritization	[45, 64]
	LIME	Local surrogate models	Fast; model-agnostic; intuitive	Unstable; local approximations may not generalize	Patient-specific explanations; clinical decision support	[65, 66]
	Permutation importance	Performance drop after feature shuffling	Simple; efficient; global ranking	Ignores interactions; assumes feature independence	Comparing omics layer importance; feature screening	[26, 28]
	Partial dependence	Marginal effect visualization	Reveals directionality; intuitive plots	Independence assumption; unrealistic extrapolation	Understanding feature-outcome relationships	[27, 28]
Model-specific	Attention weights	Learned importance scores from attention	Intrinsic interpretability; captures interactions; hierarchical	Faithfulness concerns; may not reflect true importance	Pathway-level explanations; cross-modal fusion	[67]
	Saliency maps	Gradients of output w.r.t. inputs	Computationally efficient; fine-grained	Noisy; gradient saturation; unstable	Genomic sequence interpretation; imaging-omics fusion	[68, 69]
	Sparse/Rule-based	LASSO, decision trees, rule lists	Intrinsically interpretable; faithful	May sacrifice performance; limited non-linear interactions	Feature selection; interpretable final-layer predictors	[22, 54, 58]

SHAP: SHapley Additive exPlanations; XAI: explainable artificial intelligence; LIME: Local Interpretable Model-agnostic Explanations; w.r.t.: with respect to; LASSO: Least absolute shrinkage and selection operator.

Biological interpretation of explanations

Translating XAI outputs into biological knowledge requires bridging computational attributions with established frameworks—from individual genes to pathways and disease mechanisms.

Gene-level interpretation

Feature attributions prioritize individual genes, transcripts, or proteins. SHAP, for example, has identified cancer survival drivers [70] and inflammatory bowel disease genes [71]. Interpretation requires integrating external knowledge (e.g., Gene Ontology, KEGG) for functional enrichment and accounting for feature redundancy, which sparse models help mitigate [49, 51].

Pathway-level interpretation

Pathways aggregate functionally related features, offering robustness to noise and alignment with clinical reasoning. Models such as Pathformer, DeepKEGG, and GraphPath enable direct pathway-level identification [52, 55, 72]. Pathway explanations can validate against known biology or generate novel hypotheses [53, 59].

Biomarker prioritization

XAI ranks features by predictive contribution, potentially identifying composite biomarkers that account for interactions [64, 70]. Because attribution rankings may vary across models and resampling schemes, their stability should be assessed across repeated validation runs [45]. Validation in independent cohorts and functional experiments remains essential [34].

Disease mechanism inference

Explanations capturing feature interactions—such as attention patterns highlighting co-contributing genes or pathway cross-talk—can suggest regulatory networks and causal hypotheses [27]. However, attributions

reflect correlations, not causation. Integration with causal frameworks (e.g., Mendelian randomization) is needed to move from correlation to mechanism [71, 73, 74].

From computational attribution to validated biomarker

The ultimate test of an XAI-derived hypothesis is experimental validation. Attributions may capture spurious correlations driven by batch effects, hidden confounders, or model overfitting; therefore, computational prioritization must be followed by orthogonal evidence. This can include genetic perturbation experiments (e.g., CRISPR screens), pharmacologic inhibition, or analysis of independent cohorts with orthogonal assays. Recent frameworks propose combining XAI with high-throughput validation pipelines to accelerate target discovery [45, 75].

Strengths and limitations of XAI methods in omics research

Strengths

- **Transparency and trust:** XAI fosters adoption by providing justifications that align with medical knowledge and enable accountability [23, 32].
- **Biological discovery:** Prioritizes features for experimental validation, accelerating biomarker and target identification [27, 51].
- **Regulatory alignment:** Supports auditing, bias detection, and compliance with transparency requirements [18, 34].
- **Handling dimensionality:** Distills high-dimensional omics data into interpretable summaries [28, 45].

A comprehensive taxonomy of XAI methods, including their relationship to intrinsic interpretability and post-hoc explainability, is presented in Figure 4. The figure also maps each method to typical biological interpretation scales.

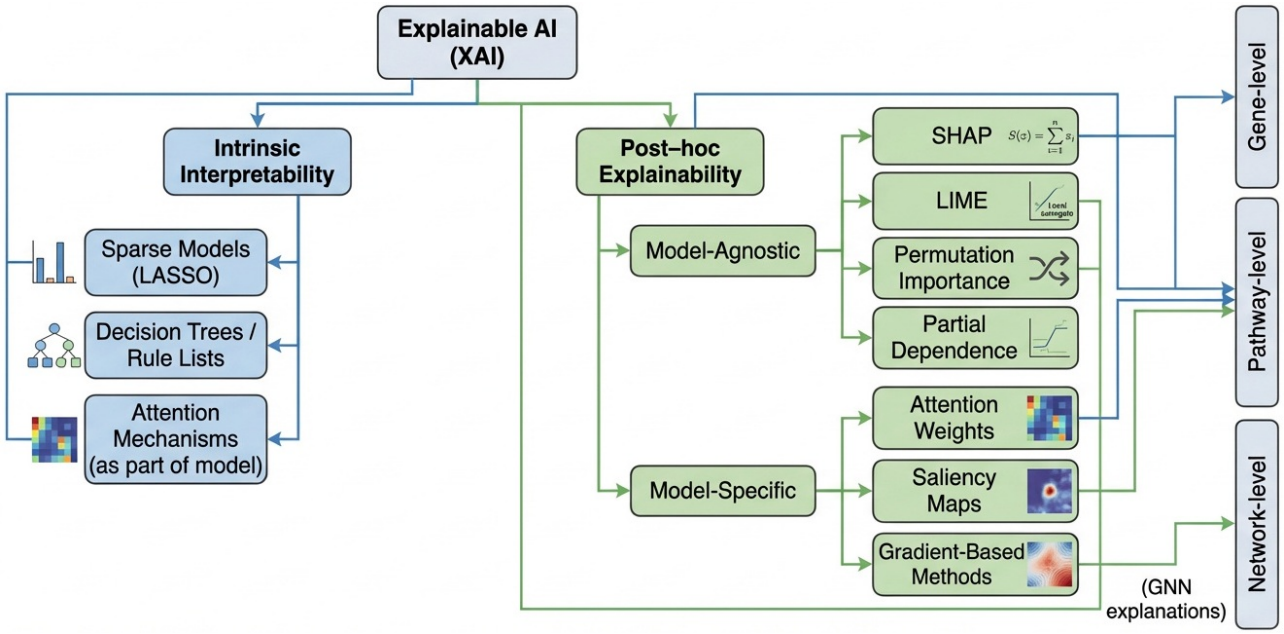

Figure 4. Taxonomy of XAI methods used in multi-omics integration, illustrating the relationship between intrinsic interpretability, post-hoc explanation techniques, and biological interpretation scales. LASSO: Least Absolute Shrinkage and Selection Operator; SHAP: SHapley Additive exPlanations; LIME: Local Interpretable Model-agnostic Explanations; GNN: graph neural network.

Limitations

- **Faithfulness and stability:** Post-hoc methods approximate model behavior and may produce unstable or contradictory explanations [21, 45]. No single method satisfies all desirable properties [21].

- **Computational cost:** Generating explanations for high-dimensional models can be prohibitively expensive [20].
- **Lack of biological validation:** Attributions may reflect spurious correlations or batch effects; standardized benchmarks for biological plausibility are lacking [19, 45].
- **Over-reliance:** Plausible but inaccurate explanations can mislead decisions; user training and triangulation across methods are essential [19].
- **Interpretability-performance trade-off:** Intrinsically interpretable models may underperform black-box models, though hybrid approaches are narrowing this gap [22, 54]. However, for high-stakes clinical decisions, some argue that interpretability should be a non-negotiable design requirement, even at the cost of a modest reduction in predictive accuracy, because post-hoc explanations may be unfaithful or misleading [22]. The 'right to explanation' and regulatory trends (e.g., EU AI Act) further prioritize transparency over maximal performance in medical AI.

Data resources, benchmarking, and evaluation

Major public omics databases and repositories

Public repositories provide the raw data essential for developing and validating multi-omics AI models.

- **The Cancer Genome Atlas (TCGA):** A landmark resource with multi-omics profiles (genomic, transcriptomic, epigenomic, proteomic) and clinical annotations across 33 cancer types. TCGA serves as the de facto benchmark for cancer-focused integration studies, enabling method comparison on common datasets [13, 42, 50, 76].
- **Gene Expression Omnibus (GEO):** Archives transcriptomic and epigenomic data across diverse organisms and conditions. Its heterogeneity—in platforms, preprocessing, and annotations—enables cross-disease analyses and serves as a stress test for integration methods [5, 39].
- **Encyclopedia of DNA Elements (ENCODE):** Provides functional genomics data (chromatin accessibility, histone modifications, transcription factor binding) across cell types and tissues. It is a critical resource for constructing biological priors (e.g., regulatory networks) that inform interpretable models [14, 63, 77].
- **Clinical Proteomic Tumor Analysis Consortium (CPTAC):** Complements TCGA with deep proteomic and phosphoproteomic tumor profiling, enabling integration with genomics to characterize cancer biology [12, 39].
- **Metabolomics repositories:** Resources such as the Metabolomics Workbench, MetaboLights, and HMDB host metabolomic datasets from disease cohorts and population studies, capturing downstream functional states [39, 40, 78]. Their high chemical diversity and platform variability pose standardization challenges [15].
- **Other repositories:** The International Cancer Genome Consortium (ICGC), Genotype-Tissue Expression (GTEx), Human Cell Atlas, and single-cell repositories (e.g., Single Cell Expression Atlas) provide complementary data across tissues, diseases, and cellular resolution [37, 38].

Benchmark datasets for multi-omics AI studies

Benchmark datasets—curated collections with standardized preprocessing, defined splits, and consensus ground truths—enable fair comparison of computational methods.

Benchmarking efforts

Systematic comparisons have established benchmark protocols. Duan et al. (2021) [50] evaluated 12 multi-omics integration methods for cancer subtyping across five TCGA cancer types, assessing clustering accuracy, stability, and interpretability. Leng et al. (2022) [42] benchmarked 12 deep learning fusion methods on TCGA for survival prediction and subtype classification, finding that deep learning generally outperformed classical approaches, though architecture choice significantly influenced results.

Benchmarking for interpretability

Reynolds and Pan (2025) [45] benchmarked feature attribution methods for predictive genomics, demonstrating that different methods yield inconsistent results with high variability—highlighting the need for standardized XAI evaluation frameworks. Azher et al. (2023) [79] assessed pretraining strategies in interpretable multimodal deep learning for cancer prognostication, providing a benchmark for both predictive performance and explanation quality.

Commonly used datasets

Several TCGA cancer types serve as standard benchmarks: breast cancer (TCGA-BRCA) for subtyping and survival [80, 81]; lung adenocarcinoma (TCGA-LUAD) and glioblastoma (TCGA-GBM) [82, 83]. These datasets typically include RNA-seq, DNA methylation, copy number variation, and clinical data, with standardized preprocessing pipelines (e.g., UCSC Xena, GDAC Firehose) facilitating reproducibility.

However, TCGA and many other public repositories predominantly consist of individuals of European ancestry, which limits the generalizability of models trained on these data and can perpetuate health disparities if not addressed [32]. This demographic skew underscores the need for diverse cohorts and fairness assessments, as discussed in the subsection "Ethical, privacy, and fairness concerns".

Benchmarking challenges

Variability in train–test splits, cross-validation strategies, and evaluation metrics complicates direct comparison [42, 50]. Batch effects and platform heterogeneity confound method evaluation [14, 75]. Many benchmarks lack independent external validation cohorts, limiting assessment of generalizability. For XAI, there is no consensus on quantitative evaluation of explanation quality [45]. Additionally, most benchmarks focus on cancer, leaving other disease domains underrepresented.

Common evaluation metrics

Rigorous evaluation depends on the prediction task—classification, survival analysis, or clustering—and must account for class imbalance, censoring, and clinical context.

Classification metrics

For binary or multiclass classification (e.g., subtype prediction, treatment response), standard metrics include:

- **Accuracy:** Intuitive but misleading under class imbalance [80, 81].
- **AUC-ROC:** Widely used for its insensitivity to class imbalance, though it may be optimistic with very small minority classes [42, 50, 80].
- **Precision-recall (AUPRC) and F1-score:** Recommended when class imbalance is severe, as they focus on minority class performance [80, 81].

Survival analysis metrics

For time-to-event outcomes (e.g., overall survival):

- **Concordance index (C-index):** The most commonly used metric, measuring the probability that the model correctly ranks the risk order of two patients. Values range from 0.5 (random) to 1.0 (perfect) [42, 77, 84].
- **Time-dependent AUC (AUC(t)):** Evaluates discrimination at specific time points, useful when performance varies over disease course [85].
- **Integrated Brier score (IBS):** Measures overall prediction accuracy, combining discrimination and calibration, though less frequently reported [79].

Calibration metrics

Calibration—agreement between predicted probabilities and observed frequencies—is critical for clinical decisions. Calibration plots and the Hosmer–Lemeshow test are used for binary outcomes; for survival models, calibration slope and Gronnesby–Borgan test assess calibration [79].

Clustering metrics

For unsupervised subtyping, internal metrics (silhouette coefficient, Davies–Bouldin index) assess cluster cohesion without ground truth [51]. When reference subtypes exist, adjusted Rand index (ARI) and normalized mutual information (NMI) measure agreement [50, 80].

Practical considerations

Evaluation protocols must account for data characteristics. Nested cross-validation prevents overfitting during hyperparameter tuning [42]. Stratified sampling maintains class proportions across folds [80]. Multiple random splits or bootstrapping provide confidence intervals [81]. For XAI, metrics for explanation quality—faithfulness, stability, biological alignment—are emerging but not yet standardized [28, 45].

Importance of external validation and reproducibility

Models trained on curated datasets may fail on cohorts with different demographics, processing protocols, or disease distributions—a generalizability gap exacerbated by batch effects and platform heterogeneity [14, 17, 19, 34]. External validation on independent, geographically or temporally distinct cohorts is the gold standard for assessing generalizability. Many influential studies now require validation on at least one independent dataset, often from a different repository (e.g., ICGC) [80, 84–86].

Reproducibility challenges

Variability in preprocessing (normalization, imputation, batch correction) dramatically affects results [14, 40]. Model training involves random seeds, hyperparameters, and hardware-specific optimizations. Explanation methods themselves exhibit instability across runs. Lack of standardized benchmarks and evaluation protocols further complicates reproducibility [42, 50].

Best practices

- Prospective validation remains the ultimate test; in its absence, rigorous external validation on independent datasets is expected. Meta-analytic approaches can pool performance across multiple cohorts [85].
- Transparent reporting following guidelines like TRIPOD—detailing data sources, preprocessing, architecture, hyperparameters, and evaluation protocols—is essential. Open-source code and data sharing enable replication [34].
- Cross-study benchmarking initiatives (e.g., DREAM Challenges, Multi-Omics Benchmark) provide standardized frameworks for fair comparison [42, 50].
- Sensitivity analyses assess robustness to preprocessing choices, hyperparameters, or dataset composition [14, 17].
- Repeated cross-validation with multiple random splits provides stable performance estimates; nested cross-validation prevents information leakage [42].

Role of explainability in validation

XAI can support validation: explanations that align with established biological knowledge or remain consistent across validation cohorts increase confidence in the model [27, 28]. Conversely, explanations that change drastically between datasets may indicate overfitting. Evaluating explanation stability across random seeds and data perturbations is therefore critical [45].

Major challenges and limitations

Although these challenges are discussed separately, they are highly interconnected. Data heterogeneity and batch effects constitute foundational bottlenecks that propagate downstream problems including overfitting, poor generalization, unstable explanations, and reproducibility failures. Similarly, limited sample sizes amplify both model instability and interpretability challenges. Consequently, improving data quality, harmonization, and preprocessing standardization should be considered a primary priority before optimization of advanced model architectures. Despite significant progress in XAI for multi-omics integration, the field faces formidable challenges that impede clinical translation and scientific reproducibility. These challenges span data characteristics, methodological limitations, evaluation practices, and ethical considerations.

Data heterogeneity and integration difficulty

Multi-omics data are inherently heterogeneous, encompassing diverse molecular layers measured on fundamentally different scales with distinct statistical properties and noise structures [39, 41]. Integrating these disparate data types into a unified analytical framework requires careful normalization to prevent one modality from dominating learning. The lack of consensus on optimal integration strategies further complicates method selection and comparison [42, 50].

Missing modalities and incomplete samples

Complete multi-omics profiling across all layers is rarely achievable. Missing modalities are common, and the pattern of missingness is often informative [14, 15]. Most integration methods require complete data or rely on imputation, which can introduce bias. Deep learning approaches with dedicated encoders can handle missing modalities by masking, but performance typically degrades when missingness is high [61]. Robust methods that gracefully handle missing data while providing reliable explanations remain an active research area.

Small sample size versus high feature dimension

The classic "$p \gg n$" problem is magnified in multi-omics settings, where total features often exceed sample size by orders of magnitude [6, 8]. This leads to overfitting, unstable feature selection, and inflated variance in performance estimates [7]. Deep learning models are particularly prone to overfitting on small cohorts unless heavily regularized or pre-trained. The curse of dimensionality also affects XAI: feature attribution methods become unstable when many correlated features are present [28, 45].

Batch effects and preprocessing variability

Batch effects—non-biological variation introduced by sample processing—are pervasive in omics studies [17]. They can systematically confound analyses, leading models to learn batch signatures rather than biological signals. In multi-omics studies, batch effects may affect each layer independently, creating complex confounding patterns that are difficult to correct [14]. Moreover, the choice of preprocessing—normalization, imputation, feature filtering—significantly influences downstream model performance and explanations [41]. Common approaches for correcting batch effects include ComBat, surrogate variable analysis (SVA), Harmony, mutual nearest neighbors (MNN), and deep learning-based domain adaptation frameworks. These methods aim to remove non-biological variation while preserving true biological signals, though overcorrection may inadvertently eliminate clinically relevant variation.

Model overfitting and poor generalization

The combination of high dimensionality, small sample sizes, and complex model architectures creates substantial risk of overfitting, where models perform well on training data but fail to generalize to independent cohorts [42, 80]. External validation often reveals dramatic performance drops [34]. Generalization is further challenged by differences in patient populations, treatment protocols, and measurement platforms across institutions.

Lack of interpretability in deep models

Deep learning models achieve state-of-the-art performance but are notoriously opaque [19]. While post-hoc XAI methods aim to explain these black boxes, their faithfulness is questionable [21]. Feature attribution methods can produce inconsistent or contradictory explanations [45]. Attention mechanisms, often cited as intrinsic interpretability, may not accurately reflect feature importance. Intrinsically interpretable models often sacrifice predictive performance [22]. The trade-off between interpretability and performance remains unresolved.

Reproducibility and benchmark inconsistency

Reproducibility is a major concern in multi-omics AI. Variability in preprocessing pipelines, software versions, random seeds, and hardware can lead to divergent outcomes [14]. Even when the same dataset is used, different studies report widely varying performance due to differences in train-test splits, evaluation metrics, and cross-validation strategies [50]. Many studies fail to share code or detailed preprocessing steps, precluding independent replication. Standardized benchmarks for XAI evaluation are lacking [28, 45].

Limited clinical validation

Few multi-omics AI models have undergone rigorous clinical validation. Most studies are retrospective, conducted on curated research cohorts that may not reflect real-world patient populations [34, 60]. Prospective validation is rare due to cost, time, and regulatory barriers. Clinical utility—whether a model improves patient outcomes—is almost never assessed. Integration into clinical workflows requires user-friendly interfaces, real-time inference, and alignment with existing electronic health record (EHR) systems, all of which are rarely addressed [31, 32].

Ethical, privacy, and fairness concerns

Multi-omics data contain sensitive genetic and health information that could be misused if not properly protected [60]. Centralized data sharing for model training poses privacy risks. Fairness is another critical concern: models trained on predominantly European-ancestry cohorts may not generalize to underrepresented populations, potentially exacerbating health disparities [32]. XAI can help detect and mitigate bias by revealing whether models rely on protected attributes, but such analyses are rarely performed. Ethical frameworks for responsible AI in healthcare are still evolving.

Additional ethical concerns include informed consent for secondary use of omics data, explainability transparency obligations, algorithmic accountability, and the potential misuse of predictive genomic information by insurers or employers. Moreover, excessive reliance on AI-generated recommendations may reduce clinician autonomy if explanation mechanisms are not critically evaluated. Ethical deployment therefore requires human oversight, transparent governance frameworks, and continuous auditing of model fairness and safety.

Applications in precision medicine

The convergence of multi-omics integration and XAI has catalyzed significant advances across the precision medicine continuum. From cancer subtyping to drug response prediction, XAI-enhanced multi-omics models are enabling more accurate, interpretable, and clinically actionable insights. This section reviews key application domains, highlighting how XAI methods contribute to both predictive performance and biological understanding, while critically examining the evidence base and remaining challenges.

Cancer classification and subtype prediction

Cancer classification has been a flagship application for multi-omics integration, as the heterogeneous nature of malignancies necessitates molecular subtyping to guide treatment decisions. XAI methods have been instrumental in moving beyond black-box predictions to provide interpretable subtype assignments that align with biological knowledge and clinical practice.

Deep learning-based integration methods have demonstrated superior performance in cancer subtype classification compared to classical approaches [42, 80]. For instance, the attention-fusion model for multi-omics (AMMO) integrates genomic, transcriptomic, and clinical data for lung adenocarcinoma classification, achieving high accuracy while providing attention weights that highlight key genes and pathways driving subtype assignments. GNNs such as DeepMoIC and hierarchical architectures like MoAGNN [87] similarly capture complex relationships across omics layers.

XAI contributes to subtype interpretability by revealing the molecular determinants of each cluster. For example, GraphPath uses graph attention on pathway–pathway interaction networks to stratify patients while highlighting key pathways and their cross-talk [53]. These interpretable subtypes facilitate clinical adoption by aligning with established molecular taxonomies and suggesting targeted therapeutic strategies.

Biomarker discovery and validation

Biomarker discovery is a primary translational outcome of multi-omics analysis, and XAI methods have become essential tools for prioritizing candidate biomarkers from high-dimensional data [27, 51]. By attributing model predictions to specific molecular features, XAI enables the identification of both individual biomarkers and composite signatures that are predictive across omics layers.

SHAP-based analyses have been widely used. In Crohn's disease, SHAP applied to an integrative ML model identified genes involved in glycerophospholipid metabolism as critical contributors to disease classification [71]. For cancer immunotherapy response, integrated explainable ML combined with multi-omics analysis identified immune-related gene signatures and tumor mutational burden features that jointly predict outcomes [70].

Explainable GNNs have been developed specifically for biomarker discovery [53], and pathway-informed frameworks such as DeepKEGG enable pathway-level biomarker discovery for cancer recurrence prediction [72]. These methods not only prioritize biomarkers but also embed them within biological context, facilitating validation and mechanistic understanding.

Drug response and treatment sensitivity prediction

Predicting individual patient responses to therapeutic interventions is a cornerstone of precision oncology. Multi-omics integration provides a comprehensive view of drug targets, resistance mechanisms, and tumor microenvironment factors, while XAI offers insights into why a patient is predicted to respond or not [88, 89].

Deep learning frameworks for drug response prediction have been developed that integrate multi-omics data with drug chemical structures or biological pathway information. For example, DeepFusionCDR employs multi-omics integration and molecule-specific transformers to enhance prediction of cancer drug responses, with attention mechanisms providing interpretability. In breast cancer, deep learning-assisted multi-omics integration has been applied to predict both survival and drug response, with feature importance analysis revealing known and novel determinants of sensitivity [90, 91].

XAI contributes to therapy prediction by elucidating mechanisms of sensitivity and resistance. SHAP attributions have been used to identify which genomic alterations or expression patterns drive predicted responses, enabling clinicians to understand the rationale behind a recommendation [64]. In immunotherapy, multi-omics integration with XAI has identified immune-related signatures predictive of response in blood cancers, informing patient selection [89].

Patient stratification and risk prediction

Patient stratification extends beyond cancer subtyping to encompass risk assessment across diverse diseases. Multi-omics integration combined with XAI enables the identification of patient subgroups with distinct clinical trajectories and the prediction of individual risk for adverse outcomes.

In cardiovascular translational research, ML and multi-omics integration have been applied to stratify patients by risk and to identify biomarkers associated with disease progression [92]. For chronic liver

disease, AI-guided multi-omics data integration has been used to predict disease severity and progression [93]. In nephrology, multi-omics integration approaches have been developed to improve patient stratification for kidney diseases, though challenges in data harmonization and validation persist [94].

XAI enhances risk prediction by providing patient-specific explanations that can be communicated to clinicians. For example, in chronic kidney disease prediction, SHAP and LIME analyses have been used to interpret ML models, generating individualized explanations that identify which clinical and omics features most influence a patient's risk score [66]. These explanations support shared decision-making and help clinicians identify modifiable risk factors.

Survival analysis and prognosis modeling

Survival analysis is a critical task in oncology and chronic disease management, and multi-omics integration has consistently been shown to improve prognostic accuracy compared to clinical variables alone [85]. XAI methods provide interpretability for survival models, identifying time-dependent biomarkers and explaining individual risk predictions.

Interpretable deep learning frameworks for survival analysis have been developed that incorporate both clinical and multi-omics data. Autosurv is an interpretable deep learning framework that integrates multi-omics and clinical data for cancer survival analysis, using attention mechanisms to identify features associated with survival outcomes [84]. Collaborative transformer models, such as CoFormerSurv, have been applied to multi-omics survival analysis, providing attention-based explanations that reveal how feature importance changes over time [56]. Pathformer incorporates biological pathway information into a transformer architecture, enabling pathway-level interpretation of prognostic predictions [55].

XAI contributes by enabling clinicians to understand which molecular features drive prognosis and to assess whether predictions align with clinical knowledge. For instance, SHAP analysis of survival models for pancreatic cancer revealed metabolic and signaling pathway alterations associated with poor prognosis, generating hypotheses for therapeutic targeting [64].

Rare disease and complex disease prediction

Rare diseases and complex diseases with heterogeneous etiologies present unique challenges for multi-omics integration due to limited sample sizes and incomplete understanding of underlying mechanisms. XAI methods can help by providing transparent justifications that support clinical decision-making even when data are scarce.

In Alzheimer's disease, interpretable ML combined with multi-omics systems biology has been used to identify personalized biomarkers and drug repurposing opportunities [95]. Cross-modal transformer reconstruction methods have been developed to integrate unpaired multi-omics data for Alzheimer's diagnosis, providing interpretability through attention mechanisms [96]. In glioblastoma, integrative multi-omics analysis using XAI identified AEBP1 and EFEMP2 as key regulators of immune heterogeneity and therapeutic response [83].

For complex diseases such as chronic obstructive pulmonary disease (COPD), transformer-based multimodal fusion has been applied to combine omics with physiological and biochemical indicators, with attention weights indicating which features drive classification [97]. These applications demonstrate that XAI can help distill meaningful signals from multi-omics data even in complex, heterogeneous disease contexts.

Single-cell and spatial omics applications

Single-cell multi-omics technologies have revolutionized our ability to profile molecular heterogeneity at cellular resolution, but they introduce additional complexity in data integration and interpretation. XAI methods are increasingly being adapted to meet these challenges.

Deep learning methods for single-cell multi-omics data integration have been reviewed, highlighting the role of VAEs, GNNs, and attention mechanisms in achieving interpretable integration [34, 92]. These methods enable the identification of cell type-specific biomarkers and the reconstruction of regulatory networks.

While spatial omics technologies are not the primary focus of this review, the principles of XAI applied to multi-omics are extendable to spatial contexts. Future developments will likely focus on integrating spatial information with molecular layers to provide interpretable maps of cellular ecosystems [78, 98, 99]. Table 4 provides representative studies in single-cell integration.

Beyond these cancer-focused applications, multi-omics and ML have also been applied to the study of intestinal ischemia/reperfusion injury [103], while attention-based deep learning has been used for protein functional annotation by integrating Gene Ontology inter-relationships [104].

Discussion

The ultimate aspiration of XAI for multi-omics integration is to improve patient outcomes through clinically deployed decision-support systems. Yet the path from research to practice is fraught with technical, organizational, and regulatory obstacles. This section examines the key considerations for translating multi-omics XAI models into clinical settings, including model development, clinician trust, workflow integration, real-time decision support, regulatory frameworks, adoption barriers, and translational readiness criteria.

From model development to clinical implementation

Translating a research model into a clinically deployable tool requires a paradigm shift from optimizing predictive accuracy to ensuring safety, usability, and sustainability. The process typically begins with retrospective validation on curated cohorts, but clinical implementation demands prospective evaluation in real-world settings [34, 60]. This includes assessing model performance across diverse patient populations, handling missing data encountered in routine care, and integrating with existing information technology infrastructure [29, 35].

A critical early step is defining the intended use case: whether the model will serve as a screening tool, a diagnostic aid, a prognostic marker, or a treatment selection guide. Each use case carries different performance requirements and regulatory pathways [18]. For example, a model used to identify patients for clinical trial enrollment may tolerate lower specificity than one used to recommend high-risk treatments. Establishing clear clinical utility—evidence that the model improves patient outcomes or reduces costs—is essential but rarely achieved in academic studies [34]. Clinical deployment should also include mechanisms for continuous performance monitoring, recalibration, and periodic updating to address data drift, evolving patient populations, and changing clinical practices throughout the AI lifecycle.

Role of explainability in clinician trust

Clinician trust is a prerequisite for adoption, and explainability is a key driver of trust [32, 33]. When physicians understand why a model makes a recommendation, they are more likely to integrate it into their clinical reasoning and to accept responsibility for the decision. XAI provides a mechanism for clinicians to verify that the model's reasoning aligns with their own knowledge and with the patient's context [23, 31].

However, trust is not automatically conferred by explanations. If explanations are inconsistent, overly complex, or biologically implausible, they may erode rather than build trust [19]. Moreover, explanations must be tailored to the clinical audience. Visualizations of pathway-level contributions, attention maps, or simplified feature summaries may be more useful than raw Shapley values [28, 64]. In antimicrobial stewardship, for instance, explainable models that highlight resistance genes and pharmacokinetic parameters have been shown to support clinician acceptance [105]. Systematic reviews indicate that explainability can increase trust when explanations are accurate, timely, and presented in a user-friendly manner, but poorly designed explanations can have the opposite effect [33].

Table 4. Representative applications in cancer, prognosis, and drug response.

Application domain	Representative studies	Methods	Omics layers	XAI contribution	Clinical relevance	References
Cancer subtype classification	AMMO; DeepMoIC; MoAGNN	Attention-fusion; GCNs; hierarchical GNN	Genomics, transcriptomics, clinical	Attention weights identify key genes and pathways driving subtypes	Guides treatment selection; identifies targetable subtypes	[63, 86, 100]
Biomarker discovery	DeepKEGG; SHAP-based analyses	Pathway-informed DL; explainable GNNs	Multi-omics (varies)	Feature attribution prioritizes genes, pathways for validation	Novel therapeutic targets; diagnostic signatures	[53, 71, 72]
Drug response prediction	DeepFusionCDR; DeepInsight-3D	Transformers; CNNs; DL	Genomics, transcriptomics, proteomics	SHAP, attention identify resistance/sensitivity mechanisms	Treatment selection; avoidance of ineffective therapies	[90, 91, 101]
Survival/Prognosis	Autosurv; CoFormerSurv; Pathformer	Interpretable DL; collaborative transformers; pathway-informed	Multi-omics + clinical	Time-dependent feature importance; pathway-level explanations	Risk stratification; treatment intensity decisions	[55, 56, 84]
Immunotherapy response	Immunotherapy survival XAI; Blood-cancer multi-omics	Explainable ML; SHAP	Multi-omics, immune signatures	Identifies immune-related gene signatures and TMB contributions	Patient selection for checkpoint inhibitors	[70, 89]
Single-cell integration	FactVAE	Factorized VAEs; multimodal DL	Single-cell multi-omics	Disentangled latent factors; cell-type specific explanations	Cellular heterogeneity; developmental trajectories	[34, 102]

XAI: explainable artificial intelligence; GCNs: graph convolutional networks; GNNs: graph neural networks; SHAP: SHapley Additive exPlanations; DL: deep learning; ML: machine learning; VAEs: variational autoencoders.

Excessively detailed or computationally intensive explanations may inadvertently increase clinician cognitive burden, particularly in time-sensitive clinical environments. Therefore, explanation interfaces should prioritize concise, clinically relevant summaries that support rapid interpretation without overwhelming end users.

Integration with EHRs and clinical workflows

For a multi-omics AI model to be used at the point of care, it must be integrated with EHRs and clinical workflows. This requires interoperability with hospital information systems, secure data transfer, and minimal disruption to existing practices [29, 35]. Many institutions are exploring real-time integration of omics data (e.g., genomic sequencing results) into EHRs, but challenges remain in standardizing data formats, ensuring data quality, and enabling automated model invocation [106].

Workflow integration also involves determining when and how model outputs are presented. A model that provides a prognostic score upon diagnosis may be deployed as a discrete alert, a clinical decision support panel, or a report integrated into the patient chart. Each approach has implications for user acceptance and cognitive load [32]. Successful implementation requires co-design with clinicians, iterative testing, and attention to how the model fits into existing decision-making processes [31].

Real-time or near-real-time decision support

Some clinical applications demand real-time or near-real-time predictions. For example, predicting acute deterioration, surgical complications, or immediate drug–drug interactions requires models that can process incoming data within seconds to minutes. For multi-omics models, real-time inference is complicated by the time required to generate omics measurements (e.g., sequencing or mass spectrometry), which can take hours to days [36]. Therefore, many multi-omics XAI systems are currently designed for pre-treatment planning or retrospective risk stratification rather than acute decision support.

Where real-time support is needed, strategies include using proxies for omics data (e.g., rapid gene expression assays), pre-computing model predictions for common scenarios, or deploying models that can operate on incomplete data while providing uncertainty estimates [106, 107]. Advances in portable sequencing and mass spectrometry may eventually enable same-day omics profiling, making real-time multi-omics AI more feasible [36].

The computational overhead associated with generating explanations, particularly for SHAP-based or attention-intensive models, may limit deployment in real-time settings. Efficient approximation strategies and lightweight explanation frameworks will therefore be essential for time-critical clinical applications.

Regulatory and ethical considerations

Regulatory approval is a major gatekeeper for clinical AI deployment. In the United States, the FDA has issued guidance on software as a medical device (SaMD), emphasizing the need for transparency, validation, and post-market surveillance. The European Union's Medical Device Regulation and the proposed Artificial Intelligence Act similarly require that high-risk AI systems be interpretable and provide explanations for their outputs [18, 34].

For multi-omics XAI models, regulatory bodies are likely to require evidence of generalizability across populations, robustness to input variability, and validation of explanation faithfulness. The lack of standardized XAI evaluation frameworks complicates this process [28]. Moreover, the dynamic nature of AI models—where continuous learning may change behavior—poses challenges for traditional approval pathways [34].

Ethical considerations extend beyond regulation to include informed consent, data privacy, and equitable access. Multi-omics data are uniquely sensitive; patients must be informed about how their data will be used and protected [60]. Models must be audited for bias across demographic groups, and deployment plans must address disparities in healthcare access that could exacerbate existing inequities [32].

Emerging regulatory frameworks increasingly emphasize ethical AI principles including transparency, traceability, non-discrimination, human oversight, and explainability. In clinical environments, explainability is not only a technical requirement but also an ethical necessity that enables informed clinical judgment and patient-centered care. Future multi-omics AI systems should therefore incorporate ethics-by-design principles throughout model development and deployment.

Barriers to adoption in hospitals and laboratories

Even when models are technically sound and regulatory approvals are obtained, adoption in hospitals and laboratories faces numerous barriers. Infrastructure limitations—including lack of computational resources, data storage, and trained personnel—can preclude deployment in smaller or resource-limited institutions [34]. Hospital IT departments may be reluctant to integrate external AI tools due to security concerns, vendor compatibility issues, and maintenance burdens [35].

Laboratory adoption of multi-omics workflows is similarly constrained by the need for standardized protocols, quality control, and turnaround times. Clinical laboratories may be hesitant to implement new assays that feed into AI models without clear evidence of cost-effectiveness and reimbursement pathways [94]. Moreover, the rapid evolution of AI technologies can conflict with the slower pace of clinical validation and adoption cycles, leading to "pilot purgatory" where promising tools never scale [36].

Requirements for translational readiness

To bridge the gap between research and clinical practice, multi-omics XAI systems must meet a set of translational readiness criteria [34]. These include:

Prospective validation on multi-center, real-world cohorts that reflect the target population, with predefined performance thresholds and prespecified subgroup analyses [60].

Interoperability with existing EHR and laboratory information systems through standard data formats [e.g., Health Level Seven Fast Healthcare Interoperability Resources (HL7 FHIR)] and secure application programming interfaces [29].

Usability and human factors engineering, ensuring that explanations are presented in a format that clinicians can quickly understand and act upon without adding excessive cognitive burden [32].

Scalability to handle increasing volumes of data and users, with cloud or on-premises architectures that maintain performance under peak loads [35].

Sustainability including clear funding models for maintenance, updates, and technical support, as well as pathways for reimbursement of AI-guided services [34].

Continuous monitoring for performance drift, bias, and safety issues, with mechanisms for updating models, when necessary, without violating regulatory requirements [57].

Ethical and legal frameworks that address data ownership, consent, liability, and equity.

Achieving these requirements demands close collaboration among computational scientists, clinicians, hospital administrators, regulators, and patients. Multi-stakeholder consortia and public–private partnerships have begun to address these challenges, but widespread clinical deployment of explainable multi-omics AI remains in its infancy. Future progress will depend on sustained investment, pragmatic evaluation frameworks, and a shared commitment to patient-centered innovation [12, 36].

The key stages required to move a multi-omics XAI model from research to clinical practice are summarized in Figure 5, which outlines the pipeline from retrospective validation through regulatory approval and post-market surveillance.

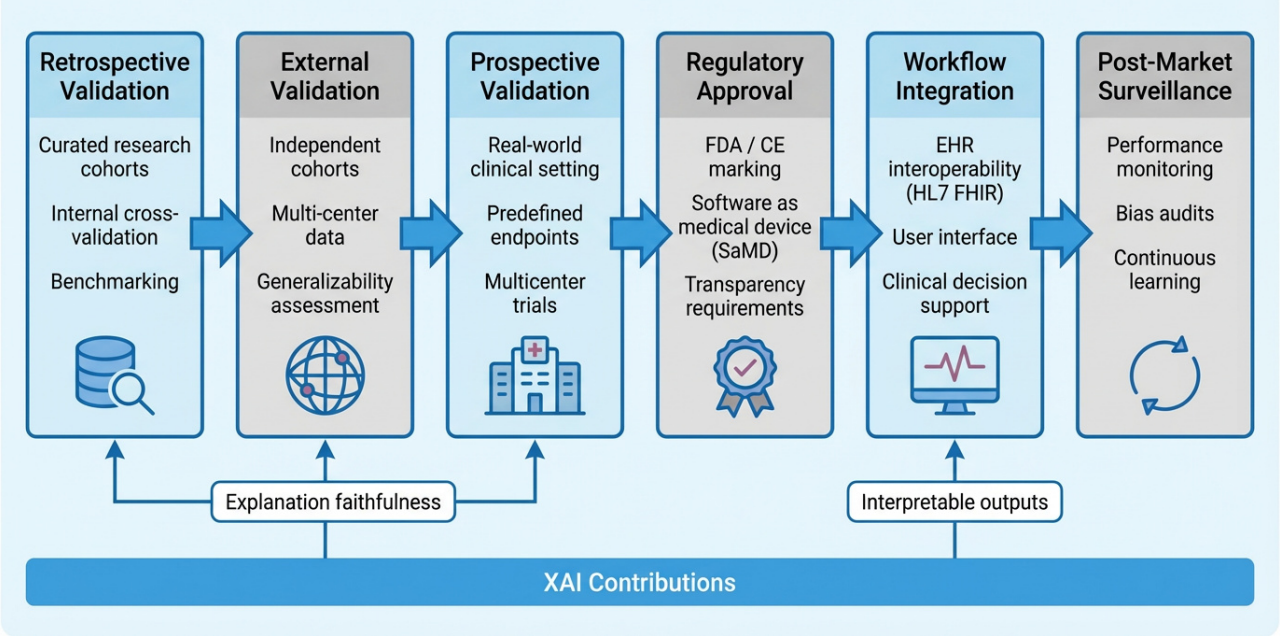

Figure 5. Clinical translation pipeline for explainable multi-omics AI systems, from retrospective validation to deployment, regulatory approval, and post-market monitoring. XAI: explainable artificial intelligence; EHR: electronic health records.

Future directions

As the fields of multi-omics integration and XAI continue to evolve, emerging paradigms promise to address current limitations while opening new frontiers for precision medicine. Table 5 summarizes these future directions alongside the challenges they aim to resolve.

Table 5. Key challenges and future research directions.

Challenge category	Specific challenge	Current limitation	Future direction	Key references
Data characteristics	High dimensionality	Overfitting; unstable feature selection; curse of dimensionality	Self-supervised learning; foundation models; synthetic data generation	[6, 8]
	Data heterogeneity	Integration difficulty; modality-specific characteristics lost	Advanced fusion strategies; modality-specific encoders with shared representations	[9, 40]
	Batch effects	Confounded models; poor generalization	Standardized preprocessing; robust normalization; batch-invariant representations	[14, 75]
	Missing modalities	Incomplete data; biased imputation	Masked modeling; generative imputation; modality-agnostic architectures	[15, 53]
Methodological	Interpretability vs. performance trade-off	Black-box models outperform interpretable ones	Hybrid models; intrinsically interpretable architectures; attention mechanisms	[19, 22]
	Explanation faithfulness	Post-hoc explanations may not reflect model behavior	Faithfulness benchmarks; causal explanation methods; model-specific XAI	[21, 45]
	Scalability	Computational cost of training and explaining large models	Efficient attention; sparse architectures; federated learning	[108, 109]
Evaluation	Lack of standardized benchmarks	Inconsistent evaluation; incomparable results	Community benchmarks (e.g., MOB); standardized metrics for XAI	[42, 50]
	Reproducibility	Variability in preprocessing, splits, seeds	Open code; shared pipelines; reproducible workflows	[14, 73]
	Clinical validation	Few prospective studies; limited external validation	Multi-center trials; real-world evidence; regulatory pathways	[34, 60]
Translational	Workflow integration	Models not integrated with EHR/clinical systems	Interoperability standards (HL7 FHIR); user-centered design	[29, 35, 97]
	Regulatory approval	Unclear pathways for AI as medical device	Engage regulators early; validation frameworks; post-market surveillance	[18]
	Ethical and fairness concerns	Bias in training data; privacy risks	Federated learning; fairness audits; inclusive cohort design	[32, 60]
Emerging frontiers	Causal inference	Correlation-based predictions limit mechanistic insight	Causal AI; Mendelian randomization; interventional predictions	[14, 110]
	Foundation models	Limited pre-trained models for multi-omics	Multi-modal foundation models; transfer learning; open-source models	[108, 111]
	Digital twins	No integrated patient simulations	Multi-modal digital twins; generative models; personalized simulations	[112, 113]
	Human-in-the-loop	AI decisions without clinician oversight	Interactive XAI; clinician feedback loops; collaborative AI	[32, 33]

XAI: explainable artificial intelligence.

Federated learning for privacy-preserving multi-omics analysis

Centralized data aggregation raises significant privacy concerns. Federated learning enables model training across distributed institutions without exchanging raw data, addressing privacy regulations while enabling larger-scale analyses [108, 114]. For multi-omics integration, federated learning must handle heterogeneous feature spaces and missing modalities across sites. Future work should focus on developing federated XAI frameworks that provide faithful explanations without compromising privacy.

Causal AI for biological inference

Current multi-omics AI models predominantly capture correlations, whereas clinical decision-making increasingly requires causal evidence. Causal AI aims to infer cause–effect relationships from observational data and support reasoning about interventions [110, 115]. Mendelian randomization uses genetic variants as instrumental variables to investigate causal relationships [73, 74]. Recent disease-focused studies have

combined Mendelian randomization with multi-omics, ML, and SHAP to identify causally associated metabolites and interpretable molecular signatures [116].

Foundation models in biomedical omics

Foundation models—large-scale pre-trained models adaptable to diverse downstream tasks—are transforming AI across domains [108, 111]. For multi-omics integration, foundation models offer the potential to overcome small sample size problems by leveraging pre-trained representations that encode rich biological knowledge. Future directions include developing truly multimodal foundation models that jointly process genomics, transcriptomics, proteomics, and clinical data, with built-in interpretability mechanisms.

Multi-modal digital twins in precision medicine

Digital twins—virtual representations of individual patients integrating multi-scale data—represent a frontier in precision medicine [113]. A patient's digital twin would incorporate multi-omics, clinical, and longitudinal data, enabling personalized simulations of therapeutic interventions. XAI is essential for digital twins to provide interpretable predictions and allow clinicians to interrogate model behavior. Future work must address computational demands, standardized interfaces, and validation against real-world outcomes.

Self-supervised and weakly supervised learning

The reliance on large, well-annotated datasets is a major bottleneck. Self-supervised learning enables models to learn representations from unlabeled data by solving pretext tasks such as masked reconstruction [79]. Weakly supervised learning uses noisy, incomplete labels. For multi-omics, these approaches can leverage vast amounts of unlabeled data and incorporate diverse evidence sources. Future directions include developing self-supervised objectives tailored to multi-omics data and ensuring learned features are biologically interpretable.

Standardized explainability benchmarks

The proliferation of XAI methods has outpaced the development of rigorous evaluation frameworks, making it difficult to compare methods or trust their outputs [28]. Standardized benchmarks are urgently needed to assess explanation faithfulness, stability, and biological validity. Future initiatives should develop common metrics—faithfulness, stability, biological alignment, computational cost—and establish community challenges (e.g., DREAM-style) specifically for XAI in multi-omics.

Human-in-the-Loop AI for clinical decision-making

The integration of AI into clinical workflows should be conceived as a partnership between human experts and machines. Human-in-the-loop AI systems augment clinician capabilities, with explanations serving as a medium for collaboration [32]. Clinicians can validate explanations against patient context, provide feedback, and retain final decision authority. Future directions include developing intuitive user interfaces that present explanations in clinically meaningful ways and designing workflows for efficient clinician–AI interaction.

The convergence of multi-omics technologies and XAI marks a decisive shift toward truly interpretable and data-driven precision medicine. By integrating heterogeneous molecular layers with transparent computational frameworks, these approaches enable the discovery of complex biological relationships while providing clinically meaningful explanations for model predictions. Such capabilities not only improve disease stratification, biomarker identification, and therapeutic response prediction, but also strengthen trust, accountability, and regulatory acceptance of AI-driven medical tools. Nevertheless, the path to clinical implementation remains constrained by persistent challenges, including data heterogeneity, limited cohort sizes, reproducibility concerns, and insufficient real-world validation. Overcoming these barriers will require standardized benchmarks, robust validation across diverse populations, and seamless integration with clinical workflows. Ultimately, the next generation of precision medicine will depend on

the development of scalable, interpretable, and clinically reliable AI systems that transform complex multi-omics data into actionable knowledge, thereby enabling more accurate diagnoses, personalized treatments, and improved patient outcomes.

Abbreviations

AI: Artificial Intelligence

EHR: electronic health record

GNNs: graph neural networks

LIME: Local Interpretable Model-agnostic Explanations

ML: machine learning

RFs: random forests

SHAP: SHapley Additive exPlanations

TCGA: The Cancer Genome Atlas

XAI: explainable artificial intelligence

Declarations

Acknowledgments

During the preparation of this work, the authors used Google AI tools to improve language and readability. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the published article.

Author contributions

MH: Conceptualization, Writing—original draft, Visualization. SH: Writing—review & editing, Supervision. KA: Methodology, Writing—review & editing, Formal analysis. MW: Methodology, Formal analysis, Writing—review & editing. All authors have read and approved the final version of the manuscript.

Conflicts of interest

The authors declare that they have no conflicts of interest.

Ethical approval

Not applicable.

Consent to participate

Not applicable.

Consent to publication

Not applicable.

Availability of data and materials

No new data were generated or analyzed in this review. All referenced datasets are publicly available from the sources cited.

Funding

Not applicable.

Publisher's note

Open Exploration maintains a neutral stance on jurisdictional claims in published institutional affiliations and maps. All opinions expressed in this article are the personal views of the author(s) and do not represent the stance of the editorial team or the publisher.

References

1. Johnson KB, Wei WQ, Weeraratne D, Frisse ME, Misulis K, Rhee K, et al. Precision Medicine, AI, and the Future of Personalized Health Care. Clin Transl Sci. 2021;14:86–93. [DOI] [PubMed] [PMC]
2. MacEachern SJ, Forkert ND. Machine learning for precision medicine. Genome. 2021;64:416–25. [DOI] [PubMed]
3. Bhinder B, Gilvary C, Madhukar NS, Elemento O. Artificial Intelligence in Cancer Research and Precision Medicine. Cancer Discov. 2021;11:900–15. [DOI] [PubMed] [PMC]
4. Liao J, Li X, Gan Y, Han S, Rong P, Wang W, et al. Artificial intelligence assists precision medicine in cancer treatment. Front Oncol. 2023;12:998222. [DOI] [PubMed] [PMC]
5. Kaur P, Singh A, Chana I. Computational Techniques and Tools for Omics Data Analysis: State-of-the-Art, Challenges, and Future Directions. Arch Comput Methods Eng. 2021;28:4595–631. [DOI]
6. Feldner-Busztin D, Firbas Nisantzis P, Edmunds SJ, Boza G, Racimo F, Gopalakrishnan S, et al. Dealing with dimensionality: the application of machine learning to multi-omics data. Bioinformatics. 2023; 39:btad021. [DOI] [PubMed] [PMC]
7. Johnson EM. Principled Analysis of High-Dimensional Data: How to Identify Signals while Under the Curse of Dimensionality [dissertation]. Evanston: Northwestern University; 2022.
8. Zaikin A, Sviridov I, Sosedka A, Linich A, Nasyrov R, Mirkes EM, et al. Overcoming the Curse of Dimensionality with Synolitic AI. Technologies. 2026;14:84. [DOI]
9. Baião AR, Cai Z, Poulos RC, Robinson PJ, Reddel RR, Zhong Q, et al. A technical review of multi-omics data integration methods: from classical statistical to deep generative approaches. Brief Bioinform. 2025;26:bbaf355. [DOI] [PubMed] [PMC]
10. Ballard JL, Wang Z, Li W, Shen L, Long Q. Deep learning-based approaches for multi-omics data integration and analysis. BioData Min. 2024;17:38. [DOI] [PubMed] [PMC]
11. Kang M, Ko E, Mersha TB. A roadmap for multi-omics data integration using deep learning. Brief Bioinform. 2021;23:bbaf355. [DOI] [PubMed] [PMC]
12. Nam Y, Kim J, Jung SH, Woerner J, Suh EH, Lee DG, et al. Harnessing Artificial Intelligence in Multimodal Omics Data Integration: Paving the Path for the Next Frontier in Precision Medicine. Annu Rev Biomed Data Sci. 2024;7:225–50. [DOI] [PubMed] [PMC]
13. Wekesa JS, Kimwele M. A review of multi-omics data integration through deep learning approaches for disease diagnosis, prognosis, and treatment. Front Genet. 2023;14:1199087. [DOI] [PubMed] [PMC]
14. Graw S, Chappell K, Washam CL, Gies A, Bird J, Robeson MS 2nd, et al. Multi-omics data integration considerations and study design for biological systems and disease. Mol Omics. 2021;17:170–85. [DOI] [PubMed] [PMC]
15. Jan R, Hussain A, Assad A, Khurshid S, Macha MA. Challenges with multi-omics data integration. In: Multi-omics technology in human health and diseases. Elsevier; 2025. pp. 223–42. [DOI]
16. Mirza B, Wang W, Wang J, Choi H, Chung NC, Ping P. Machine Learning and Integrative Analysis of Biomedical Big Data. Genes. 2019;10:87. [DOI] [PubMed] [PMC]
17. Yu Y, Mai Y, Zheng Y, Shi L. Assessing and mitigating batch effects in large-scale omics studies. Genome Biol. 2024;25:254. [DOI] [PubMed] [PMC]
18. Ennab M, Mcheick H. Enhancing interpretability and accuracy of AI models in healthcare: a comprehensive review on challenges and future directions. Front Robot AI. 2024;11:1444763. [DOI] [PubMed] [PMC]

19. Ghassemi M, Oakden-Rayner L, Beam AL. The false hope of current approaches to explainable artificial intelligence in health care. Lancet Digit Health. 2021;3:e745–50. [DOI] [PubMed]
20. Hossain MI, Zamzmi G, Mouton PR, Salekin MS, Sun Y, Goldgof D. Explainable AI for Medical Data: Current Methods, Limitations, and Future Directions. ACM Comput Surv. 2025;57:1–46. [DOI]
21. Bilodeau B, Jaques N, Koh PW, Kim B. Impossibility theorems for feature attribution. Proc Natl Acad Sci. 2024;121:e2304406120. [DOI] [PubMed] [PMC]
22. Rudin C, Chen C, Chen Z, Huang H, Semenova L, Zhong C. Interpretable machine learning: Fundamental principles and 10 grand challenges. Stat Surv. 2022;16:e16. [DOI]
23. Allgaier J, Mulansky L, Draelos RL, Pryss R. How does the model make predictions? A systematic literature review on the explainability power of machine learning in healthcare. Artif Intell Med. 2023;143:102616. [DOI] [PubMed]
24. Bharati S, Mondal MRH, Podder P. A Review on Explainable Artificial Intelligence for Healthcare: Why, How, and When? IEEE Trans Artif Intell. 2023;5:1429–42. [DOI]
25. Hulsen T. Explainable Artificial Intelligence (XAI): Concepts and Challenges in Healthcare. AI. 2023;4: 652–66. [DOI]
26. Mienye ID, Obaido G, Jere N, Mienye E, Aruleba K, Emmanuel ID, et al. A survey of explainable artificial intelligence in healthcare: Concepts, applications, and challenges. Inform Med Unlocked. 2024;51:101587. [DOI]
27. Novakovsky G, Dexter N, Libbrecht MW, Wasserman WW, Mostafavi S. Obtaining genetics insights from deep learning via explainable artificial intelligence. Nat Rev Genet. 2023;24:125–37. [DOI] [PubMed]
28. Toussaint PA, Leiser F, Thiebes S, Schlesner M, Brors B, Sunyaev A. Explainable artificial intelligence for omics data: a systematic mapping study. Brief Bioinform. 2023;25:bbad453. [DOI] [PubMed] [PMC]
29. Tong L, Shi W, Isgut M, Zhong Y, Lais P, Gloster L, et al. Integrating Multi-Omics Data With EHR for Precision Medicine Using Advanced Artificial Intelligence. IEEE Rev Biomed Eng. 2024;17:80–97. [DOI] [PubMed]
30. Hussein S, Ramos V, Liu W, Kechris KJ, Lange L, Bowler RP, et al. Learning from multi-omics networks to enhance disease prediction: an optimized network embedding and fusion approach. In: 2024 IEEE International Conference on Bioinformatics and Biomedicine (BIBM); 2024 Dec 3–6; Lisbon, Portugal. IEEE; 2024. pp. 4371–8. [DOI]
31. Noor AA, Manzoor A, Mazhar Qureshi MD, Qureshi MA, Rashwan W. Unveiling Explainable AI in Healthcare: Current Trends, Challenges, and Future Directions. WIREs Data Min Knowl Discov. 2025; 15:e70018. [DOI]
32. MacDonald S, Steven K, Trzaskowski M. Interpretable AI in healthcare: Enhancing fairness, safety, and trust. In: Artificial Intelligence in Medicine: Applications, Limitations and Future Directions. Springer; 2022. pp. 241–58. [DOI]
33. Rosenbacke R, Melhus Å, McKee M, Stuckler D. How Explainable Artificial Intelligence Can Increase or Decrease Clinicians' Trust in AI Applications in Health Care: Systematic Review. JMIR AI. 2024;3: e53207. [DOI] [PubMed] [PMC]
34. Carini C, Seyhan AA. Tribulations and future opportunities for artificial intelligence in precision medicine. J Transl Med. 2024;22:411. [DOI] [PubMed] [PMC]
35. Chen YM, Hsiao TH, Lin CH, Fann YC. Unlocking precision medicine: clinical applications of integrating health records, genetics, and immunology through artificial intelligence. J Biomed Sci. 2025;32:16. [DOI] [PubMed] [PMC]
36. Hsu CY, Askar S, Alshkarchy SS, Nayak PP, Attabi KAL, Khan MA, et al. AI-driven multi-omics integration in precision oncology: bridging the data deluge to clinical decisions. Clin Exp Med. 2025; 26:29. [DOI] [PubMed] [PMC]

37. Athaya T, Ripan RC, Li X, Hu H. Multimodal deep learning approaches for single-cell multi-omics data integration. Brief Bioinform. 2023;24:bbad313. [DOI] [PubMed] [PMC]
38. Baysoy A, Bai Z, Satija R, Fan R. The technological landscape and applications of single-cell multi-omics. Nat Rev Mol Cell Biol. 2023;24:695–713. [DOI] [PubMed] [PMC]
39. Sanches PHG, de Melo NC, Porcari AM, de Carvalho LM. Integrating Molecular Perspectives: Strategies for Comprehensive Multi-Omics Integrative Data Analysis and Machine Learning Applications in Transcriptomics, Proteomics, and Metabolomics. Biology. 2024;13:848. [DOI] [PubMed] [PMC]
40. Sussulini A, Xia J, Orešič M. Editorial: Multi-omics: Trends and applications in clinical research. Front Mol Biosci. 2022;9:994239. [DOI] [PubMed] [PMC]
41. Martínez-García M, Hernández-Lemus E. Data Integration Challenges for Machine Learning in Precision Medicine. Front Med. 2022;8:784455. [DOI] [PubMed] [PMC]
42. Leng D, Zheng L, Wen Y, Zhang Y, Wu L, Wang J, et al. A benchmark study of deep learning-based multi-omics data fusion methods for cancer. Genome Biol. 2022;23:171. [DOI] [PubMed] [PMC]
43. Liu Z, Park T. DMOIT: denoised multi-omics integration approach based on transformer multi-head self-attention mechanism. Front Genet. 2024;15:1488683. [DOI] [PubMed] [PMC]
44. Labarga A. An Interpretable Graph Neural Network for Multi-omics Data Integration and Biomarker Discovery. In: Lecture Notes in Computer Science. Springer; 2026. pp. 438–48. [DOI]
45. Reynolds J, Pan C. Benchmarking interpretability of deep learning for predictive genomics: Recall, precision, and variability of feature attribution. PLOS Comput Biol. 2025;21:e1013784. [DOI] [PubMed] [PMC]
46. Choi SR, Lee M. Transformer Architecture and Attention Mechanisms in Genome Data Analysis: A Comprehensive Review. Biology. 2023;12:1033. [DOI] [PubMed] [PMC]
47. Stahlschmidt SR, Ulfenborg B, Synnergren J. Multimodal deep learning for biomedical data fusion: a review. Brief Bioinform. 2022;23:bbab569. [DOI] [PubMed] [PMC]
48. Zhao C, Su KJ, Wu C, Cao X, Sha Q, Li W, et al. Multi-scale variational autoencoder for imputation of missing values in untargeted metabolomics using whole-genome sequencing data. Comput Biol Med. 2024;179:108813. [DOI] [PubMed] [PMC]
49. Cascianelli S, Galzerano A, Masseroli M. Supervised Relevance-Redundancy assessments for feature selection in omics-based classification scenarios. J Biomed Inform. 2023;144:104457. [DOI] [PubMed]
50. Duan R, Gao L, Gao Y, Hu Y, Xu H, Huang M, et al. Evaluation and comparison of multi-omics data integration methods for cancer subtyping. PLOS Comput Biol. 2021;17:e1009224. [DOI] [PubMed] [PMC]
51. Qi L, Wang W, Wu T, Zhu L, He L, Wang X. Multi-Omics Data Fusion for Cancer Molecular Subtyping Using Sparse Canonical Correlation Analysis. Front Genet. 2021;12:607817. [DOI] [PubMed] [PMC]
52. Ma T, Wang J. GraphPath: a graph attention model for molecular stratification with interpretability based on the pathway–pathway interaction network. Bioinformatics. 2024;40:btae165. [DOI] [PubMed] [PMC]
53. Tripathy RK, Frohock Z, Wang H, Cary GA, Keegan S, Carter GW, et al. Effective integration of multi-omics with prior knowledge to identify biomarkers via explainable graph neural networks. NPJ Syst Biol Appl. 2025;11:43. [DOI] [PubMed] [PMC]
54. Valous NA, Popp F, Zörnig I, Jäger D, Charoentong P. Graph machine learning for integrated multi-omics analysis. Br J Cancer. 2024;131:205–11. [DOI] [PubMed] [PMC]
55. Liu X, Tao Y, Cai Z, Bao P, Ma H, Li K, et al. Pathformer: a biological pathway informed transformer for disease diagnosis and prognosis using multi-omics data. Bioinformatics. 2024;40:btae316. [DOI] [PubMed] [PMC]
56. Wen G, Li L. CoFormerSurv: Collaborative transformer for multi-omics survival analysis. PLOS Comput Biol. 2026;22:e1013875. [DOI] [PubMed] [PMC]

57. Ennab M, Mcheick H. Designing an Interpretability-Based Model to Explain the Artificial Intelligence Algorithms in Healthcare. Diagnostics. 2022;12:1557. [DOI] [PubMed] [PMC]
58. Peloso A, Naesens M, Thaunat O. The Dawn of a New Era in Kidney Transplantation: Promises and Limitations of Artificial Intelligence for Precision Diagnostics. Transpl Int. 2023;36:12010. [DOI] [PubMed] [PMC]
59. Santorsola M, Lescai F. The promise of explainable deep learning for omics data analysis: Adding new discovery tools to AI. New Biotechnol. 2023;77:1–11. [DOI] [PubMed]
60. Corti C, Cobanaj M, Dee EC, Criscitiello C, Tolaney SM, Celi LA, et al. Artificial intelligence in cancer research and precision medicine: Applications, limitations and priorities to drive transformation in the delivery of equitable and unbiased care. Cancer Treat Rev. 2023;112:102498. [DOI] [PubMed]
61. Benkirane H, Vakalopoulou M, Planchard D, Adam J, Olaussen K, Michiels S, et al. Multimodal CustOmics: A unified and interpretable multi-task deep learning framework for multimodal integrative data analysis in oncology. PLOS Comput Biol. 2025;21:e1013012. [DOI] [PubMed] [PMC]
62. Onoja A. An integrated interpretable machine learning framework for high-dimensional multi-omics datasets [dissertation]. SCUOLA NORMALE SUPERIORE DI PISA; 2023.
63. Li H, Zhou Y, Zhao N, Wang Y, Lai Y, Zeng F, et al. ISMI-VAE: A deep learning model for classifying disease cells using gene expression and SNV data. Comput Biol Med. 2024;175:108485. [DOI] [PubMed]
64. Sureshkumar S, Sureshkumar N. A SHAP-based Interpretation for Leveraging MultiOmic Data and Pancreatic Cancer Prognostication using Deep Learning Techniques. In: Proceedings of 2024 5th International Conference on Smart Electronics and Communication (ICOSEC); 2024 Sep 18–20; Trichy, India. IEEE; 2024, pp. 1640–4. [DOI]
65. Gabbay F, Bar-Lev S, Montano O, Hadad N. A LIME-Based Explainable Machine Learning Model for Predicting the Severity Level of COVID-19 Diagnosed Patients. Appl Sci. 2021;11:10417. [DOI]
66. Mehdi Chouit E, Rachdi M, Bellafkih M, Raouyane B. Interpretable machine learning for chronic kidney disease prediction: Insights from SHAP and LIME analyses. PLOS One. 2026;21:e0343205. [DOI]
67. Benkirane H, Pradat Y, Michiels S, Cournède PH. CustOmics: A versatile deep-learning based strategy for multi-omics integration. PLOS Comput Biol. 2023;19:e1010921. [DOI] [PubMed] [PMC]
68. Ukwuoma CC, Cai D, Eziefuna EO, Oluwasanmi A, Abdi SF, Muoka GW, et al. Enhancing histopathological medical image classification for Early cancer diagnosis using deep learning and explainable AI – LIME & SHAP. Biomed Signal Process Control. 2025;100:107014. [DOI]
69. Grešová K, Vaculík O, Alexiou P. Using Attribution Sequence Alignment to Interpret Deep Learning Models for miRNA Binding Site Prediction. Biology. 2023;12:369. [DOI] [PubMed] [PMC]
70. Hounye AH, Xiong L, Hou M. Integrated explainable machine learning and multi-omics analysis for survival prediction in cancer with immunotherapy response. Apoptosis. 2025;30:364–88. [DOI] [PubMed]
71. Chen C, Weng Y. P1305 Explore key genes of Crohn's disease based on glycerophospholipid metabolism: A comprehensive analysis utilizing mendelian randomization, multi-omics integration, machine learning, and SHAP methodology. J Crohn's Colitis. 2026;20:jjaf231.148. [DOI]
72. Lan W, Liao H, Chen Q, Zhu L, Pan Y, Chen YP. DeepKEGG: a multi-omics data integration framework with biological insights for cancer recurrence prediction and biomarker discovery. Brief Bioinform. 2024;25:bbae185. [DOI] [PubMed] [PMC]
73. Davey Smith G, Hemani G. Mendelian randomization: genetic anchors for causal inference in epidemiological studies. Hum Mol Genet. 2014;23:R89–98. [DOI] [PubMed] [PMC]
74. Lawlor DA, Harbord RM, Sterne JAC, Timpson N, Davey Smith G. Mendelian randomization: Using genes as instruments for making causal inferences in epidemiology. Stat Med. 2008;27:1133–63. [DOI] [PubMed]

75. Jiang Z, Zhang H, Gao Y, Sun Y. Multi-omics strategies for biomarker discovery and application in personalized oncology. Mol Biomed. 2025;6:115. [DOI] [PubMed] [PMC]
76. Abdelaziz EH, Ismail R, Mabrouk MS, Amin E. Multi-omics data integration and analysis pipeline for precision medicine: Systematic review. Comput Biol Chem. 2024;113:108254. [DOI] [PubMed]
77. Xiao S, Lin H, Wang C, Wang S, Rajapakse JC. Graph Neural Networks With Multiple Prior Knowledge for Multi-Omics Data Analysis. IEEE J Biomed Health Inform. 2023;27:4591–600. [DOI] [PubMed]
78. Papadaki E, Kakkos I, Vlamos P, Petropoulou O, Miloulis ST, Palamas S, et al. Recent Web Platforms for Multi-Omics Integration Unlocking Biological Complexity. Appl Sci. 2024;15:329. [DOI]
79. Azher ZL, Suvarna A, Chen JQ, Zhang Z, Christensen BC, Salas LA, et al. Assessment of emerging pretraining strategies in interpretable multimodal deep learning for cancer prognostication. BioData Min. 2023;16:23. [DOI] [PubMed] [PMC]
80. Omran MM, Emam M, Gamaleldin M, Abushady AM, Elattar MA, El-Hadidi M. Comparative analysis of statistical and deep learning-based multi-omics integration for breast cancer subtype classification. J Transl Med. 2025;23:709. [DOI] [PubMed] [PMC]
81. Tembhare K, Sharma T, Kasibhatla SM, Achalere A, Joshi R. Multi-ensemble machine learning framework for omics data integration: A case study using breast cancer samples. Inform Med Unlocked. 2024;47:101507. [DOI]
82. Meng XY, Zhou XH, Li S, Shi MJ, Li XH, Yang BY, et al. Machine Learning-Based Detection of Bladder Cancer by Urine cfDNA Fragmentation Hotspots that Capture Cancer-Associated Molecular Features. Clin Chem. 2024;70:1463–73. [DOI] [PubMed]
83. Yin Y, Fu X, Gong S, Xie Y, Wu W, Li Z, et al. Integrative multi-omics analysis identifies AEBP1 and EFEMP2 as key regulators of immune heterogeneity and therapeutic response in glioblastoma. Discov Oncol. 2025;16:1792. [DOI] [PubMed] [PMC]
84. Jiang L, Xu C, Bai Y, Liu A, Gong Y, Wang YP, et al. Autosurv: interpretable deep learning framework for cancer survival analysis incorporating clinical and multi-omics data. NPJ Precis Oncol. 2024;8:4. [DOI] [PubMed] [PMC]
85. Tran D, Nguyen H, Pham VD, Nguyen P, Nguyen Luu H, Minh Phan L, et al. A comprehensive review of cancer survival prediction using multi-omics integration and clinical variables. Brief Bioinform. 2025;26:bbaf150. [DOI] [PubMed] [PMC]
86. Lin CP, Ho YJ, Chiu YP, Tang Y, Paik YS, Chen GT, et al. MoAGNN: a multi-omics hierarchical graph neural network for subtype classification and prognosis prediction in lung adenocarcinoma. Brief Bioinform. 2026;27:bbaf735. [DOI] [PubMed] [PMC]
87. Li Y, Shi X, Wang L, Zhang L. DeepMoDRP: A Multi-Omics-Based Deep Learning Framework for Drug Response Prediction in Brain Cancer. Mol Inform. 2026;45:e70020. [DOI] [PubMed]
88. Zhang J. Emerging Trends in Multi-Omics Data Integration: Challenges and Future Directions. Comput Mol Biol. 2024;14:64–75. [DOI]
89. Pirouzbakht M, Hamzeh S, Soleimani Samarkhazan H. Beyond single biomarkers: multi-omics strategies to predict immunotherapy outcomes in blood cancers. Clin Exp Med. 2025;25:355. [DOI] [PubMed] [PMC]
90. Hu X, Zhang P, Zhang J, Deng L. DeepFusionCDR: Employing Multi-Omics Integration and Molecule-Specific Transformers for Enhanced Prediction of Cancer Drug Responses. IEEE J Biomed Health Inform. 2024;28:6248–58. [DOI] [PubMed]
91. Malik V, Kalakoti Y, Sundar D. Deep learning assisted multi-omics integration for survival and drug-response prediction in breast cancer. BMC Genom. 2021;22:214. [DOI] [PubMed] [PMC]
92. Lin M, Guo J, Gu Z, Tang W, Tao H, You S, et al. Machine learning and multi-omics integration: advancing cardiovascular translational research and clinical practice. J Transl Med. 2025;23:388. [DOI] [PubMed] [PMC]

93. Lin B, Ma Y, Wu S. Multi-Omics and Artificial Intelligence-Guided Data Integration in Chronic Liver Disease: Prospects and Challenges for Precision Medicine. OMICS: J Integr Biol. 2022;26:415–21. [DOI] [PubMed]
94. Saliba A, Du Y, Feng T, Garmire L. Multi-Omics Integration in Nephrology: Advances, Challenges, and Future Directions. Semin Nephrol. 2024;44:151584. [DOI] [PubMed] [PMC]
95. Mottaqi M, Zhang P, Xie L. Integrating explainable AI with multiomics systems biology and EHR data mining for personalized drug repurposing in Alzheimer's disease. bioRxiv 2025.03.24.644676 [Preprint]. 2025 [cited 2026 May 20]. Available from: https://www.biorxiv.org/content/10.1101/2025.03.24.644676v2
96. Liao K, Du D, Li J, Huang J, Fan X, Chen C, et al. Interpretable integration of unpaired multi-omics for Alzheimer's diagnosis via cross-modal transformer reconstruction. PLOS Comput Biol. 2026;22:e1014074. [DOI] [PubMed] [PMC]
97. Xie W, Fang Y, Yang G, Yu K, Li W. Transformer-Based Multi-Modal Data Fusion Method for COPD Classification and Physiological and Biochemical Indicators Identification. Biomolecules. 2023;13:1391. [DOI] [PubMed] [PMC]
98. Nunes FVM, Behrens LMP, Weimer RD, Gonçalves GF, da Silva Fernandes G, Dorn M. Deep learning methods and applications in single-cell multimodal data integration. Mol Omics. 2025;21:545–65. [DOI] [PubMed]
99. Xu Z, Li W, Dong X, Chen Y, Zhang D, Wang J, et al. Precision medicine in colorectal cancer: Leveraging multi-omics, spatial omics, and artificial intelligence. Clin Chim Acta. 2024;559:119686. [DOI] [PubMed]
100. Li W, Muneer A, Waqas M, Zhou X, Wu J. Attention-Fusion Model for Multi-omics (AMMO) Data Integration in Lung Adenocarcinoma. In: International Workshop on Computational Mathematics Modeling in Cancer Analysis. Springer; 2024. pp. 52–60. [DOI]
101. Sharma A, Lysenko A, Boroevich KA, Tsunoda T. DeepInsight-3D architecture for anti-cancer drug response prediction with deep-learning on multi-omics. Sci Rep. 2023;13:2483. [DOI] [PubMed] [PMC]
102. Dai H, Yu Z, Zhao Y, Jiang K, Hang Z, Huang X, et al. Integrating machine learning models with multi-omics analysis to decipher the prognostic significance of mitotic catastrophe heterogeneity in bladder cancer. Biol Direct. 2025;20:56. [DOI] [PubMed] [PMC]
103. Chen X, Liu K. Multi-omics analysis and machine learning in the study of intestinal ischemia/reperfusion injury. Anesthesiol Perioper Sci. 2025;3:17. [DOI]
104. Pan T, Li C, Bi Y, Wang Z, Gasser RB, Purcell AW, et al. PFresGO: an attention mechanism-based deep-learning approach for protein annotation by integrating gene ontology inter-relationships. Bioinformatics. 2023;39:btad094. [DOI] [PubMed] [PMC]
105. Giacobbe DR, Marelli C, Guastavino S, Mora S, Rosso N, Signori A, et al. Explainable and Interpretable Machine Learning for Antimicrobial Stewardship: Opportunities and Challenges. Clin Ther. 2024;46:474–80. [DOI] [PubMed]
106. Kodali S. Leveraging AI for Real-Time Clinical Decision Support Systems in Oncology: Utilizing Machine Learning for Cancer Diagnosis, Prognosis, and Treatment Planning Based on Multi-Modal Patient Data. Los Angeles J Intell Syst Pattern Recognit. 2023;3:521–8.
107. Gouru G. Precision oncology: Computational methods for multi-omics data integration to improve drug response prediction. Camb Prisms: Precis Med. 2025;3:e6. [DOI] [PubMed] [PMC]
108. He Y, Huang F, Jiang X, Nie Y, Wang M, Wang J, et al. Foundation Model for Advancing Healthcare: Challenges, Opportunities and Future Directions. IEEE Rev Biomed Eng. 2025;18:172–91. [DOI] [PubMed]
109. Wang L, Chen X, Zhang L, Li L, Huang Y, Sun Y, et al. Artificial intelligence in clinical decision support systems for oncology. Int J Med Sci. 2023;20:79–86. [DOI] [PubMed] [PMC]

110. Wu Y, Xie L. AI-driven multi-omics integration for multi-scale predictive modeling of genotype-environment-phenotype relationships. Comput Struct Biotechnol J. 2025;27:265–77. [DOI] [PubMed] [PMC]
111. Vavekanand R. A Comprehensive Review of Multimodal Large Language Models for Medical Imaging and Omics Data. Arch Comput Methods Eng. 2026;33:6149–71. [DOI]
112. Sidorenko D, Pushkov S, Sakip A, Leung GHD, Lok SWY, Urban A, et al. Precious2GPT: the combination of multiomics pretrained transformer and conditional diffusion for artificial multi-omics multi-species multi-tissue sample generation. NPJ Aging. 2024;10:37. [DOI] [PubMed] [PMC]
113. Mardinoglu A, Turkez H, Shong M, Srinivasulu VP, Nielsen J, Palsson BO, et al. Longitudinal big biological data in the AI era. Mol Syst Biol. 2025;21:1147–65. [DOI] [PubMed] [PMC]
114. Calvino G, Peconi C, Strafella C, Trastulli G, Megalizzi D, Andreucci S, et al. Federated Learning: Breaking Down Barriers in Global Genomic Research. Genes. 2024;15:1650. [DOI] [PubMed] [PMC]
115. Allen GI, Gan L, Zheng L. Interpretable Machine Learning for Discovery: Statistical Challenges and Opportunities. Annu Rev Stat Appl. 2024;11:97–121. [DOI]
116. Chen C, Liu W, Lan Y, Li F, Li X. Exploring key genes in NAFLD linked to glutamine metabolism: A comprehensive analysis combining multi-omics, machine learning and SHAP. Asia Pac J Clin Nutr. 2026;35:81–90. [DOI] [PubMed] [PMC]