Multi-Omics Approaches in Gastric Cancer: A Survey

www.surveyx.cn

Abstract

Gastric cancer (GC) remains a formidable challenge in oncology due to its complex etiology and significant heterogeneity, necessitating advanced research methodologies for effective diagnosis and treatment. This survey examines the integration of multi-omics approaches, encompassing genomics, proteomics, and metabolomics, to enhance understanding of GC's molecular landscape. The advent of high-throughput sequencing technologies has revolutionized GC research, enabling comprehensive molecular profiling and the identification of novel biomarkers critical for precision medicine. Integrative analyses facilitate the development of personalized therapeutic strategies by elucidating tumor heterogeneity and identifying actionable genetic variants. Despite significant advancements, challenges persist in data integration, standardization, and computational analysis, necessitating sophisticated frameworks and machine learning models to manage high-dimensional datasets. Innovative methodologies, such as graph neural networks and deep learning models, have shown promise in enhancing predictive accuracy and biological interpretation. The survey underscores the importance of addressing standardization and reproducibility issues to ensure the reliability of multi-omics findings. Future directions include the development of non-invasive biomarkers, validation of identified biomarkers across diverse populations, and exploration of combination therapies. By refining these approaches, the potential of multi-omics to revolutionize GC research and treatment can be fully realized, paving the way for the widespread adoption of personalized medicine in oncology.

1 Introduction

1.1 Overview of Gastric Cancer and Multi-Omics

Gastric cancer (GC) poses significant challenges in oncology due to its intricate etiology and considerable heterogeneity, complicating diagnosis and treatment [1]. As the fifth most prevalent malignancy worldwide, GC is a leading cause of cancer-related mortality, highlighting the urgent need for advanced research methodologies to clarify its pathogenesis [2]. The molecular alterations associated with GC, influenced by genetic, epigenetic, and environmental factors, necessitate comprehensive research strategies [3].

High-throughput sequencing technologies have revolutionized gastric cancer research by enabling extensive multi-omics data collection, allowing for refined investigations into cancer molecular profiles and taxonomy based on molecular subtypes [4]. Multi-omics approaches have become essential in this field, providing a holistic framework for understanding the molecular mechanisms underlying the disease [5]. By integrating genomics, transcriptomics, proteomics, and metabolomics, researchers achieve a comprehensive view of the biological systems involved in cancer development and progression [6].

These integrative analyses are vital for identifying novel biomarkers and therapeutic targets, advancing precision medicine and personalized therapeutic strategies [7]. Addressing drug resistance in cancer therapy emphasizes the need for predicting effective drug combinations, crucial for discovering combinatory therapies [8]. Innovations in computational methods enhance the utility of multi-omics

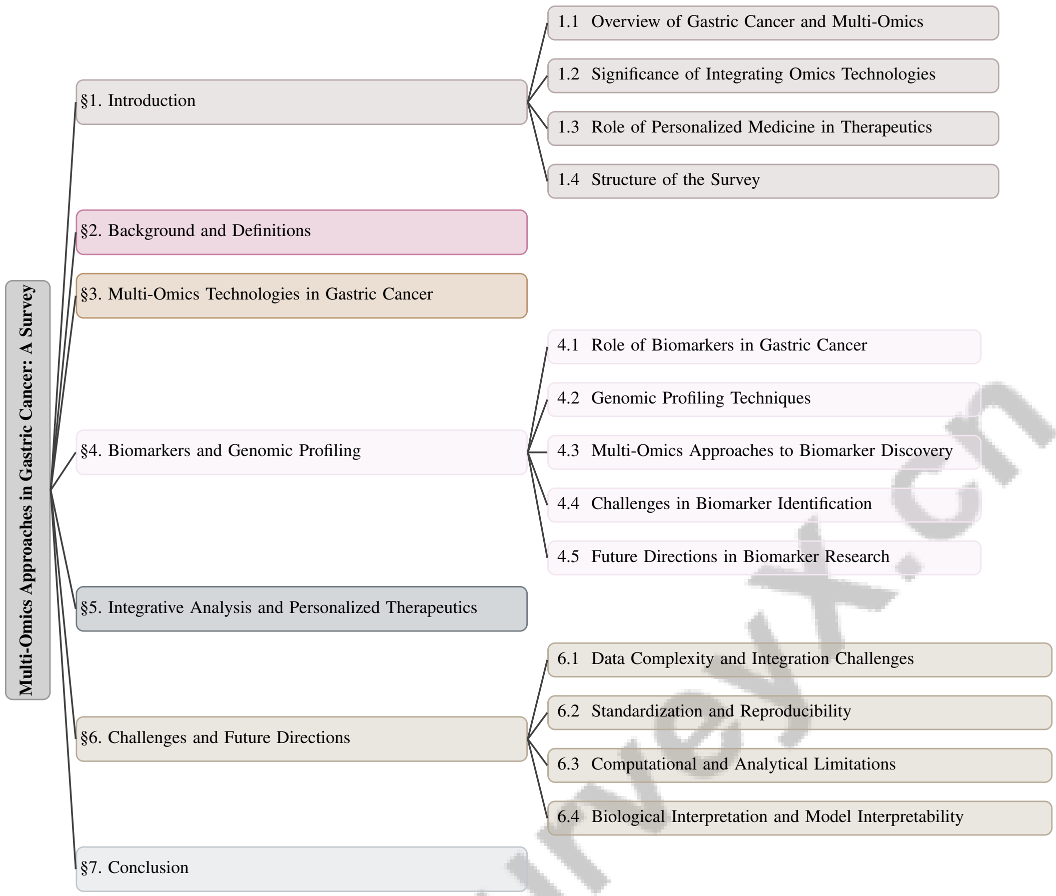


Figure 1: chapter structure

by facilitating the integration and analysis of high-dimensional data, enabling predictions of complex traits and cancer types [9].

Moreover, integrating large-scale biomedical datasets through network medicine approaches enriches our understanding of disease mechanisms and the development of targeted therapies [10]. Identifying multi-gene prognostic signatures through multi-omics research offers deeper insights into gastric cancer progression compared to traditional single-gene biomarkers, paving the way for innovative therapeutic interventions [11]. This comprehensive approach not only enhances our understanding of gastric cancer heterogeneity and progression but also underscores the transformative potential of multi-omics in cancer research and treatment paradigms [12]. The integration of multi-omics data facilitates personalized medicine, enabling tailored therapeutic strategies that align with individual patient profiles [7].

1.2 Significance of Integrating Omics Technologies

Integrating diverse omics technologies—genomics, transcriptomics, proteomics, and metabolomics—is crucial for unraveling the complex biological systems associated with gastric cancer. This comprehensive approach enhances our understanding of molecular interactions and pathways, essential for identifying therapeutic targets and biomarkers vital for advancing precision oncology [13]. By leveraging multi-omics data, researchers can address tumor biology's intrinsic heterogeneity and the variable responses of patients to treatments, thereby improving individualized therapeutic strategies.

The complexity and high dimensionality of multi-omics data present significant challenges, necessitating sophisticated computational strategies to extract meaningful insights. Traditional methods, such as all-vs-all correlation analyses, often yield unstructured outcomes that are difficult to interpret and may diminish statistical power due to excessive hypothesis testing [14]. Recent advancements in machine learning and deep learning techniques provide structured frameworks for analyzing complex multi-omics datasets, thereby enhancing our understanding of gastric cancer.

Integrating diverse biomedical data with biological networks is vital for understanding gastric cancer, as highlighted by network medicine approaches [10]. This integration not only enriches our understanding of disease mechanisms but also aids in developing targeted therapies. Resources like The Cancer Genome Atlas (TCGA) serve as invaluable platforms for developing computational methods that integrate multi-omics data, facilitating the discovery of novel therapeutic strategies and improving clinical outcomes [7].

Innovative methods, such as the AuDNNsynergy, exemplify the potential of integrating multi-omics data with chemical structure data to predict effective drug combinations, addressing multifaceted challenges in gastric cancer research [8]. The integration of whole exome sequencing with other omics technologies further enhances our understanding of gastric cancer by tackling the challenges posed by big data in genomic information [15]. This integrative approach is pivotal in developing personalized medicine strategies, allowing tailored therapeutic interventions that target specific molecular characteristics of individual tumors [16].

Furthermore, integrating omics technologies has significant implications for non-invasive biomarker identification, crucial for early-stage gastric cancer detection and improved patient outcomes [17]. By modeling interactions among features and integrating multiple data sources, researchers can overcome challenges associated with multi-omics data, unraveling the complex biological landscape of gastric cancer and offering new avenues for diagnosis, treatment, and personalized medicine approaches. The integration of diverse omics data types enhances predictive power for complex traits, emphasizing the importance of integrating omics technologies [6].

1.3 Role of Personalized Medicine in Therapeutics

Personalized medicine signifies a paradigm shift in gastric cancer therapeutics, tailoring treatments to the unique molecular and genetic profiles of individual patients and moving away from traditional one-size-fits-all approaches [1]. This approach is vital for addressing the heterogeneity in cancer responses, which often limits the efficacy of conventional treatment regimens [18]. By leveraging gastric cancer's molecular diversity, personalized medicine aims to enhance therapeutic outcomes through more precise and effective interventions.

Integrating whole exome sequencing (WES) into personalized therapeutic strategies has significantly improved diagnostic accuracy and facilitated the development of targeted therapies for gastric cancer [15]. WES allows for identifying specific genetic mutations that can be targeted with existing drugs, optimizing treatment efficacy for patients with well-defined mutations [19]. Despite these advancements, personalized treatments have not yet become the standard of care, underscoring the need for ongoing research and integration of multi-omics data to establish predictive and personalized therapeutic strategies.

Machine learning and data science methodologies are pivotal in advancing personalized medicine by enhancing the speed and efficacy of therapeutic development [20]. Knowledge-informed machine learning (KIML) methods, in particular, improve cancer diagnosis and prognosis, supporting the development of personalized treatment plans aligned with individual patient characteristics [21]. Furthermore, integrating genomic features with drug chemical properties refines drug sensitivity predictions, allowing for more accurate personalization of treatment regimens [22].

The potential of personalized medicine to revolutionize therapeutic strategies in gastric cancer is underscored by its ability to select optimal treatments for each patient based on unique biometric characteristics [23]. This approach enhances therapeutic efficacy while minimizing adverse effects by avoiding unnecessary treatments. Interactions between different omics layers provide valuable insights into disease mechanisms and potential therapeutic targets, critical for developing personalized therapeutic strategies [24].

However, implementing personalized medicine must be equitable to prevent exacerbating existing health disparities [25]. While the promises of personalized medicine are substantial, the field must address several challenges to fully realize its potential in enhancing therapeutic strategies for complex diseases like gastric cancer [26]. Continued advancements in multi-omics technologies and integrative analysis are essential for overcoming these challenges and establishing personalized medicine as a standard practice in cancer therapy.

1.4 Structure of the Survey

This survey is structured into several sections, each addressing key aspects of multi-omics research in gastric cancer. The introduction provides a comprehensive overview of gastric cancer and the significance of multi-omics approaches, setting the stage for subsequent discussions on the integration of omics technologies and the role of personalized medicine in therapeutics. The background section explores the prevalence of gastric cancer and current treatment challenges, along with definitions of critical concepts such as multi-omics, personalized medicine, biomarkers, genomic profiling, and integrative analysis.

Following this, the survey examines various multi-omics technologies employed in gastric cancer research, detailing advancements in high-throughput technologies and their specific applications. The discussion then shifts to biomarkers and genomic profiling, analyzing their roles in diagnosis and treatment, challenges in identification, and future research directions. The section on integrative analysis and personalized therapeutics highlights the importance of genomic profiling and data integration in developing personalized treatment strategies, including machine learning and innovative frameworks.

The survey concludes by addressing the challenges and future directions in multi-omics research, emphasizing data complexity, standardization, computational limitations, and biological interpretation. Each section is meticulously crafted to provide a well-rounded understanding of the current landscape and future potential of multi-omics in advancing gastric cancer research and treatment.The following sections are organized as shown in Figure 1.

2 Background and Definitions

2.1 Current Challenges in Gastric Cancer Treatment

Gastric cancer (GC) treatment is fraught with challenges owing to its complexity and the limitations of existing therapies. Late-stage diagnoses often lead to poor outcomes, highlighting the need for more effective systemic treatments [2]. The heterogeneity of GC underscores the necessity for precision medicine strategies that are tailored to the diverse genomic profiles of tumors. Resource and budget constraints further complicate the identification of suitable therapeutic strategies for individual patients [23]. The high costs associated with emerging biotechnologies exacerbate health disparities, particularly in low- and middle-income countries, necessitating equitable access to advancements in personalized medicine.

Current therapeutic approaches often overlook overlapping clusters among cancer cell lines and the interdependencies of multiple drugs, limiting their predictive accuracy and effectiveness [27]. The complexity of signaling pathways in GC tumorigenesis and progression complicates the development of targeted therapies [11]. Additionally, variability in immune responses among individuals is revealed through the interactive networks of immune cells and mediators [5]. Integrating phenotype-specific information into existing networks and addressing the incompleteness of current interactomes present substantial challenges [10]. The vast number of potential drug combinations and heterogeneous mechanisms of synergy further complicate experimental screening [8]. These challenges necessitate sophisticated analytical tools and innovative approaches to enhance diagnostic accuracy and treatment strategies [12]. Addressing these multifaceted issues is crucial for advancing the diagnosis, treatment, and management of gastric cancer.

2.2 Defining Multi-Omics and Its Relevance

Multi-omics involves the integration of diverse biological data types—genomic, transcriptomic, proteomic, and metabolomic—to gain a comprehensive understanding of cancer biology [28]. This approach is especially pertinent in gastric cancer, where the disease's complexity and heterogeneity require a multi-dimensional analytical perspective to elucidate intricate molecular interactions and pathways [10]. Multi-omics strategies facilitate insights into disease mechanisms and potential therapeutic interventions, advancing network medicine [10].

AI-generated, for reference only.

The integration of multi-omics data addresses the high-dimensionality and variability inherent in biological datasets, which complicate biomarker discovery and therapeutic target identification [9]. Advanced computational frameworks, such as graph neural networks, enhance the classification accuracy of cancer molecular subtypes, improving gastric cancer diagnosis and treatment precision [4]. However, aligning and merging diverse datasets, such as those from LC-MS studies, present significant challenges, necessitating innovative methods for effective data integration [29].

Despite these challenges, multi-omics integration is a powerful tool for advancing personalized medicine, enabling the development of tailored therapeutic strategies that cater to individual patients' unique molecular profiles [7]. This nuanced understanding of gastric cancer's molecular landscape enhances the identification of potential therapeutic targets and contributes to improved clinical outcomes through more effective interventions. Such an approach is essential for overcoming the limitations of existing methods that fail to fully utilize the biological context of the data, thereby optimizing patient survival predictions and treatment efficacy [7].

In recent years, the field of gastric cancer research has witnessed significant progress, largely attributed to the emergence of multi-omics technologies. These advancements have enabled researchers to approach the disease from a more holistic perspective, integrating various biological data types to enhance understanding and treatment outcomes. As illustrated in Figure 2, this figure demonstrates the structured advancements and applications of multi-omics technologies in gastric cancer research. It categorizes the major developments in high-throughput technologies and their applications, highlighting significant tools and frameworks for data integration, drug development, biomarker identification, and predictive modeling. This comprehensive overview not only underscores the importance of multi-omics in advancing gastric cancer research but also provides a visual representation of the interconnectedness of various technological advancements and their practical applications.

3 Multi-Omics Technologies in Gastric Cancer

3.1 Advancements in High-Throughput Technologies

Recent developments in high-throughput omics technologies have transformed gastric cancer research by facilitating the integration of complex datasets, thereby deepening our understanding of the disease's molecular underpinnings. Tools like OmiEmbed enable the embedding of high-dimensional multi-omics data into more interpretable low-dimensional spaces, improving data analysis [9]. The Multimodal Graph Neural Network Framework (MGNNF) exemplifies the application of graph-based techniques in classifying cancer molecular subtypes through multi-omics data integration [4].

Graph contrastive learning methods, such as MOGCL, leverage graph structures to surpass traditional integration techniques, addressing the inherent complexity and variability of multi-omics data [30]. AuDNNsynergy, a deep learning model, illustrates the potential of high-throughput technologies in drug development by predicting synergistic drug combinations through the integration of multi-omics datasets and drug chemical properties [8].

The Self-omics framework utilizes autoencoders and self-supervised learning to analyze high-dimensional multi-omics data, providing enhanced insights into immune mechanisms and improving cancer patient stratification [28, 31]. As illustrated in Figure 3, these advancements encompass key methodologies and frameworks that have contributed significantly to the field, specifically highlighting data integration, drug development, and patient stratification. The figure categorizes these advancements, showcasing OmiEmbed, MGNNF, and MOGCL for data integration, AuDNNsynergy for drug development, and Self-omics and IntegrAO for patient stratification.

These advancements underscore the pivotal role of multi-omics integration in addressing gastric cancer complexities, facilitating targeted therapeutic interventions. The alignment of multi-omics analysis with contemporary methodologies in feature selection, clinical outcome prediction, and clustering techniques enhances stratified medicine, biomarker identification, and drug repurposing efforts [32, 20].

3.2 Applications of Multi-Omics in Gastric Cancer

Multi-omics approaches have significantly advanced gastric cancer research by integrating diverse biological data types to elucidate complex molecular interactions and pathways. This integration is

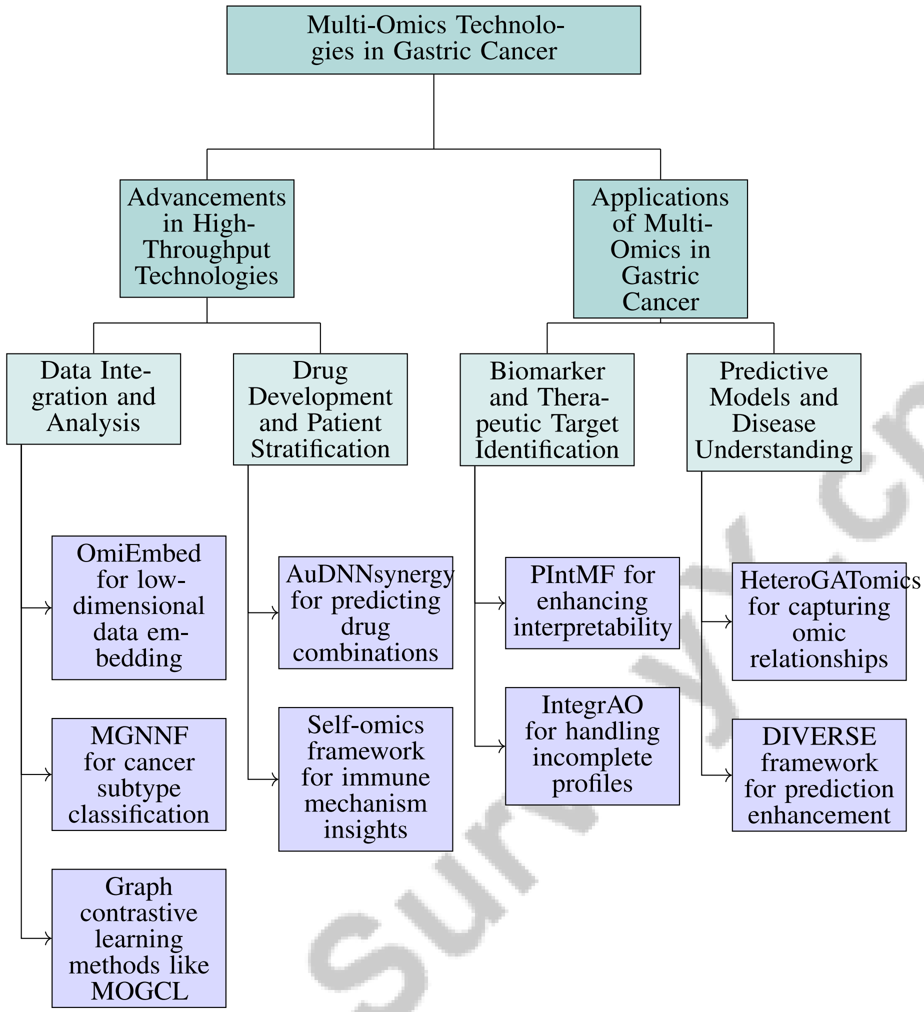


Figure 2: This figure illustrates the structured advancements and applications of multi-omics technologies in gastric cancer research. It categorizes the major developments in high-throughput technologies and their applications, highlighting significant tools and frameworks for data integration, drug development, biomarker identification, and predictive modeling.

crucial for precise genomic profiling and the development of targeted therapies, as demonstrated by successful immunotherapies and targeted treatments that improve outcomes for specific patient groups [2]. Moreover, multi-omics data integration enhances disease prediction and biomarker discovery, which are essential for personalized healthcare [33].

Computational frameworks like PIntMF employ matrix factorization to enhance interpretability and clustering, aiding in biomarker and therapeutic target identification [34]. The IntegrAO framework addresses the challenge of incomplete multi-omics profiles by enabling the classification of new samples with missing data, a common issue in gastric cancer research [31].

Advanced models such as HeteroGATomics capture intra- and inter-omic relationships, demonstrating superior predictive performance in multi-omics integration tasks crucial for understanding gastric cancer's biological mechanisms [35]. The DIVERSE framework systematically integrates biologically relevant datasets, enhancing prediction models and identifying potential therapeutic interventions [36].

Additionally, methods like Self-omics leverage unlabelled data through innovative pretext tasks for cancer type classification, improving gastric cancer diagnosis precision [28]. The DeePathology framework showcases the potential of multi-omics by inferring mRNA and miRNA expression profiles, as well as tissue and disease types, offering comprehensive insights into cancer biology [3].

AI-generated, for reference only.

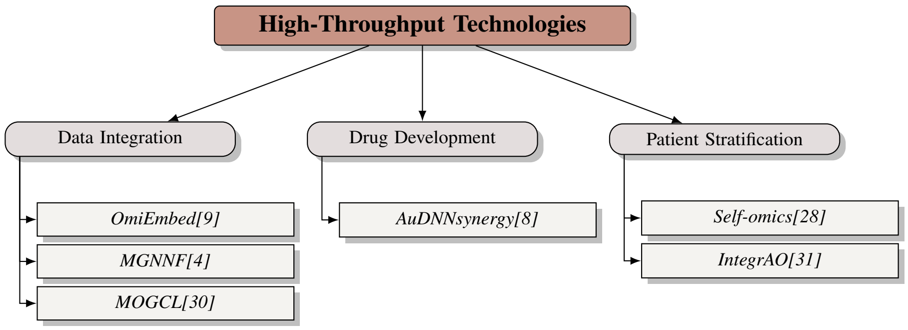
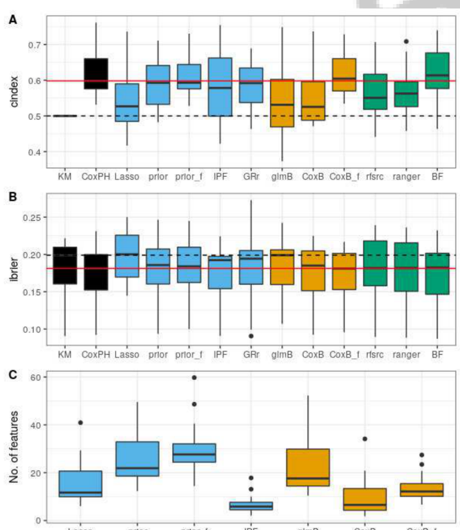


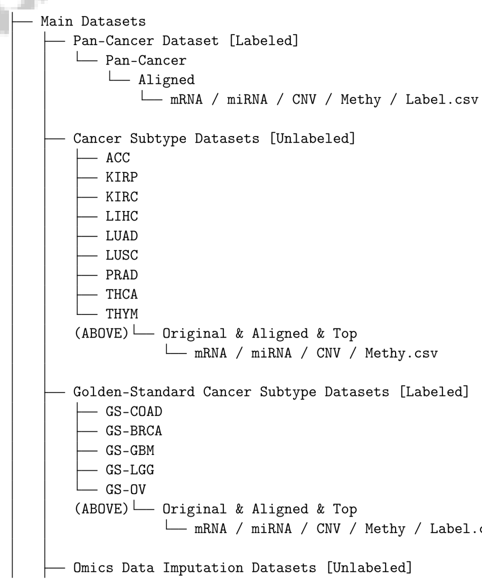
Figure 3: This figure illustrates the key advancements in high-throughput technologies, focusing on data integration, drug development, and patient stratification. Each category highlights specific methodologies and frameworks that have contributed to the field, such as OmiEmbed, MGNNF, and MOGCL for data integration, AuDNNsynergy for drug development, and Self-omics and IntegrAO for patient stratification.

Omics Stacking combines intermediate and late integration techniques, enhancing predictive performance in multi-omics studies [37]. GromovMatcher automates the matching of metabolic features across datasets using optimal transport, facilitating effective integration of metabolic data [29]. OmiEmbed captures complex relationships in omics data through a deep embedding module, demonstrating the utility of advanced embedding techniques in multi-omics analysis [9].

The integration of multi-omics approaches in gastric cancer research significantly enhances our understanding of the disease's complex biological mechanisms and facilitates the identification of new biomarkers, supporting the development of targeted and personalized therapeutic strategies. This ultimately improves patient stratification and outcomes in precision medicine [38, 39, 32, 33]. By integrating diverse datasets and employing advanced computational models, researchers can achieve a comprehensive understanding of gastric cancer biology, leading to more effective interventions and improved patient care.



(a) Comparison of Different Machine Learning Algorithms in Predicting Survival Rates[40]



(b) Main Datasets[41]

Figure 4: Examples of Applications of Multi-Omics in Gastric Cancer

As illustrated in Figure 4, the application of multi-omics technologies in gastric cancer research employs advanced computational methods to deepen understanding and treatment of this complex disease. The first example, a comparative analysis of machine learning algorithms, highlights the utility of these

AI-generated, for reference only.

technologies in predicting patient survival rates. By categorizing algorithms such as Lasso, prior, and prior$_{f}andevaluatingthembasedonmetricslikecindex, ibrier, andfeaturecount, researcherscanidentifyeffectivecom$
$Cancer, CancerSubtype, andOmicsDataImputation, furtherbreakingdownthesedatasetsintosubcategoriesbasedon$
$omicsapproachesinunravelingthecomplexitiesofgastriccancer, pavingthewayformorepreciseandpersonalizedther$
41].

4 Biomarkers and Genomic Profiling

4.1 Role of Biomarkers in Gastric Cancer

Biomarkers play a pivotal role in diagnosing and treating gastric cancer by elucidating the molecular characteristics essential for patient stratification and therapeutic decision-making. Their identification has propelled personalized medicine, enhancing patient outcomes [42]. Circular RNAs (circRNAs) have emerged as promising biomarkers, implicated in gastric cancer development, offering novel diagnostic and therapeutic avenues [43]. Despite these advancements, challenges such as the lack of standardized biomarker definitions and understanding persist, leading to potential resource misallocation and suboptimal clinical applications [44]. The integration of multi-omics data is hindered by variability in immune responses and regulatory networks [5], with manual alignment of metabolic features prone to errors, necessitating innovative solutions for compatibility across diverse analytical protocols [29].

Flexible predictive biomarker identification methodologies highlight the necessity of precise biomarker identification to improve patient outcomes in precision medicine [45]. The enhancement of classification accuracy through inter-omics and intra-omic connections within multi-layer graphs aids in identifying clinically relevant biomarkers [4]. Understanding interactions across omics layers further refines personalized medicine, emphasizing biomarkers' critical role in gastric cancer management [38]. Identifying biomarkers and molecular classifications has profound clinical implications, paving the way for innovative therapeutic interventions [11]. Recent research on cMET in gastric cancer has identified potential therapeutic targets for biomarker-driven strategies [46]. Advancing biomarker research is crucial for overcoming challenges in gastric cancer management, enabling early diagnosis, precise patient stratification, and targeted therapies [47].

4.2 Genomic Profiling Techniques

Genomic profiling techniques are crucial for detecting genetic alterations associated with gastric cancer, facilitating targeted therapeutic strategies. Whole-exome sequencing (WES) enables comprehensive analysis of coding regions to identify actionable mutations [15]. Integrating WES with other omics data provides a holistic view of gastric cancer's genetic landscape [7]. High-throughput sequencing technologies like next-generation sequencing (NGS) have revolutionized genomic profiling by allowing rapid, cost-effective sequencing of large DNA volumes, uncovering genetic alterations such as single nucleotide variations, insertions, deletions, and copy number variations [9]. These technologies aid in identifying molecular subtypes within gastric cancer, enhancing patient stratification and personalized treatment [4].

Advanced genomic profiling techniques, such as RNA sequencing, offer insights into gene expression patterns and regulatory mechanisms contributing to gastric cancer pathogenesis [3]. Integrating transcriptomic data with genomic information enhances understanding of the functional consequences of genetic alterations, informing targeted therapy development [7]. Machine learning and deep learning models applied to genomic data have improved the identification of genetic alterations and their implications in gastric cancer [9]. These computational approaches facilitate the discovery of novel genetic markers and therapeutic targets [4]. Genomic profiling techniques are integral to understanding gastric cancer's genetic underpinnings, forming a foundation for personalized medicine and targeted interventions [7].

4.3 Multi-Omics Approaches to Biomarker Discovery

Multi-omics approaches have significantly advanced biomarker discovery in gastric cancer by integrating diverse biological datasets, facilitating the identification of key biomarkers for diagnostic and therapeutic strategies. The DeePathology method exemplifies this integration by encoding the entire transcriptome into a low-dimensional latent vector, aiding in identifying crucial biomarkers for

AI-generated, for reference only.

gastric cancer diagnosis and treatment [3]. In metabolomics, the iCARH model effectively identifies metabolic biomarkers and infers perturbed pathways, elucidating complex biological interactions in gastric cancer [48]. Integrating these insights with other omics data enhances biomarker discovery precision, providing a comprehensive understanding of the disease. Innovations like the OmiEmbed framework have successfully integrated multiple omics data types, outperforming existing methods in biomarker discovery [9]. This framework illustrates the potential of advanced computational models in enhancing biomarker identification across omics layers.

The Pathway Dysregulation Score (PDS) captures the biological context of genomic data, providing accurate survival predictions and highlighting key biomarkers essential for personalized medicine [7]. This approach emphasizes contextualizing genomic alterations within biological pathways to improve prognostic accuracy. Additionally, GromovMatcher has demonstrated superior performance in aligning metabolomic datasets, achieving high precision and recall rates, crucial for accurately integrating metabolomic data in biomarker discovery [29]. This capability is vital for ensuring the reliability and reproducibility of biomarker identification across diverse datasets. Integrating multi-omics data in biomarker discovery offers a comprehensive understanding of gastric cancer biology by combining various biological features such as genomics, proteomics, and epigenomics. This integrative approach facilitates identifying critical molecular signatures that elucidate mechanisms underlying disease progression and enhance potential for precision medicine through improved patient stratification and targeted therapeutic strategies [32, 33, 24]. By leveraging advanced computational techniques and diverse datasets, researchers can enhance the precision of biomarker discovery, ultimately informing personalized therapeutic strategies and improving patient outcomes.

4.4 Challenges in Biomarker Identification

Biomarker identification in gastric cancer faces challenges due to the complexity of biological processes and the heterogeneity of omics data. The curse of dimensionality complicates extracting meaningful insights from high-dimensional datasets [49], often leading to overfitting when conventional cross-validation methods fail to account for heterogeneity between legacy and future trials, resulting in overestimation of biomarker effect sizes [50]. Missing data further complicates these challenges, impacting the robustness and reliability of biomarker discovery [51]. Dataset heterogeneity and imbalanced data distributions hinder the development of robust models that generalize across diverse patient populations [49]. Overfitting due to high data dimensionality necessitates robust validation methods to accurately assess the predictive power of identified biomarkers [52]. Sophisticated computational techniques are required to integrate multi-omics data, reduce dimensionality, and preserve biological relevance, a task that remains challenging despite recent advancements.

Addressing these challenges is essential for improving clinical outcomes by enhancing biomarkers' sensitivity and specificity, enabling accurate early diagnosis, prognosis, and treatment response monitoring. This will facilitate developing personalized therapeutic strategies tailored to individual patient profiles, ultimately reducing mortality rates associated with this aggressive and heterogeneous disease [53, 54, 27].

4.5 Future Directions in Biomarker Research

Future research in gastric cancer biomarker discovery should focus on developing non-invasive biomarkers for early detection and improved patient outcomes. Enhancing existing biomarkers' specificity and sensitivity is crucial for accurate diagnosis and effective disease progression monitoring [53]. Molecular classification of gastric cancer holds promise for tailoring treatment strategies to individual profiles, advancing personalized medicine. Integrating advanced computational techniques, such as machine learning and deep learning, with multi-omics data can enhance biomarker discovery precision by identifying novel molecular signatures indicative of disease progression and therapeutic response. This approach leverages recent advancements in molecular analysis, including circulating molecules like microRNAs and long non-coding RNAs, to deepen understanding of the mechanisms driving gastric cancer, facilitating non-invasive diagnostic tests and personalized treatment strategies [53, 32, 54].

Future studies should validate identified biomarkers across diverse populations to ensure generalizability and clinical applicability. Collaborative efforts between research institutions and clinical settings are crucial for translating biomarker discoveries into practice. By integrating multi-omics

AI-generated, for reference only.

analyses and employing advanced machine learning methods, these collaborations enhance robust biomarker identification and validation, ultimately improving patient care and outcomes. Addressing heterogeneity challenges in clinical trials and ensuring clarity in biomarker definitions, as emphasized by the FDA and NIH, can streamline the development and application of effective diagnostics and therapeutics in personalized medicine [44, 50, 32]. A comprehensive strategy integrating cutting-edge technologies, thorough validation processes, and collaborative initiatives is essential to enhance gastric cancer biomarker research. This approach aims to drive advancements in early detection, accurate diagnosis, and personalized treatment plans. Recent studies highlight the need for innovative, non-invasive biomarkers from body fluids like blood and gastric juice to improve diagnostic sensitivity and specificity. Leveraging multi-omics analyses and generative AI can facilitate identifying predictive biomarkers, ultimately leading to improved patient outcomes and tailored therapeutic strategies [32, 45, 55, 54, 53].

5 Integrative Analysis and Personalized Therapeutics

5.1 Genomic Profiling and Integrative Analysis

Genomic profiling and integrative analysis are pivotal in advancing personalized therapeutics for gastric cancer by elucidating tumor heterogeneity and identifying critical molecular alterations. These methodologies facilitate the identification of actionable genetic variants through integrated multi-omics analyses, enhancing our understanding of disease mechanisms and supporting tailored treatment strategies. Leveraging advancements in omics technologies and causal inference methods, these approaches enable comprehensive analyses of complex biological systems and the discovery of novel biomarkers for precise diagnostics and targeted interventions [56, 16, 32, 38]. The integration of genomics, transcriptomics, and proteomics allows for personalized treatment plans aligned with patient-specific molecular profiles.

Advanced computational frameworks have enhanced genomic profiling precision. The Self-omics method improves profiling by deriving robust representations from unlabelled multi-omics data, aiding personalized therapeutic strategies [28]. The DeePathology method exemplifies genomic profiling integration by accurately predicting tissue-of-origin and cancer types, facilitating targeted treatment regimens [3]. The iCARH model captures interdependencies between metabolites and other omics variables, identifying therapeutic targets in gastric cancer [48]. Additionally, the Pathway Dysregulation Score (PDS) method enhances interpretability and accuracy in survival predictions by capturing the biological significance of pathways, informing personalized therapeutic strategies [7].

The OmiEmbed framework improves predictive accuracy by learning a shared representation of multi-omics data that captures relevant biological signals across tasks, enhancing therapeutic target identification [9]. The uniCATE method offers a flexible nonparametric inference procedure to evaluate the predictive potential of biomarkers, contributing to personalized therapeutics [45]. Despite advancements, challenges remain regarding the transparency of deep learning models used in multi-omics analysis, complicating their decision-making processes [38]. Addressing these issues is crucial for improving interpretability and reliability in personalized medicine.

Genomic profiling and integrative analysis are critical for advancing personalized therapeutics in gastric cancer, enabling identification of specific biomarkers and molecular characteristics that inform early diagnosis, prognosis, and tailored treatment strategies [53, 32, 54]. Future research should focus on integrating multi-omics data and artificial intelligence to enhance cancer treatment precision and identify novel therapeutic strategies.

5.2 Integration and Analysis of Multi-Omics Data

Integration and analysis of multi-omics data are vital for advancing therapeutic development in gastric cancer, providing a comprehensive understanding of intricate biological systems. This approach encompasses genomics, transcriptomics, proteomics, and epigenomics, allowing researchers to uncover relationships among genes, proteins, and epigenetic factors influencing disease phenotypes. Employing diverse data integration methods, such as statistical and model-based approaches, enables the identification of biomarkers, stratification of patient subgroups based on therapeutic responses, and exploration of potential drug repurposing opportunities, enhancing precision and efficacy of treatment strategies [32, 24]. Synthesizing diverse omics data elucidates molecular interactions and

AI-generated, for reference only.

pathways driving disease progression, with advanced computational frameworks playing a crucial role in effective integration and extraction of meaningful insights for personalized therapeutic strategies.

The SeNMo framework, designed for multi-omics data analysis, integrates various molecular data modalities to predict patient outcomes [57]. This framework enhances integration by capturing unique characteristics of each omics dataset while facilitating unified biological insights extraction. Similarly, the IntegrAO method combines partially overlapping patient graphs from diverse omics sources and utilizes graph neural networks to produce unified patient embeddings, enhancing diverse omics data integration [31]. The HeteroGATomics method models complex relationships among features and patients through a heterogeneous graph structure, enhancing expressive power and improving multi-omics data integration [35]. The MOGCL framework constructs graphs from multi-omics data and employs contrastive learning as a pre-training strategy to enhance classification performance, showcasing graph-based approaches' utility in multi-omics integration [30].

Additionally, the PIntMF approach extracts patterns and clusters samples from multi-omics data through a penalized matrix factorization method, facilitating meaningful biological insights [34]. The AuDNNsynergy model utilizes comprehensive genomic data to identify patterns enhancing drug combination predictions, underscoring multi-omics data integration's significance for therapeutic development [8]. Integrating multi-omics data to enhance predictive power of network models is emphasized as a key insight from network medicine approaches, highlighting integrative strategies' importance in advancing therapeutic development [10]. By leveraging these approaches, researchers can develop tailored therapeutic strategies aligned with individual patients' unique molecular profiles, improving clinical outcomes and advancing personalized medicine. Advanced computational methods underscore multi-omics integration's significance in enhancing gastric cancer therapeutic development precision and efficacy.

5.3 Machine Learning and Deep Learning Approaches

Machine learning (ML) and deep learning (DL) have transformed multi-omics data analysis, providing essential capabilities for studying and treating gastric cancer. These computational approaches address complexity and high dimensionality of multi-omics datasets, facilitating meaningful patterns and relationships extraction crucial for advancing personalized medicine strategies [21]. The integration of machine learning, particularly knowledge-informed machine learning (KIML), has been pivotal in enhancing therapeutic applications' predictive accuracy [20].

Recent advancements highlight integrating biological knowledge into deep learning models, significantly improving their interpretability and applicability in cancer biology [58]. This integration allows nuanced interpretations of multi-omics data, supporting novel therapeutic targets identification and enhancing personalized treatment strategies precision [59]. Causal-based feature selection combined with machine learning classifiers has further improved gastric cancer diagnostic accuracy, demonstrating these approaches' potential in clinical applications [60].

In single-cell multi-omics, artificial intelligence (AI) and machine learning methods dominate computational tools landscape, significantly enhancing biological interpretations, such as clustering [61]. These methodologies facilitate complex datasets integration, enabling researchers to uncover intricate molecular interactions and pathways driving gastric cancer progression.

Machine learning and deep learning approaches integration in multi-omics research is instrumental in advancing gastric cancer understanding. Employing advanced computational techniques across multiple omics layers—including genomics, transcriptomics, proteomics, and radiomics—researchers gain a detailed molecular landscape understanding of diseases. This comprehensive insight facilitates novel therapeutic targets identification, improves patient stratification based on individual disease characteristics, and enhances therapeutic interventions' precision and efficacy, ultimately leading to more personalized and effective treatment plans in clinical practice [62, 20, 59].

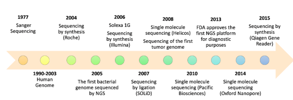
5.4 Innovative Frameworks and Models

Innovative frameworks and models are crucial for advancing integrative analysis in multi-omics research, particularly concerning gastric cancer. These models enhance complex datasets synthesis and facilitate meaningful insights extraction essential for understanding intricate biological interactions characterizing the disease. Developing scalable fusion methods and enhancing open-source software availability is vital for integrating diverse omics data types, enabling comprehensive multi-omics analysis [49].

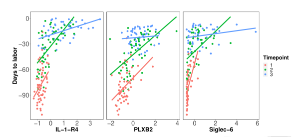
Introducing bio-centric interpretability in deep learning models underscores aligning computational approaches with biological mechanisms' importance. This alignment enhances model interpretability by integrating domain knowledge, improving deep learning applicability in cancer biology [58]. Future research should leverage these advancements by developing universal pipelines incorporating biological knowledge, enhancing multi-omics data integration [16].

Exploring emerging precision medicine trends, such as integrating genomic profiling and targeted therapies, is pivotal for improving patient outcomes in gastric cancer [11]. These advancements are anticipated to refine deep learning methods for multi-omics imputation, leveraging each omics type's unique features and developing adaptable frameworks to biological data complexities.

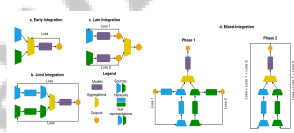
Innovative frameworks and models discussed in recent literature are critical for advancing integrative analysis in gastric cancer research. Leveraging artificial intelligence and multi-omics approaches, these methodologies enhance diagnostic accuracy and enable specific biomarkers identification from various biological fluids. This foundational work supports personalized therapeutic strategies development, addressing the urgent need for more effective and less invasive diagnostic tools and treatments for gastric cancer, a leading cause of cancer-related mortality worldwide [63, 32, 54, 64]. By addressing current limitations and exploring new methodologies, researchers can enhance multi-omics analyses' precision and efficacy, ultimately improving patient outcomes.



(a) The Evolution of Single-Molecule Sequencing Technologies[65]



(b) Timepoint-dependent changes in IL-1-R4, PLXB2, and Siglec-6 expression in human peripheral blood mononuclear cells[66]



(c) Four Different Integration Strategies in Deep Learning Models[67]

Figure 5: Examples of Innovative Frameworks and Models

As shown in Figure 5, the example on "Integrative Analysis and Personalized Therapeutics; Innovative Frameworks and Models" presents a comprehensive overview of cutting-edge methodologies and technological advancements in personalized medicine. The figure comprises three distinct visual representations that collectively underscore the evolution and integration of various scientific approaches. The first subfigure illustrates the historical progression of single-molecule sequencing technologies, highlighting significant milestones from 1977 to 2015. This timeline serves as a testament to the rapid advancements in sequencing capabilities, pivotal in enhancing our understanding of genetic information at an unprecedented resolution. The second subfigure focuses on the dynamic expression changes of specific genes—IL-1-R4, PLXB2, and Siglec-6—in human peripheral blood mononuclear cells across different timepoints, offering insights into temporal gene expression patterns crucial for developing targeted therapeutic strategies. The third subfigure demonstrates four distinct integration strategies in deep learning models, emphasizing the versatility and potential of artificial intelligence in processing complex biological data. Together, these examples illustrate the innovative frameworks and models at the forefront of personalized therapeutics, paving the way for tailored interventions and treatments informed by integrative analysis [65, 66, 67].

6 Challenges and Future Directions

In the realm of multi-omics research, particularly concerning gastric cancer, the integration of diverse data types presents substantial challenges that must be overcome to improve the quality and applicability of research findings. The inherent complexity of multi-omics data complicates analytical processes and requires a deep understanding of the underlying biological mechanisms. This section addresses specific challenges related to data complexity and integration.

6.1 Data Complexity and Integration Challenges

The integration of multi-omics data in gastric cancer research is impeded by the complexity, high dimensionality, and heterogeneity of these datasets. A significant hurdle is the limited transparency and explainability of deep learning models in biomedical contexts, which restricts their application in oncology [58]. This lack of interpretability complicates analysis, especially as existing methods struggle to utilize unlabeled data, exacerbating the "large p small n" issue where predictors outnumber samples [45].

The diversity of multi-omics modalities further complicates integration, with accuracy and generalizability across datasets remaining key challenges [5]. The computational demands required for integrating these datasets pose limitations, potentially hindering model applicability to larger datasets [48]. Moreover, non-uniform benchmarks in experimental conditions lead to ambiguous results regarding the performance of various integration methods [37].

A critical challenge is the integration of incomplete datasets, which can significantly reduce sample sizes when excluding samples with missing data [31]. Traditional methods often fail to account for interdependencies among omics layers, complicating predictive analytics in gastric cancer [36]. The curse of dimensionality further complicates the analysis of genome-wide omics data, hindering the capture of meaningful biological patterns [9].

Integrating multi-omics data also requires distinguishing primary perturbations from downstream effects within biological networks, a task traditional methods struggle to accomplish due to their tendency to analyze data types separately [11]. Additionally, reliance on existing datasets, which may not encompass all drug interactions or cancer types, limits the scope of integrative analyses [8]. The necessity for standardized methodologies across different omics studies and the underrepresentation of diverse populations further compound integration challenges [5].

Addressing these challenges is crucial for advancing multi-omics data integration in gastric cancer research. Developing scalable, reproducible, and interpretable models is essential for overcoming these obstacles and translating findings into clinical applications, ultimately enhancing the precision and efficacy of therapeutic interventions in gastric cancer and advancing personalized medicine [7].

6.2 Standardization and Reproducibility

Standardization and reproducibility are critical in multi-omics research, especially in gastric cancer, where diverse data types demand consistent methodologies. A primary obstacle is the lack of standardized protocols for data collection, processing, and analysis, leading to variability in outcomes and hindering reproducibility across studies. This variability is exacerbated by the diverse technologies and platforms employed in multi-omics studies, each with unique parameters and potential sources of error [7].

Developing standardized workflows is essential for ensuring reproducibility and comparability of findings across research groups. Establishing uniform criteria for data normalization, quality control, and statistical analysis is critical for minimizing technical biases and enhancing result reliability [9]. Implementing open-source software and databases can facilitate data sharing and collaboration, promoting transparency and reproducibility in multi-omics research [16].

Reproducibility is further challenged by the complexity of integrating multi-omics data, necessitating sophisticated computational tools capable of handling high-dimensional datasets while preserving biological relevance [4]. While machine learning and deep learning models are advantageous for analyzing complex data, they introduce additional challenges related to model interpretability and validation [21]. Ensuring transparency and reproducibility of these models is crucial for their successful clinical application [38].

Addressing standardization and reproducibility issues is vital for advancing the field and translating findings into clinical practice. By adhering to standardized methodologies, researchers can enhance the reliability and impact of multi-omics studies in gastric cancer, contributing to the advancement of personalized medicine [7].

6.3 Computational and Analytical Limitations

Integrating multi-omics data in gastric cancer research encounters significant computational and analytical limitations, primarily due to the complexity and high dimensionality of these datasets. Traditional methods, such as canonical correlation analysis, often struggle to model nonlinear interactions and are limited to handling only two modalities, restricting their utility in multi-omics studies [68]. The high dimensionality, where predictors exceed sample numbers, complicates capturing complex relationships [69]. Additionally, the integration process is hindered by the complexity and interdependence of various genomic assays [70].

Current studies face limitations such as inadequate sample sizes, overfitting due to high dimensionality, and the need for extensive computational resources [33]. The computational demands of current algorithms present major barriers to their application in big data scenarios, limiting scalability and efficiency in multi-omics integration [71]. Moreover, the accuracy of predictions from models like COMO depends on the quality and completeness of input omics data and reference models, potentially restricting applicability [72].

The requirement for patients to have all omics data available, as seen in models like HeteroGATomics, may limit applicability in cases of missing data [35]. This challenge is compounded by the complexity of data and difficulties in integrating multiple treatment versions within precision medicine contexts [18]. Despite progress, many studies still lack comprehensive integration of biological knowledge, resulting in models that may remain opaque and difficult to interpret [58].

Future research should focus on enhancing model stability, exploring single-cell multi-omics technologies, and integrating clinical information to improve input data quality for multi-omics integration [7]. Expanding datasets for drug combinations and incorporating additional genomic data types can further enhance predictive capabilities [8]. By addressing these computational and analytical limitations, researchers can advance personalized medicine in gastric cancer, ultimately improving treatment outcomes and patient care.

6.4 Biological Interpretation and Model Interpretability

Biological interpretation and model interpretability are critical challenges in multi-omics analysis, particularly in gastric cancer research. The complexity and high dimensionality of multi-omics datasets, including genomics, transcriptomics, and proteomics, often obscure meaningful biological insights, necessitating robust methodologies to manage heterogeneous and incomplete data [31]. Traditional models frequently struggle with interpretability, especially when large datasets are required to mitigate overfitting risks, potentially leading to misleading conclusions [45]. The intricate nature of deep learning models further complicates understanding of underlying biological mechanisms and interactions [3].

Assumptions of independence among omic data can limit the applicability of certain predictive models, hindering biological interpretation of complex interactions [38]. Additionally, limited understanding of signaling pathways and their interactions poses significant barriers to developing effective therapies for gastric cancer [11]. Reliance on unverifiable assumptions in existing methods can lead to biased estimates, complicating the identification of predictive biomarkers [45].

Frameworks such as Self-omics highlight the importance of ensuring that features present in pre-training datasets are available in downstream datasets, which presents challenges in biological interpretation and limits generalizability [28]. Biases in the interactome and potential reductions in sample sizes when incorporating clinical phenotypes can adversely affect model predictability and interpretability.

Moreover, models like GromovMatcher may restrict their application in multi-study analyses due to their inability to align more than two datasets simultaneously [29]. The performance and reliability of analyses heavily depend on the quality and representativeness of input multi-omics data, as demonstrated in models such as MOGCL [30]. Despite advancements, questions remain regarding optimal methods for integrating biological interpretability and the practical effectiveness of various explainability techniques [58].

Addressing the challenges of early diagnosis and the identification of specific biomarkers is crucial for advancing personalized medicine in gastric cancer, given the high mortality rates associated with late-stage diagnoses and the need for more effective, tailored treatment strategies [15, 54, 64]. By refining current methods and expanding their applicability to other omics data types, researchers can enhance both the precision and efficacy of multi-omics analyses, ultimately improving the development of tailored therapeutic strategies.

7 Conclusion

The convergence of multi-omics methodologies has significantly advanced the understanding and treatment of gastric cancer by detailing its intricate molecular architecture. Through the integration of diverse omics data—encompassing genomics, transcriptomics, proteomics, and metabolomics—researchers have been able to delineate intricate biological interactions and uncover novel therapeutic targets, thereby propelling the field of precision oncology forward. The adoption of advanced computational techniques, particularly Knowledge-Informed Machine Learning, has markedly enhanced the predictive capabilities of multi-omics analyses, paving the way for the development of personalized therapeutic strategies. Notable frameworks such as CustOmics and CSAE have demonstrated superior interpretability and predictive precision, offering robust solutions for managing complex biological datasets. These developments underscore the transformative potential of multi-omics data integration in improving therapeutic outcomes and advancing personalized medicine. Moreover, the synthesis of multi-omics data with existing knowledge bases enhances the accuracy and clarity of network inference, presenting a promising direction for refining methodologies in network analysis.

References

[1] Julita Machlowska, Jacek Baj, Monika Sitarz, Ryszard Maciejewski, and Robert Sitarz. Gastric cancer: epidemiology, risk factors, classification, genomic characteristics and treatment strategies. *International journal of molecular sciences*, 21(11):4012, 2020.

[2] Wen-Long Guan, Ye He, and Rui-Hua Xu. Gastric cancer treatment: recent progress and future perspectives. *Journal of hematology & oncology*, 16(1):57, 2023.

[3] Behrooz Azarkhalili, Ali Saberi, Hamidreza Chitsaz, and Ali Sharifi-Zarchi. Deepathology: Deep multi-task learning for inferring molecular pathology from cancer transcriptome, 2019.

[4] Bingjun Li and Sheida Nabavi. A multimodal graph neural network framework of cancer molecular subtype classification, 2024.

[5] Xiaojing Chu, Bowen Zhang, Valerie ACM Koeken, Manoj Kumar Gupta, and Yang Li. Multi-omics approaches in immunological research. *Frontiers in Immunology*, 12:668045, 2021.

[6] Heather E. Wheeler, Keston Aquino-Michaels, Eric R. Gamazon, Vassily V. Trubetskoy, M. Eileen Dolan, R. Stephanie Huang, Nancy J. Cox, and Hae Kyung Im. Poly-omic prediction of complex traits: Omickriging, 2013.

[7] Lana X Garmire. Strategies to integrate multi-omics data for patient survival prediction, 2020.

[8] Tianyu Zhang, Liwei Zhang, Philip R. O. Payne, and Fuhai Li. Synergistic drug combination prediction by integrating multi-omics data in deep learning models, 2018.

[9] Xiaoyu Zhang, Yuting Xing, Kai Sun, and Yike Guo. Omiembed: a unified multi-task deep learning framework for multi-omics data, 2021.

[10] Abhijeet R. Sonawane, Scott T. Weiss, Kimberly Glass, and Amitabh Sharma. Network medicine in the age of biomedical big data, 2019.

[11] Zi-Ning Lei, Qiu-Xu Teng, Qin Tian, Wei Chen, Yuhao Xie, Kaiming Wu, Qianlin Zeng, Leli Zeng, Yihang Pan, Zhe-Sheng Chen, et al. Signaling pathways and therapeutic interventions in gastric cancer. *Signal transduction and targeted therapy*, 7(1):358, 2022.

[12] Sanad Aburass, Osama Dorgham, and Jamil Al Shaqsi. A hybrid machine learning model for classifying gene mutations in cancer using lstm, bilstm, cnn, gru, and glove, 2024.

[13] Mingon Kang, Euiseong Ko, and Tesfaye B Mersha. A roadmap for multi-omics data integration using deep learning. *Briefings in Bioinformatics*, 23(1):bbab454, 2022.

[14] Thomaz F. S. Bastiaanssen, Thomas P. Quinn, and John F. Cryan. Knowledge-based integration of multi-omic datasets with anansi: Annotation-based analysis of specific interactions, 2023.

[15] Pawel Suwinski, ChuangKee Ong, Maurice HT Ling, Yang Ming Poh, Asif M Khan, and Hui San Ong. Advancing personalized medicine through the application of whole exome sequencing and big data analytics. *Frontiers in genetics*, 10:49, 2019.

[16] Nasim Vahabi and George Michailidis. Unsupervised multi-omics data integration methods: a comprehensive review. *Frontiers in genetics*, 13:854752, 2022.

[17] Jack Kelly, Carlo Berzuini, Bernard Keavney, Maciej Tomaszewski, and Hui Guo. Discovery methods for systematic analysis of causal molecular networks in modern omics datasets, 2022.

[18] Jonas Béal and Aurélien Latouche. Causal inference with multiple versions of treatment and application to personalized medicine, 2020.

[19] Abbi Abdel-Rehim, Oghenejokpeme Orhobor, Gareth Griffiths, Larisa Soldatova, and Ross D. King. Personalised medicine: Establishing predictive machine learning models for drug responses in patient derived cell culture, 2024.

[20] Kexin Huang, Cao Xiao, Lucas M. Glass, Cathy W. Critchlow, Greg Gibson, and Jimeng Sun. Machine learning applications for therapeutic tasks with genomics data, 2021.

[21] Lingchao Mao, Hairong Wang, Leland S. Hu, Nhan L Tran, Peter D Canoll, Kristin R Swanson, and Jing Li. Knowledge-informed machine learning for cancer diagnosis and prognosis: A review, 2024.

[22] Michael P. Menden, Francesco Iorio, Mathew Garnett, Ultan McDermott, Cyril Benes, Pedro J. Ballester, and Julio Saez-Rodriguez. Machine learning prediction of cancer cell sensitivity to drugs based on genomic and chemical properties, 2013.

[23] Jianzhong Du, Siyang Gao, and Chun-Hung Chen. A contextual ranking and selection method for personalized medicine, 2023.

[24] Stefan Graw, Kevin Chappell, Charity L Washam, Allen Gies, Jordan Bird, Michael S Robeson, and Stephanie D Byrum. Multi-omics data integration considerations and study design for biological systems and disease. *Molecular omics*, 17(2):170–185, 2021.

[25] Kishi Kobe Yee Francisco, Andrane Estelle Carnicer Apuhin, Myles Joshua Toledo Tan, Mickael Cavanaugh Byers, Nicholle Mae Amor Tan Maravilla, Hezerul Abdul Karim, and Nouar AlDahoul. Can personalized medicine coexist with health equity? examining the cost barrier and ethical implications, 2024.

[26] Jorge Alberto Bernstein Iriart. Precision medicine/personalized medicine: a critical analysis of movements in the transformation of biomedicine in the early 21st century. *Cadernos de saúde publica*, 35:e00153118, 2019.

[27] Hongmei Liu and J. Sunil Rao. Precision therapeutic biomarker identification with application to the cancer genome project, 2017.

[28] Sayed Hashim, Karthik Nandakumar, and Mohammad Yaqub. Self-omics: A self-supervised learning framework for multi-omics cancer data, 2022.

[29] Marie Breeur, George Stepaniants, Pekka Keski-Rahkonen, Philippe Rigollet, and Vivian Viallon. Optimal transport for automatic alignment of untargeted metabolomic data, 2024.

[30] Nishant Rajadhyaksha and Aarushi Chitkara. Graph contrastive learning for multi-omics data, 2023.

[31] Shihao Ma, Andy G. X. Zeng, Benjamin Haibe-Kains, Anna Goldenberg, John E Dick, and Bo Wang. Integrate any omics: Towards genome-wide data integration for patient stratification, 2024.

[32] Giovanna Nicora, Francesca Vitali, Arianna Dagliati, Nophar Geifman, and Riccardo Bellazzi. Integrated multi-omics analyses in oncology: a review of machine learning methods and tools. *Frontiers in oncology*, 10:1030, 2020.

[33] Parminder S Reel, Smarti Reel, Ewan Pearson, Emanuele Trucco, and Emily Jefferson. Using machine learning approaches for multi-omics data analysis: A review. *Biotechnology advances*, 49:107739, 2021.

[34] Morgane Pierre-Jean, Florence Mauger, Jean-François Deleuze, and Edith Le Floch. Pintmf: Penalized integrative matrix factorization method for multi-omics data, 2021.

[35] Sina Tabakhi, Charlotte Vandermeulen, Ian Sudbery, and Haiping Lu. Heterogeneous graph attention network improves cancer multiomics integration, 2024.

[36] Betül Güvenç Paltun, Samuel Kaski, and Hiroshi Mamitsuka. Diverse: Bayesian data integrative learning for precise drug response prediction, 2021.

[37] Tony Hauptmann and Stefan Kramer. A fair experimental comparison of neural network architectures for latent representations of multi-omics for drug response prediction, 2022.

[38] Ahmad Hussein, Mukesh Prasad, and Ali Braytee. Explainable ai methods for multi-omics analysis: A survey, 2024.

[39] Xiaojie Liu, Ting Peng, Miaochun Xu, Shitong Lin, Bai Hu, Tian Chu, Binghan Liu, Yashi Xu, Wencheng Ding, Li Li, et al. Spatial multi-omics: deciphering technological landscape of integration of multi-omics and its applications. *Journal of Hematology & Oncology*, 17(1):72, 2024.

[40] Moritz Herrmann, Philipp Probst, Roman Hornung, Vindi Jurinovic, and Anne-Laure Boulesteix. Large-scale benchmark study of survival prediction methods using multi-omics data, 2020.

[41] Ziwei Yang, Rikuto Kotoge, Xihao Piao, Zheng Chen, Lingwei Zhu, Peng Gao, Yasuko Matsubara, Yasushi Sakurai, and Jimeng Sun. Mlomics: Benchmark for machine learning on cancer multi-omics data, 2025.

[42] Naim U. Rashid, Daniel J. Luckett, Jingxiang Chen, Michael T. Lawson, Longshaokan Wang, Yunshu Zhang, Eric B. Laber, Yufeng Liu, Jen Jen Yeh, Donglin Zeng, and Michael R. Kosorok. High dimensional precision medicine from patient-derived xenografts, 2019.

[43] Rong Li, Jiajia Jiang, Hui Shi, Hui Qian, Xu Zhang, and Wenrong Xu. Circrna: a rising star in gastric cancer. *Cellular and Molecular Life Sciences*, 77:1661–1680, 2020.

[44] Robert M Califf. Biomarker definitions and their applications. *Experimental biology and medicine*, 243(3):213–221, 2018.

[45] Philippe Boileau, Nina Ting Qi, Mark J. van der Laan, Sandrine Dudoit, and Ning Leng. A flexible approach for predictive biomarker discovery, 2022.

[46] The role of cmet in gastric canc.

[47] Sebastian Canzler, Jana Schor, Wibke Busch, Kristin Schubert, Ulrike E Rolle-Kampczyk, Hervé Seitz, Hennicke Kamp, Martin von Bergen, Roland Buesen, and Jörg Hackermüller. Prospects and challenges of multi-omics data integration in toxicology. *Archives of Toxicology*, 94:371–388, 2020.

[48] Takoua Jendoubi and Timothy M. D. Ebbels. Integrative analysis of time course metabolic data and biomarker discovery, 2018.

[49] Sina Tabakhi, Mohammod Naimul Islam Suvon, Pegah Ahadian, and Haiping Lu. Multimodal learning for multi-omics: A survey, 2022.

[50] Yichen Lu, Jane Fridlyand, Tiffany Tang, Ting Qi, Noah Simon, and Ning Leng. The future will be different than today: Model evaluation considerations when developing translational clinical biomarker, 2021.

[51] Meng Song, Jonathan Greenbaum, Joseph Luttrell IV, Weihua Zhou, Chong Wu, Hui Shen, Ping Gong, Chaoyang Zhang, and Hong-Wen Deng. A review of integrative imputation for multi-omics datasets. *Frontiers in genetics*, 11:570255, 2020.

[52] Tim Downing and Nicos Angelopoulos. A primer on correlation-based dimension reduction methods for multi-omics analysis, 2023.

[53] Tasuku Matsuoka and Masakazu Yashiro. Biomarkers of gastric cancer: Current topics and future perspective. *World journal of gastroenterology*, 24(26):2818, 2018.

[54] Laura Necula, Lilia Matei, Denisa Dragu, Ana I Neagu, Cristina Mambet, Saviana Nedeianu, Coralia Bleotu, Carmen C Diaconu, and Mihaela Chivu-Economescu. Recent advances in gastric cancer early diagnosis. *World journal of gastroenterology*, 25(17):2029, 2019.

[55] Wangyang Ying, Dongjie Wang, Xuanming Hu, Ji Qiu, Jin Park, and Yanjie Fu. Revolutionizing biomarker discovery: Leveraging generative ai for bio-knowledge-embedded continuous space exploration, 2024.

[56] Minhao Yao and Zhonghua Liu. An introduction to causal inference methods with multi-omics data, 2024.

[57] Asim Waqas, Aakash Tripathi, Sabeen Ahmed, Ashwin Mukund, Hamza Farooq, Matthew B. Schabath, Paul Stewart, Mia Naeini, and Ghulam Rasool. Self-normalizing foundation model for enhanced multi-omics data analysis in oncology, 2024.

[58] Magdalena Wysocka, Oskar Wysocki, Marie Zufferey, Dónal Landers, and André Freitas. A systematic review of biologically-informed deep learning models for cancer: fundamental trends for encoding and interpreting oncology data, 2023.

[59] Holger Fröhlich, Rudi Balling, Niko Beerenwinkel, Oliver Kohlbacher, Santosh Kumar, Thomas Lengauer, Marloes H Maathuis, Yves Moreau, Susan A Murphy, Teresa M Przytycka, et al. From hype to reality: data science enabling personalized medicine. *BMC medicine*, 16:1–15, 2018.

[60] Divyagna Bavikadi, Ayushi Agarwal, Shashank Ganta, Yunro Chung, Lusheng Song, Ji Qiu, and Paulo Shakarian. Machine learning driven biomarker selection for medical diagnosis, 2024.

[61] Stefan Stanojevic, Yijun Li, and Lana X. Garmire. Computational methods for single-cell multi-omics integration and alignment, 2022.

[62] Liangrui Pan, Zhichao Feng, and Shaoliang Peng. A review of machine learning approaches, challenges and prospects for computational tumor pathology, 2022.

[63] Hamideh Kerdegari, Kyle Higgins, Dennis Veselkov, Ivan Laponogov, Inese Polaka, Miguel Coimbra, Junior Andrea Pescino, Marcis Leja, Mario Dinis-Ribeiro, Tania Fleitas Kanonnikoff, and Kirill Veselkov. Foundational models for pathology and endoscopy images: Application for gastric inflammation, 2024.

[64] Rachel E Sexton, Mohammed Najeeb Al Hallak, Maria Diab, and Asfar S Azmi. Gastric cancer: a comprehensive review of current and future treatment strategies. *Cancer and Metastasis Reviews*, 39(4):1179–1203, 2020.

[65] Liya Popova and Valerie J. Carabetta. The use of next-generation sequencing in personalized medicine, 2024.

[66] Matteo D'Alessandro, Theophilus Quachie Asenso, and Manuela Zucknick. Integrating multiple data sources with interactions in multi-omics using cooperative learning, 2024.

[67] Hakim Benkirane, Yoann Pradat, Stefan Michiels, and Paul-Henry Cournède. Customics: A versatile deep-learning based strategy for multi-omics integration, 2022.

[68] Jeongyoung Hwang, Sehwan Moon, and Hyunju Lee. Sdgcca: Supervised deep generalized canonical correlation analysis for multi-omics integration, 2022.

[69] Stephen T Rush, Christine H Lee, Washington Mio, and Peter T Kim. The phylogenetic lasso and the microbiome, 2016.

[70] Sijia Huang, Kumardeep Chaudhary, and Lana X Garmire. More is better: recent progress in multi-omics data integration methods. *Frontiers in genetics*, 8:84, 2017.

[71] You Wu and Lei Xie. Ai-driven multi-omics integration for multi-scale predictive modeling of causal genotype-environment-phenotype relationships, 2024.

[72] Brandt Bessell, Josh Loecker, Zhongyuan Zhao, Sara Sadat Aghamiri, Sabyasachi Mohanty, Rada Amin, Tomáš Helikar, and Bhanwar Lal Puniya. Como: A pipeline for multi-omics data integration in metabolic modeling and drug discovery, 2023.

Disclaimer:

SurveyX is an AI-powered system designed to automate the generation of surveys. While it aims to produce high-quality, coherent, and comprehensive surveys with accurate citations, the final output is derived from the AI's synthesis of pre-processed materials, which may contain limitations or inaccuracies. As such, the generated content should not be used for academic publication or formal submissions and must be independently reviewed and verified. The developers of SurveyX do not assume responsibility for any errors or consequences arising from the use of the generated surveys.

AI-generated, for reference only.