EGFRm NSCLC and Adjuvant Targeted Therapy: A Survey

Abstract

This survey provides an in-depth examination of epidermal growth factor receptor-mutant (EGFRm) non-small cell lung cancer (NSCLC), focusing on adjuvant targeted therapy's role in improving patient outcomes. The paper outlines the prevalence and clinical implications of EGFR mutations, emphasizing the heterogeneity within NSCLC and the necessity for personalized medicine strategies. The survey explores the mechanisms by which targeted therapies improve survival, alongside the technological advancements enhancing their efficacy. It also addresses the development of resistance to these therapies, highlighting genetic mutations, alternative signaling pathways, and non-genetic mechanisms as key contributors. Biomarker identification is discussed as a critical component for predicting and monitoring resistance, with an emphasis on multi-omics data integration and machine learning applications. The survey concludes with a discussion on personalized medicine approaches, focusing on predictive biomarkers and the integration of advanced computational techniques to tailor treatments to individual genetic profiles. Future directions include refining multi-omics data integration, exploring combination therapies, and leveraging technological innovations to overcome resistance and improve therapeutic precision. These advancements promise significant improvements in the management of EGFRm NSCLC, ultimately enhancing patient outcomes through more effective and personalized treatment strategies.

1 Introduction

1.1 Structure of the Survey

This survey provides a comprehensive examination of EGFRm NSCLC and the role of adjuvant targeted therapy, emphasizing resistance pathways and biomarker identification for personalized medicine. It begins with an **Introduction** that highlights the significance of personalized approaches in treating EGFRm NSCLC. The subsequent **Background on EGFRm NSCLC** section offers an overview of non-small cell lung cancer, detailing the prevalence and clinical implications of EGFR mutations.

The survey then explores **Adjuvant Targeted Therapy**, discussing its concept, applications, and mechanisms that enhance patient outcomes, alongside recent technological innovations that improve the efficacy of these therapies. The section on **Mechanisms of Targeted Therapy Resistance** investigates genetic, phenotypic, and non-genetic factors contributing to resistance, as noted in prior studies [1].

In the **Biomarker Identification** section, the focus is on the role of biomarkers in predicting and monitoring resistance, highlighting advances in technology and the integration of multi-omics data. The survey further addresses **Personalized Medicine Approaches**, emphasizing predictive biomarkers, machine learning, and omics data integration.

The synthesizes key findings and outlines promising directions for technological advancements aimed at overcoming treatment resistance and enhancing personalized medicine strategies for patients

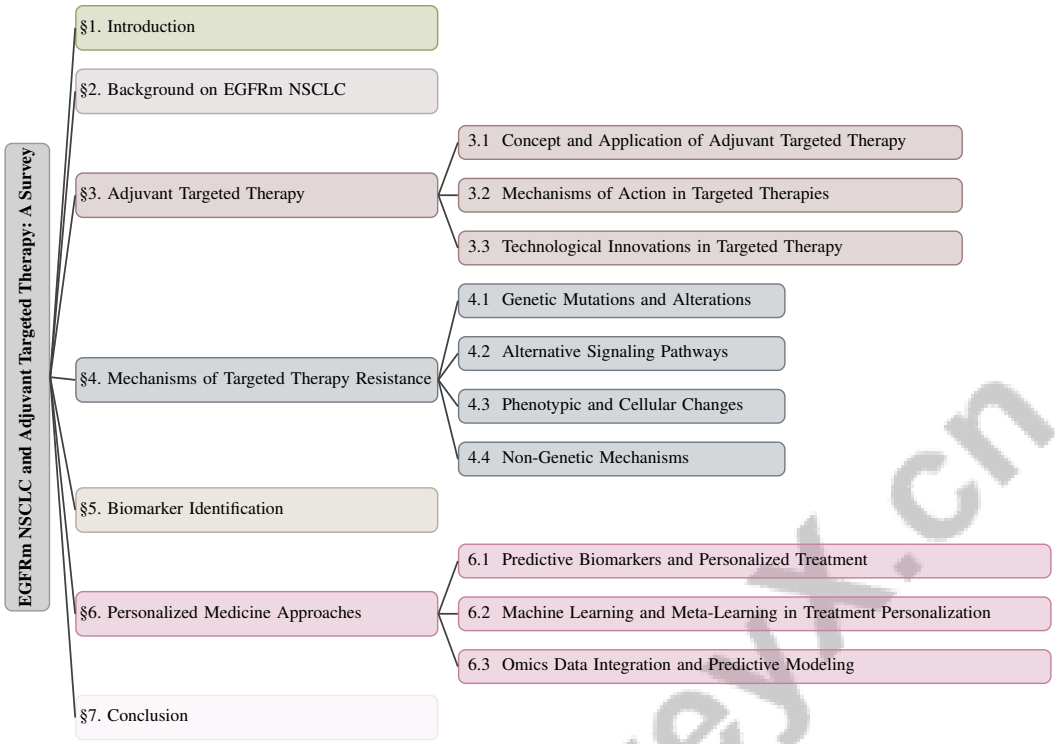


Figure 1: chapter structure

with EGFR-mutant non-small-cell lung cancer (EGFRm NSCLC). This includes novel approaches such as next-generation sequencing for precise molecular profiling, development of combination therapies, and formation of multidisciplinary molecular tumor boards to optimize treatment selection and improve patient outcomes [2, 3].The following sections are organized as shown in Figure 1.

2 Background on EGFRm NSCLC

2.1 Overview of NSCLC

Non-small cell lung cancer (NSCLC) accounts for approximately 85% of lung cancer cases, comprising adenocarcinoma, squamous cell carcinoma, and large cell carcinoma. Each subtype presents distinct biological behaviors, molecular changes, and clinical outcomes, influencing treatment strategies and prognosis. Genomic medicine and precision oncology have ushered in tailored therapeutic approaches based on specific molecular characteristics, enhancing patient management and survival rates [4, 5, 3, 6]. Factors such as smoking, environmental exposures, and genetic predispositions contribute to NSCLC's high global incidence and mortality rates.

A significant challenge in NSCLC management is integrating diverse data types to improve diagnosis, treatment, and prognosis. Addressing data interoperability, heterogeneous data management, and privacy issues is crucial for advancing NSCLC research and clinical practice [7]. The molecular diversity within tumors necessitates precise diagnostic and therapeutic strategies tailored to individual genetic profiles.

Recent advancements in molecular biology and bioinformatics have improved protein structure prediction methods, although achieving high accuracy across diverse protein sequences remains challenging [8]. These developments promise deeper insights into NSCLC pathogenesis and facilitate the discovery of novel therapeutic targets, thereby improving patient outcomes.

2.2 EGFR Mutations in NSCLC

The epidermal growth factor receptor (EGFR) is pivotal in NSCLC pathogenesis, with mutations in the EGFR gene serving as key oncogenic drivers. These mutations, primarily occurring in the

kinase domain, include exon 19 deletions and the L858R point mutation in exon 21. Activating EGFR mutations lead to constitutive receptor activation, driving uncontrolled cell proliferation and enhancing cell survival, significantly contributing to tumorigenesis. Secondary resistance mutations and alternative bypass signaling pathways complicate treatment as tumor cells adapt despite targeted therapies. Understanding intratumoral heterogeneity, including minimal residual disease cells that evolve and dominate the tumor population, is crucial for developing effective strategies to combat drug resistance and improve NSCLC treatment outcomes [6, 9, 10, 11, 12].

Clinically, EGFR mutations are predictive biomarkers for EGFR tyrosine kinase inhibitors (TKIs) efficacy, guiding treatment strategies in precision medicine and potential adjuvant therapies for resected cases [5, 13]. Patients with EGFR-mutant NSCLC exhibit significant responses to first-generation TKIs like gefitinib and erlotinib and second-generation inhibitors like afatinib. However, acquired resistance remains a challenge, limiting these therapies' long-term efficacy.

The ADAURA trial underscores the effectiveness of adjuvant targeted therapies in early-stage EGFR-mutant NSCLC, demonstrating significant improvements in disease-free survival with osimertinib, a third-generation EGFR TKI, post-surgery, highlighting the importance of EGFR mutation testing in guiding treatment decisions [13].

In addition to common mutations, rare and compound EGFR mutations exist, conferring varying sensitivities to different TKIs. This heterogeneity necessitates comprehensive molecular testing to tailor therapeutic strategies effectively. Ongoing research into alternative signaling pathways, such as PI3K/AKT/mTOR and RAS-MAPK, is essential for understanding resistance mechanisms and developing novel therapeutic approaches to overcome resistance in EGFR-mutant NSCLC [3].

2.3 Prevalence and Clinical Significance of EGFR Mutations

EGFR mutations are prevalent in a significant subset of NSCLC patients, especially among Asian populations, where frequencies reach 40%-50%, compared to 10%-15% in Western populations. These mutations predominantly occur in adenocarcinoma histology and are more common among non-smokers, females, and individuals with a history of light smoking. The most frequent mutations, exon 19 deletions and the L858R point mutation in exon 21, are associated with sensitivity to EGFR TKIs [5].

The clinical significance of EGFR mutations in NSCLC is profound, serving as critical biomarkers for targeted therapy selection. Patients with activating EGFR mutations, such as exon 19 deletions or the L858R mutation, typically respond well to EGFR TKIs like gefitinib, erlotinib, and osimertinib, enhancing response rates, prolonging progression-free survival, and significantly improving overall survival compared to traditional chemotherapy. However, resistance to these treatments often develops within 10 to 14 months, emphasizing the need for ongoing research into overcoming resistance mechanisms and optimizing treatment strategies [1, 11, 13]. Resistance mechanisms, including secondary mutations like T790M, MET amplification, and alternative pathway activation such as RAS, pose significant challenges to sustained treatment success.

The identification of EGFR mutations has revolutionized NSCLC management, enabling personalized treatment strategies that improve patient outcomes. The ADAURA trial further demonstrates the efficacy of adjuvant osimertinib in prolonging disease-free survival in patients with resectable EGFR-mutant NSCLC, highlighting the importance of early mutation testing and intervention [5]. Despite these advances, resistance development remains a core issue, necessitating ongoing research into novel therapeutic approaches and integrating multi-omics data to refine predictive models and enhance treatment personalization.

2.4 Molecular Testing and Identification of EGFR Mutations

Molecular testing for EGFR mutations is essential in NSCLC management, guiding targeted therapy selection and influencing treatment outcomes. Various methodologies, including polymerase chain reaction (PCR)-based assays, next-generation sequencing (NGS), and liquid biopsy techniques, enable the detection of both common and rare EGFR mutations, pivotal for determining patient eligibility for EGFR TKIs [5].

NGS has emerged as a powerful tool in precision oncology, offering comprehensive genomic profiling that captures a wide array of mutations, including exon 19 deletions and the L858R mutation. Its

ability to analyze multiple genes simultaneously makes it invaluable for personalized medicine, providing insights into tumor molecular landscapes and facilitating actionable mutation identification [14]. Liquid biopsy offers a non-invasive alternative for mutation detection, allowing real-time monitoring of tumor dynamics and resistance mutations, such as T790M, through circulating tumor DNA (ctDNA) analysis [15].

Integrating multi-omics data, including genomic, transcriptomic, and proteomic information, enhances the precision of molecular testing by providing a holistic view of tumor biology. However, the complexity and heterogeneity of multi-omics data present significant challenges for integration, necessitating advanced computational approaches to effectively explore cancer subtypes and guide treatment decisions [16]. Supervised clustering techniques, such as the CSMR algorithm, address these challenges by modeling high-dimensional data and accounting for patient subpopulation heterogeneity [17].

Radiomics, involving quantitative feature extraction from medical images, complements molecular testing by providing additional data to enhance clinical decision-making. This approach can improve tumor characterization and predict treatment responses by integrating imaging data with molecular profiles [18]. Despite these technologies' promise, significant challenges remain in managing and integrating big data, necessitating efficient bioinformatics tools and standardized data integration practices [19].

Molecular testing for EGFR mutations is indispensable for personalizing NSCLC treatment. It informs targeted therapy selection and enables dynamic adaptation of treatment strategies in response to evolving tumor profiles. As technological advancements continue to refine these methodologies, integrating comprehensive molecular testing into clinical practice holds the potential to revolutionize NSCLC management and improve patient outcomes [20].

3 Adjuvant Targeted Therapy

Category	Feature	Method
Concept and Application of Adjuvant Targeted Therapy	Innovative Data Analysis	CSMR[17]
Mechanisms of Action in Targeted Therapies	Adaptive Learning Strategies	Meta-learners[21]
	Pathway Analysis Enhancement	PF[22]
	Neural Network Integration	DF[8]
Technological Innovations in Targeted Therapy	Graph-Based Modeling	PML[23]

Table 1: This table provides a comprehensive overview of the innovative methodologies employed in adjuvant targeted therapy for EGFR-mutant non-small cell lung cancer. It categorizes the methods into three main areas: the concept and application of adjuvant targeted therapy, mechanisms of action in targeted therapies, and technological innovations. Each category is further detailed with specific features and methods, highlighting recent advancements and computational strategies in the field.

Adjuvant targeted therapy marks a pivotal advancement in treating epidermal growth factor receptor-mutant (EGFRm) non-small cell lung cancer (NSCLC), enhancing traditional surgical interventions by focusing on tumor-specific molecular characteristics. Understanding cancer stem cells (CSCs) and tumor heterogeneity is crucial for optimizing targeted therapies to disrupt tumor progression and improve patient outcomes. Identifying intrinsic and extrinsic signaling pathways contributing to therapy resistance and recognizing surviving cell subpopulations facilitates innovative strategies to enhance therapeutic efficacy and reduce relapse rates, guiding the development of effective treatment regimens and new agents targeting resistant cancer cell populations [4, 10, 3, 6].

To further elucidate these concepts, Figure 2 provides a comprehensive illustration of the hierarchical structure of adjuvant targeted therapy in EGFRm NSCLC. This figure details the key concepts, mechanisms, and technological innovations associated with personalized treatment approaches. It categorizes important trials and methodologies, elucidates the relevant signaling pathways, and highlights innovative approaches. Additionally, the diagram underscores the significant role of machine learning, data integration, and emerging tools in advancing targeted therapies, thereby enhancing our understanding of the therapeutic landscape in this context. Moreover, Table 1 delineates the various methodologies and innovations in adjuvant targeted therapy, offering insights into the conceptual framework, mechanisms of action, and technological advancements that underpin contemporary treatment strategies for EGFR-mutant non-small cell lung cancer. Furthermore, Table 2 provides a detailed overview of the methodologies and innovations in adjuvant targeted therapy, elucidating

the conceptual framework, mechanisms of action, and technological advancements that underpin contemporary treatment strategies for EGFR-mutant non-small cell lung cancer.

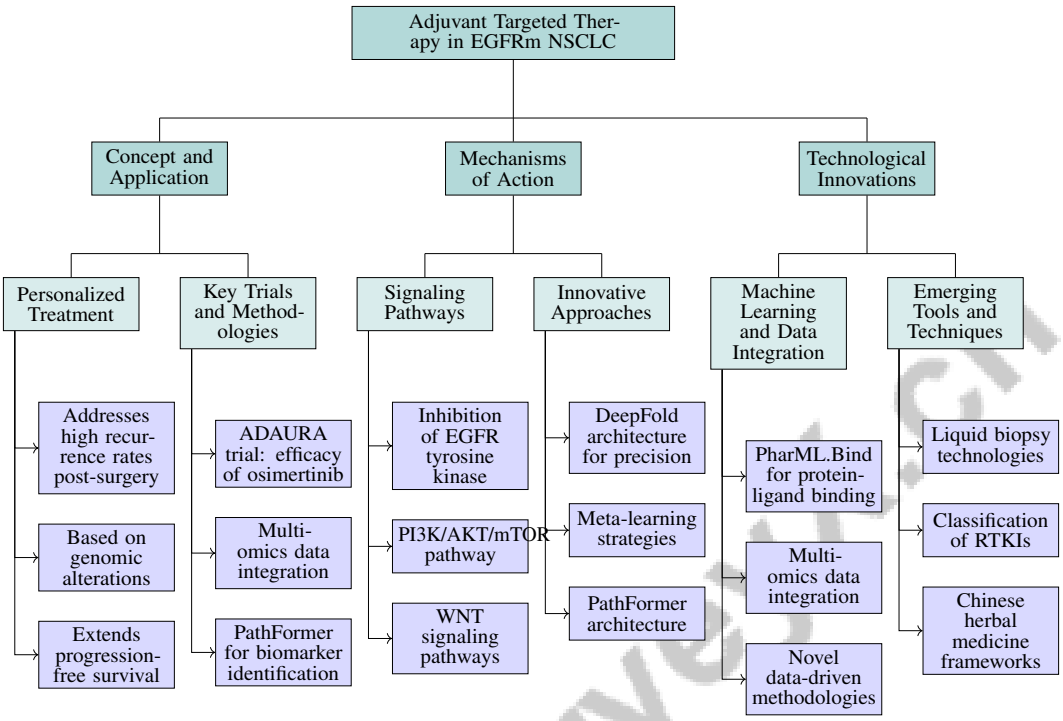


Figure 2: This figure illustrates the hierarchical structure of adjuvant targeted therapy in EGFRm NSCLC, detailing its concepts, mechanisms, and technological innovations. The diagram categorizes personalized treatment approaches, highlights key trials and methodologies, elucidates signaling pathways and innovative approaches, and underscores the role of machine learning, data integration, and emerging tools in advancing targeted therapies.

3.1 Concept and Application of Adjuvant Targeted Therapy

Adjuvant targeted therapy has transformed early-stage EGFRm NSCLC management by addressing high recurrence rates post-surgery and offering personalized treatment based on genomic alterations to enhance outcomes and potentially extend progression-free survival [1, 3, 6, 13]. Administered post-surgery, these targeted agents aim to eliminate residual disease and reduce recurrence risks by specifically inhibiting aberrant signaling pathways driven by EGFR mutations, thereby improving disease-free survival and overall patient outcomes.

The ADAURA trial highlights the efficacy of osimertinib, a third-generation EGFR tyrosine kinase inhibitor (TKI), in the adjuvant setting, demonstrating notable improvements in disease-free survival [5]. This underscores the need for precise molecular profiling to identify candidates for adjuvant therapy, ensuring treatments align with individual tumor genetic landscapes. Advancements in multi-omics data integration and computational methodologies, such as PathFormer, enhance biomarker identification and disease outcome predictions [22]. Frameworks examining resistance mechanisms, including synthetic lethality and drug synergy, provide insights for optimizing combinatorial adjuvant strategies [24]. Novel clustering algorithms like CSMR advance understanding of treatment responses and resistance mechanisms [17]. The integration of traditional therapies with novel agents, including Chinese herbal medicines, expands therapeutic options for EGFRm NSCLC [25].

3.2 Mechanisms of Action in Targeted Therapies

Targeted therapies for EGFRm NSCLC selectively inhibit aberrant signaling pathways critical for tumor growth and survival. They primarily inhibit the EGFR tyrosine kinase, whose heightened activity due to mutations like exon 19 deletions and L858R point mutations disrupts downstream

pathways, notably PI3K/AKT and RAS/RAF/MEK/ERK, reducing proliferation and increasing apoptosis in cancer cells [9]. Recent advancements emphasize the PI3K/AKT/mTOR (PAM) pathway's role in cancer pathogenesis, prompting the development of targeted therapies showing promise in clinical trials [26]. Strategies targeting cancer-associated fibroblasts (CAFs) and WNT signaling pathways aim to enhance therapeutic efficacy by addressing tumor growth, therapy resistance, and CSC eradication [27, 28].

Innovative computational methods like the DeepFold architecture, integrating convolutional and recurrent neural networks, enhance therapeutic precision [8]. Meta-learning strategies employing weighted pseudo-outcome regressions improve treatment effect estimation, enhancing therapeutic strategy efficiency and personalization [21]. Complementary approaches, such as integrating Chinese herbal medicines with conventional therapies, show promise in reducing adverse effects, boosting immune function, and improving overall efficacy [25]. PathFormer architecture's pathway-enhanced attention mechanism plays a crucial role in analyzing gene networks to identify critical pathways involved in cancer progression [22].

Figure 3 illustrates the hierarchical categorization of targeted therapy mechanisms in EGFRm NSCLC, highlighting key strategies such as EGFR inhibition, advanced computational models, and complementary approaches. This figure reflects the integration of molecular, technological, and traditional methods to enhance cancer treatment efficacy. Understanding targeted therapies' mechanisms is crucial for effectively combating cancer cells while minimizing harm to normal cells. The illustrations highlight CSC dynamics in tumor recurrence, evolutionary dynamics of cancer cells emphasizing treatment resistance, and a timeline of cancer drug approvals, reflecting the ongoing evolution and refinement of targeted therapies [10, 29, 30].

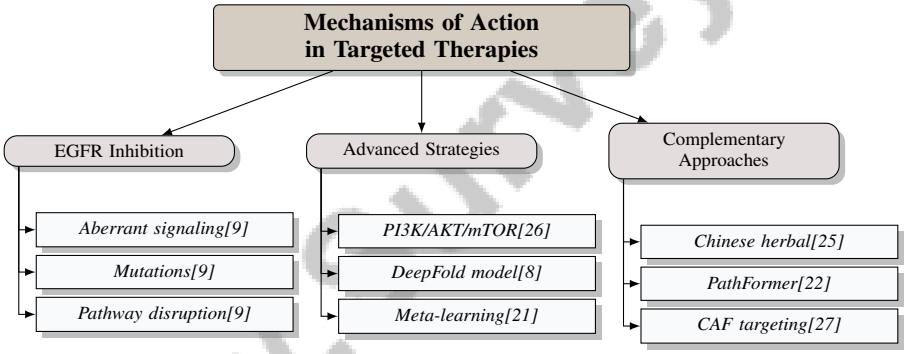


Figure 3: This figure illustrates the hierarchical categorization of targeted therapy mechanisms in EGFRm NSCLC, highlighting key strategies such as EGFR inhibition, advanced computational models, and complementary approaches, reflecting the integration of molecular, technological, and traditional methods to enhance cancer treatment efficacy.

3.3 Technological Innovations in Targeted Therapy

Technological advancements have transformed targeted therapies for EGFRm NSCLC, enhancing treatment precision and efficacy. Machine learning techniques like PharML.Bind, utilizing graph neural networks to predict protein-ligand binding affinities, improve drug-target interaction understanding and facilitate effective therapeutic agent design [23]. Integration of multi-omics data provides a comprehensive view of cancer's molecular underpinnings, with deep learning methodologies uncovering complex biological patterns, enabling novel therapeutic target identification [31]. These approaches are complemented by frameworks categorizing Chinese herbal medicines based on pharmacological properties, offering potential adjunctive benefits when combined with conventional therapies [25].

Innovative data integration frameworks synthesize omics and non-omics data using independent, conditional, and joint modeling approaches, crucial for understanding cancer's multifaceted nature and optimizing therapeutic strategies [19]. Novel data-driven methodologies improve targeted therapy efficacy, highlighting the importance of leveraging technological advances to address therapeutic challenges [4]. Classification of receptor tyrosine kinase inhibitors (RTKIs) based on their mechanisms of action is critical for refining targeted therapy strategies, aiding in selecting appropriate inhibitors to target specific mutations and resistance mechanisms [12]. Liquid biopsy technologies

have emerged as promising tools for real-time tumor dynamics monitoring and resistance mutation detection, enabling timely therapeutic regimen adjustments [1].

Feature	Concept and Application of Adjuvant Targeted Therapy	Mechanisms of Action in Targeted Therapies	Technological Innovations in Targeted Therapy
Purpose	Reduce Recurrence Risks	Inhibit Tumor Growth	Enhance Treatment Precision
Key Mechanism	Egfr Inhibition	Egfr Tyrosine Kinase	Drug-target Interaction
Technological Innovation	Multi-omics Integration	Deepfold Architecture	Machine Learning Techniques

Table 2: This table presents a comparative analysis of key features in adjuvant targeted therapy for EGFR-mutant non-small cell lung cancer, highlighting the purpose, mechanisms, and technological innovations. It delineates the objectives of reducing recurrence risks, inhibiting tumor growth, and enhancing treatment precision, alongside the mechanisms involving EGFR inhibition and advanced drug-target interactions. Technological advancements such as multi-omics integration and machine learning techniques are also emphasized, underscoring their role in advancing targeted therapy strategies.

4 Mechanisms of Targeted Therapy Resistance

Deciphering targeted therapy resistance is pivotal for enhancing treatment strategies. Resistance in EGFRm NSCLC stems from genetic mutations, alternative signaling pathways, and phenotypic changes, which collectively undermine therapeutic efficacy.

4.1 Genetic Mutations and Alterations

Genetic mutations play a critical role in resistance to targeted therapies in EGFRm NSCLC, manifesting as intrinsic or acquired resistance. Intrinsic resistance involves pre-existing mutations in downstream pathways or alternative oncogenic drivers that bypass EGFR inhibition [32]. Acquired resistance, often emerging during treatment, is frequently linked to secondary mutations like T790M in EGFR, which alter drug binding and reduce TKI efficacy [33]. The complexity of resistance signaling is further compounded by feedback loops and diverse resistance mechanisms [26]. Cancer cells exhibit adaptability through genetic mutations and tumor heterogeneity [32], intensified by the heterogeneity of CAFs and CSCs, which influence the tumor microenvironment and promote resistance [28].

Advanced computational frameworks like DeepFold enhance prediction accuracy by integrating diverse data sources, providing insights into the genetic landscape of resistance [8]. However, traditional linear models often fail to capture the heterogeneous and non-linear treatment effects in patient data, highlighting the need for sophisticated modeling approaches [34]. Despite the benefits of RTKIs, their limitations often lead to resistance and adverse effects, necessitating their integration with multi-omics data to fully understand genetic alterations contributing to resistance [12, 19].

4.2 Alternative Signaling Pathways

Resistance in EGFRm NSCLC is frequently driven by the activation of alternative signaling pathways that bypass EGFR inhibition, enabling cancer cell survival and proliferation despite TKIs. Key pathways include alternative RTKs and downstream cascades like PI3K/Akt and MAPK, crucial for cell growth and survival [26, 9, 10, 11, 12]. This complexity is illustrated in Figure 4, which highlights the alternative signaling pathways contributing to resistance in EGFR-mutant NSCLC, showcasing not only the alternative RTKs and the PI3K/Akt and MAPK pathways but also the mechanisms of cancer stem cell (CSC) resistance and EGFR-independent pathways. Cancer cells' adaptability under therapeutic pressure leads to resistant clones that evade existing treatments. Wnt, Notch, and Hedgehog pathways mediate CSC resistance, with reactivation of dormant CSCs posing significant treatment challenges [6].

The activation of EGFR-independent pathways, including dysregulation of other RTKs and downstream cascades like PI3K/AKT/mTOR and RAS/RAF/MEK/ERK, complicates therapeutic strategy development [4, 6, 9, 11, 35]. These pathways often upregulate in response to EGFR inhibition, providing alternative survival routes. The interplay between intrinsic and acquired resistance mechanisms underscores the necessity for comprehensive strategies addressing genetic variability and tumor heterogeneity.

Advanced computational methods, such as GNNs for predicting protein-ligand binding affinities, offer promising avenues for identifying novel targets within these alternative pathways, facilitating the development of new therapeutic agents [23].

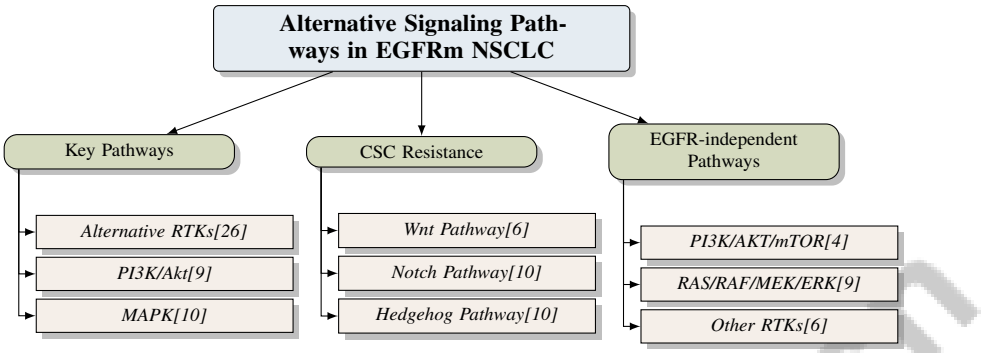


Figure 4: This figure illustrates the alternative signaling pathways contributing to resistance in EGFR-mutant NSCLC, highlighting key pathways such as alternative RTKs, PI3K/Akt, and MAPK, as well as CSC resistance mechanisms and EGFR-independent pathways.

4.3 Phenotypic and Cellular Changes

Phenotypic and cellular changes significantly contribute to resistance in EGFRm NSCLC, driven by tumor heterogeneity and resulting in resistant clones under therapeutic pressure. Lim et al. categorize these changes into intrinsic and acquired resistance mechanisms, highlighting tumor heterogeneity's complexity in drug resistance [6]. A key phenotypic change is EMT, where epithelial cells gain mesenchymal traits, enhancing migratory and invasive capabilities. EMT is linked to resistance against EGFR TKIs, contributing to stem cell-like properties and activating alternative pathways [11]. These include RTK amplifications and downstream activation of PI3K and MAPK, providing bypass mechanisms for tumor survival [11].

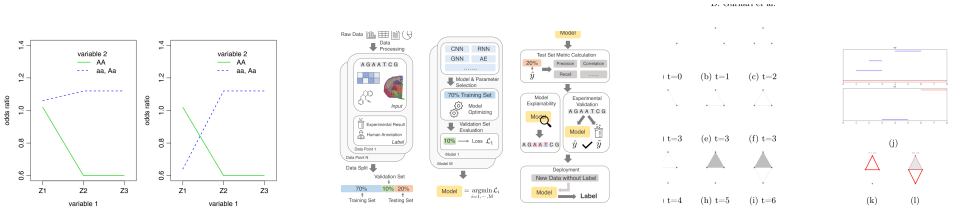
The intricate nature of drug interactions and challenges in predicting toxicity complicate combination therapies for resistance. Tumor heterogeneity, compensatory pathways, and resistance evolution necessitate innovative strategies, such as adaptive clinical trial designs and serial tumor profiling, to optimize treatment outcomes [24, 36]. Research focuses on categorizing these interactions and understanding molecular alterations like PTEN loss and mutations in PIK3CA and BRAF that contribute to phenotypic plasticity. The tumor microenvironment, influenced by factors like hypoxia and CAF presence, significantly affects drug resistance, creating a dynamic environment supporting resistant clones [6].

4.4 Non-Genetic Mechanisms

Non-genetic mechanisms substantially contribute to resistance in EGFRm NSCLC, encompassing diverse factors influencing resistance without genetic alterations. These mechanisms challenge effective therapeutic strategy development [29]. A prominent non-genetic mechanism involves tumor microenvironment alterations, enhancing cancer cell survival and promoting resistance. Factors like hypoxia, CAF presence, and immune cell infiltration create a supportive niche, enabling tumor cells to evade therapeutic pressures. These changes activate alternative pathways and induce phenotypic plasticity, allowing cancer cells to adapt and survive despite targeted therapy [6].

Epigenetic modifications, including DNA methylation and histone alterations, contribute to non-genetic resistance by altering gene expression patterns without modifying the DNA sequence. These changes can activate bypass pathways and suppress pro-apoptotic genes, facilitating cancer cell survival [11]. The role of non-coding RNAs in regulating gene expression and signaling pathways is increasingly recognized as a contributor to resistance mechanisms.

Despite advances in understanding non-genetic factors, significant questions remain about bypass pathways across cancer types and the long-term efficacy of targeting these pathways [11]. The complexity of biological data and insufficient predictive model performance hinder model interpretation and validation, complicating the translation of research findings into clinical practice [37].



(a) Odds Ratios for Two Different Allele Groups in a Binary Trait Model[38]

(b) Deep Learning Model Training and Deployment Process[39]

(c) The image shows a sequence of diagrams and graphs related to a mathematical concept.[40]

Figure 5: Examples of Non-Genetic Mechanisms

As illustrated in Figure 5, understanding non-genetic mechanisms in targeted therapy resistance is essential. The figure presents examples emphasizing the interplay between genetic and non-genetic factors in therapeutic resistance. The first subfigure shows odds ratios for allele groups within a binary trait model, elucidating probabilistic relationships influencing treatment outcomes based on genetic variability. The second subfigure explores the deep learning model training and deployment process, showcasing computational techniques predicting and analyzing resistance patterns through large dataset training. The third subfigure depicts diagrams and graphs representing a mathematical concept over time, highlighting the dynamic nature of non-genetic mechanisms in therapy resistance. These examples underscore the multifaceted nature of resistance mechanisms, extending beyond genetic mutations to include a range of environmental and cellular factors modeled through sophisticated analytical frameworks [38, 39, 40].

5 Biomarker Identification

5.1 Technological Advances in Biomarker Discovery

Technological advancements have significantly enhanced biomarker discovery for EGFRm NSCLC, enabling precise and personalized treatment approaches. Central to these advancements is the integration of multi-omics data, offering a comprehensive molecular view of cancer. The generalized finite mixture of multivariate regression (FMMR) model exemplifies improved biomarker interaction understanding through overlapping clustering [41].

High-throughput biomedical technologies facilitate diverse omics data collection, necessitating novel computational frameworks for integrative analysis. Approaches like Stacked Autoencoder-based multi-omics integration (SAEsurv-net) improve biomarker discovery by managing data heterogeneity through dimensionality reduction, particularly when molecular features are abundant but samples are limited [12]. Machine learning and AI, exemplified by computer-assisted biomedical image analysis, enhance biomarker identification by combining deep learning with traditional image processing [42]. Radiomics complements traditional methods by extracting quantitative features from medical images, aiding clinical decision-making.

Graph neural networks, such as the PathFormer model, integrate signaling networks with omics data, improving disease diagnosis and biomarker identification. This integration captures biological system complexities, identifying actionable biomarkers that guide therapeutic decisions. The need for omics monitoring to personalize targeted therapies underscores the critical role of diverse data types in biomarker discovery [12].

5.2 Challenges in Biomarker Validation

Biomarker validation for clinical use in EGFRm NSCLC encounters several challenges. A significant obstacle is the insufficient characterization of genetic and epigenetic alterations' relationship with drug response, which hinders accurate drug sensitivity predictions [43]. Developing comprehensive models integrating diverse biological data is essential for reliable therapeutic outcome predictions.

Machine learning applications in biomarker validation face technical and practical challenges, including resource constraints, model mistrust, and privacy and fairness concerns [39]. These issues hinder model adoption in clinical settings, highlighting the need for transparent frameworks ensuring ethical compliance and fostering trust.

Adapting tools like the oncoPredict package to new datasets presents hurdles, requiring continuous updates for relevance and usability [44]. The absence of comparable methods and datasets complicates benchmarking efforts and novel biomarker validation [40].

The biomarker validation process is further complicated by the instability of standard variance estimators in high-dimensional settings, leading to unreliable Type-I error control during hypothesis testing. This statistical instability undermines findings' robustness and reproducibility, necessitating innovative statistical methods [45]. The GERBIL framework, leveraging a multi-agent system and generative model approach, automates effective biomarker subset identification [46], though scalability and computational complexity concerns persist, especially with large-scale genomic datasets.

5.3 Integration of Multi-Omics Data

Integrating multi-omics data is crucial for enhancing biomarker discovery and validation in EGFRm NSCLC. This approach incorporates genomics, proteomics, and metabolomics data, providing a comprehensive understanding of cancer progression and therapeutic responses [2]. By synthesizing diverse data types, researchers can identify novel biomarkers essential for personalized medicine strategies.

Advanced computational frameworks like MOGONET use graph convolutional networks for omics-specific learning and a View Correlation Discovery Network (VCDN) for multi-omics data integration, significantly improving biomarker identification precision [47]. Similarly, the GERBIL framework combines a multi-agent system with a variational transformer model to optimize biomarker subset identification, embedding intricate biological knowledge into a continuous space for enhanced discovery [46].

Integrating multi-omics datasets captures intricate biological layer interactions, often overlooked by single-omics approaches. This comprehensive analysis is valuable in oncology, where identifying critical gene signatures can improve survival prediction accuracy and inform personalized therapeutic strategies. Challenges remain in ensuring model interpretability and generative methodologies' robustness in clinical settings [20].

Novel computational methods, such as Mendelian randomization (MR) tailored for multi-omics data analysis, offer promising avenues for overcoming biomarker discovery limitations [48]. Integrating omics with non-omics data, including clinical and epidemiological information, enhances biomarker models' predictive power and applicability [19]. Future research should refine integrative methods and explore diverse datasets to improve biomarker discovery robustness and accuracy [40]. Advancing multi-omics data integration will enable more effective, personalized treatment strategies in EGFRm NSCLC, enhancing patient outcomes.

6 Personalized Medicine Approaches

6.1 Predictive Biomarkers and Personalized Treatment

Predictive biomarkers are pivotal in tailoring treatments for EGFRm NSCLC patients, as they enable the customization of therapeutic strategies based on individual genetic and molecular profiles. Advanced computational techniques, notably artificial intelligence (AI), have enhanced patient data analysis and treatment outcome prediction, with some AI methods outperforming traditional statistical approaches in clinical response prediction and treatment optimization [49]. Individualized treatment rules optimize patient outcomes by assigning therapies based on clinical characteristics and predicted responses. The High-dimensional Individualized Treatment Selection (HITS) framework introduces a hypothesis testing approach to identify suitable treatments based on patient covariates, thereby enhancing personalized medicine precision [50]. Techniques like DTR-CT and DTR-CF use causal tree methods to derive heterogeneous treatment effects from patient data, providing accurate individualized recommendations [34].

Advancements in predictive imaging biomarkers have been achieved through computer-assisted analysis tools utilizing sophisticated computational techniques to interpret extensive datasets, improving clinical decision-making accuracy [42]. Future research should refine models to address non-linear relationships and explore additional imaging modalities to enhance predictive biomarker discovery [51]. Generative AI offers revolutionary potential for treatment customization, though interdisciplinary collaboration is essential to fully leverage its clinical applications [20]. Network-based approaches, such as NetBiTE, simulate biological interactions around drug targets, providing a more accurate depiction of perturbation propagation through cellular networks, vital for identifying effective treatment strategies [43].

6.2 Machine Learning and Meta-Learning in Treatment Personalization

Machine learning and meta-learning integration in EGFRm NSCLC treatment personalization has transformed therapeutic strategies, utilizing advanced computational techniques to tailor treatments to individual profiles. These models have significantly improved predictive accuracy and efficiency, enhancing patient outcomes [39], and adeptly manage complex datasets, integrating multi-omics data to inform personalized strategies [4]. Meta-learning addresses challenges related to time-varying treatment effects, with model-agnostic meta-learners enhancing adaptability and precision [21]. The dual-stage optimizer system, DOSA-MO, refines model selection by adjusting performance expectations based on solution characteristics, improving treatment personalization model robustness [52].

The Residual Weighted Learning (RWL) framework enhances finite sample performance in individualized treatment rules (ITRs) by applying residuals from regression models [53], complementing high-dimensional methods like HITS and contributing to personalized medicine by improving individualized treatment effects estimation [50]. Future research should focus on machine learning for predictive analytics, aligning with advanced computational techniques in personalizing strategies [54]. Developing hybrid models integrating mechanistic and machine learning approaches can enhance interpretability and address ethical and regulatory challenges associated with AI in healthcare [37].

6.3 Omics Data Integration and Predictive Modeling

Omics data integration is crucial for developing predictive models in personalized medicine, particularly for EGFRm NSCLC. This process synthesizes diverse data types, including genomics, epigenomics, transcriptomics, proteomics, and metabolomics, to create comprehensive molecular profiles that inform personalized strategies [48]. Advanced frameworks, such as SAEsurv-net, reduce dimensionality while preserving critical information, enhancing predictive performance [55]. Constructing patient similarity networks using graph-based approaches, like the Personalized Graph Feature-based Multi-omics Integration method, generates detailed patient profiles by employing random walks to capture relationships among omic layers [16]. Similarly, MOGONET integrates multi-omics data by combining Graph Convolutional Networks (GCNs) and View Correlation Discovery Networks (VCDN), leveraging omics-specific features and cross-omics correlations to enhance classification accuracy [47].

The need for improved integration techniques in multi-omics data is emphasized by the potential to develop predictive models reflecting complex cancer biology interactions [31]. Continuous optimization and generative modeling methodologies streamline optimal biomarker combination identification [46]. Additionally, methods like RWL, employing a smoothed ramp loss function and a difference of convex algorithm, address optimization limitations, contributing to robust model development [53]. Future research should prioritize enhancing model generalizability and addressing ethical concerns related to data usage, while exploring novel machine learning methodologies to boost therapeutic discovery [39]. Challenges in data standardization and big data technology integration remain critical development areas, as ongoing discussions highlight [19]. Advancing these integration strategies can significantly improve predictive accuracy in personalized medicine models, leading to more effective, tailored treatments for EGFRm NSCLC patients.

7 Conclusion

7.1 Future Directions and Technological Advances

The landscape of epidermal growth factor receptor-mutant (EGFRm) non-small cell lung cancer (NSCLC) research is poised for transformative advancements, driven by sophisticated computational methodologies and the seamless integration of multi-omics datasets. A pivotal focus will be the enhancement of genetic instrument selection and the optimization of algorithms for large-scale data analysis, crucial for advancing personalized medicine. The exploration of combination therapies, especially those targeting the WNT signaling pathway, holds potential for overcoming therapeutic resistance and improving efficacy.

Future research endeavors should concentrate on developing robust detection techniques for resistance mechanisms and investigating combination therapy strategies to enhance treatment efficacy. Additionally, the intricate interactions within the PI3K/AKT/mTOR (PAM) pathway and other signaling cascades offer fertile ground for innovation, with opportunities to discover new inhibitors and devise combination therapies that address resistance challenges.

The integration of 3D printing technology into clinical pharmacy represents a promising avenue for personalized medicine, necessitating the creation of standardized regulations and improved material compatibility. Concurrently, the enhancement of machine learning models' resilience to data noise and their application to diverse biological prediction tasks are vital for advancing computational cancer research.

Furthermore, extending the application of dynamic treatment models over longer periods and incorporating them into healthcare decision-making is critical for optimizing clinical treatment strategies. Advances in pharmacogenetics, coupled with enhanced data management systems, will support the incorporation of lifestyle factors into personalized medicine, thereby refining patient-specific treatment approaches.

References

[1] Shang-Gin Wu and Jin-Yuan Shih. Management of acquired resistance to egfr tki–targeted therapy in advanced non-small cell lung cancer. *Molecular cancer*, 17:1–14, 2018.

[2] Valentina Gambardella, Noelia Tarazona, Juan Miguel Cejalvo, Pasquale Lombardi, Marisol Huerta, Susana Roselló, Tania Fleitas, Desamparados Roda, and Andres Cervantes. Personalized medicine: recent progress in cancer therapy. *Cancers*, 12(4):1009, 2020.

[3] Min Yuan, Li-Li Huang, Jian-Hua Chen, Jie Wu, and Qing Xu. The emerging treatment landscape of targeted therapy in non-small-cell lung cancer. *Signal transduction and targeted therapy*, 4(1):61, 2019.

[4] 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.

[5] T Thamrongjirapat, D Muntham, P Incharoen, N Trachu, P Sae-Lim, N Sarachai, K Khiewngam, N Monnamo, N Kantathut, M Ngodngamthaweesuk, et al. Molecular alterations and clinical prognostic factors in resectable non-small cell lung cancer. *BMC cancer*, 24(1):200, 2024.

[6] Zuan-Fu Lim and Patrick C Ma. Emerging insights of tumor heterogeneity and drug resistance mechanisms in lung cancer targeted therapy. *Journal of hematology & oncology*, 12(1):134, 2019.

[7] Anastasia Krithara, Fotis Aisopos, Vassiliki Rentoumi, Anastasios Nentidis, Konstantinos Bougatiotis, Maria-Esther Vidal, Ernestina Menasalvas, Alejandro Rodriguez-Gonzalez, Eleftherios G. Samaras, Peter Garrard, Maria Torrente, Mariano Provencio Pulla, Nikos Dimakopoulos, Rui Mauricio, Jordi Rambla De Argila, Gian Gaetano Tartaglia, and George Paliouras. iasis: Towards heterogeneous big data analysis for personalized medicine, 2024.

[8] 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.

[9] Shannon Lee, Jens Rauch, and Walter Kolch. Targeting mapk signaling in cancer: mechanisms of drug resistance and sensitivity. *International journal of molecular sciences*, 21(3):1102, 2020.

[10] He-Ming Zhou, Ji-Gang Zhang, Xue Zhang, and Qin Li. Targeting cancer stem cells for reversing therapy resistance: mechanism, signaling, and prospective agents. *Signal transduction and targeted therapy*, 6(1):62, 2021.

[11] Qian Liu, Shengnan Yu, Weiheng Zhao, Shuang Qin, Qian Chu, and Kongming Wu. Egfr-tkis resistance via egfr-independent signaling pathways. *Molecular cancer*, 17:1–9, 2018.

[12] Charles Pottier, Margaux Fresnais, Marie Gilon, Guy Jérusalem, Rémi Longuespée, and Nor Eddine Sounni. Tyrosine kinase inhibitors in cancer: breakthrough and challenges of targeted therapy. *Cancers*, 12(3):731, 2020.

[13] Amit A Kulkarni, Abdul Rafeh Naqash, Sonam Puri, and Rodrigo Dienstmann. Is it time to implement adjuvant targeted therapy in egfr-mutant non–small-cell lung cancer? *JCO Precision Oncology*, 5:PO–20, 2021.

[14] 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.

[15] Mariantonia Di Sanzo, Luigi Cipolloni, Marina Borro, Raffaele La Russa, Alessandro Santurro, Matteo Scopetti, Maurizio Simmaco, and Paola Frati. Clinical applications of personalized medicine: a new paradigm and challenge. *Current pharmaceutical biotechnology*, 18(3):194–203, 2017.

[16] Saiful Islam and Md. Nahid Hasan. Personalized graph feature-based multi-omics data integration for cancer subtype identification, 2024.

[17] Wennan Chang, Changlin Wan, Yong Zang, Chi Zhang, and Sha Cao. Supervised clustering of high dimensional data using regularized mixture modeling, 2020.

[18] Philippe Lambin, Ralph TH Leijenaar, Timo M Deist, Jurgen Peerlings, Evelyn EC De Jong, Janita Van Timmeren, Sebastian Sanduleanu, Ruben THM Larue, Aniek JG Even, Arthur Jochems, et al. Radiomics: the bridge between medical imaging and personalized medicine. *Nature reviews Clinical oncology*, 14(12):749–762, 2017.

[19] Evangelina López de Maturana, Lola Alonso, Pablo Alarcón, Isabel Adoración Martín-Antoniano, Silvia Pineda, Lucas Piorno, M Luz Calle, and Núria Malats. Challenges in the integration of omics and non-omics data. *Genes*, 10(3):238, 2019.

[20] Isaias Ghebrehiwet, Nazar Zaki, Rafat Damseh, and Mohd Saberi Mohamad. Revolutionizing personalized medicine with generative ai: a systematic review. *Artificial Intelligence Review*, 57(5):128, 2024.

[21] Dennis Frauen, Konstantin Hess, and Stefan Feuerriegel. Model-agnostic meta-learners for estimating heterogeneous treatment effects over time, 2025.

[22] Zehao Dong, Qihang Zhao, Philip R. O. Payne, Michael A Province, Carlos Cruchaga, Muhan Zhang, Tianyu Zhao, Yixin Chen, and Fuhai Li. Highly accurate disease diagnosis and highly reproducible biomarker identification with pathformer, 2024.

[23] Aaron D. Vose, Jacob Balma, Damon Farnsworth, Kaylie Anderson, and Yuri K. Peterson. Pharml.bind: Pharmacologic machine learning for protein-ligand interactions, 2019.

[24] Juanita S Lopez and Udai Banerji. Combine and conquer: challenges for targeted therapy combinations in early phase trials. *Nature reviews Clinical oncology*, 14(1):57–66, 2017.

[25] Zhixue Wang, Fanghua Qi, Yangang Cui, Lin Zhao, Xiaogang Sun, Wei Tang, and Pingping Cai. An update on chinese herbal medicines as adjuvant treatment of anticancer therapeutics. *Bioscience trends*, 12(3):220–239, 2018.

[26] Antonino Glaviano, Aaron SC Foo, Hiu Y Lam, Kenneth CH Yap, William Jacot, Robert H Jones, Huiyan Eng, Madhumathy G Nair, Pooyan Makvandi, Birgit Geoerger, et al. Pi3k/akt/mtor signaling transduction pathway and targeted therapies in cancer. *Molecular cancer*, 22(1):138, 2023.

[27] Fanglong Wu, Jin Yang, Junjiang Liu, Ye Wang, Jingtian Mu, Qingxiang Zeng, Shuzhi Deng, and Hongmei Zhou. Signaling pathways in cancer-associated fibroblasts and targeted therapy for cancer. *Signal transduction and targeted therapy*, 6(1):218, 2021.

[28] Masaru Katoh. Canonical and non-canonical wnt signaling in cancer stem cells and their niches: Cellular heterogeneity, omics reprogramming, targeted therapy and tumor plasticity. *International journal of oncology*, 51(5):1357–1369, 2017.

[29] Jean-Christophe Marine, Sarah-Jane Dawson, and Mark A Dawson. Non-genetic mechanisms of therapeutic resistance in cancer. *Nature Reviews Cancer*, 20(12):743–756, 2020.

[30] Lei Zhong, Yueshan Li, Liang Xiong, Wenjing Wang, Ming Wu, Ting Yuan, Wei Yang, Chenyu Tian, Zhuang Miao, Tianqi Wang, et al. Small molecules in targeted cancer therapy: advances, challenges, and future perspectives. *Signal transduction and targeted therapy*, 6(1):201, 2021.

[31] 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.

[32] Xuan Wang, Haiyun Zhang, and Xiaozhuo Chen. Drug resistance and combating drug resistance in cancer. *Cancer drug resistance*, 2(2):141, 2019.

[33] Toshimitsu Yamaoka, Motoi Ohba, and Tohru Ohmori. Molecular-targeted therapies for epidermal growth factor receptor and its resistance mechanisms. *International journal of molecular sciences*, 18(11):2420, 2017.

[34] Theresa Blümlein, Joel Persson, and Stefan Feuerriegel. Learning optimal dynamic treatment regimes using causal tree methods in medicine, 2023.

[35] Behzad Mansoori, Ali Mohammadi, Sadaf Davudian, Solmaz Shirjang, and Behzad Baradaran. The different mechanisms of cancer drug resistance: a brief review. *Advanced pharmaceutical bulletin*, 7(3):339, 2017.

[36] Neil Vasan, José Baselga, and David M Hyman. A view on drug resistance in cancer. *Nature*, 575(7782):299–309, 2019.

[37] 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.

[38] Charles Kooperberg, Michael LeBlanc, James Y. Dai, and Indika Rajapakse. Structures and assumptions: Strategies to harness gene $\times$ gene and gene $\times$ environment interactions in gwas, 2010.

[39] 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.

[40] Davide Gurnari, Aldo Guzmán-Sáenz, Filippo Utro, Aritra Bose, Saugata Basu, and Laxmi Parida. Probing omics data via harmonic persistent homology, 2024.

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

[42] Leonardo Rundo. Computer-assisted analysis of biomedical images, 2021.

[43] Ali Oskooei, Matteo Manica, Roland Mathis, and Maria Rodriguez Martinez. Network-based biased tree ensembles (netbite) for drug sensitivity prediction and drug sensitivity biomarker identification in cancer, 2019.

[44] Danielle Maeser, Robert F Gruener, and Rong Stephanie Huang. oncopredict: an r package for predicting in vivo or cancer patient drug response and biomarkers from cell line screening data. *Briefings in bioinformatics*, 22(6):bbab260, 2021.

[45] Nima S. Hejazi, Philippe Boileau, Mark J. van der Laan, and Alan E. Hubbard. A generalization of moderated statistics to data adaptive semiparametric estimation in high-dimensional biology, 2022.

[46] 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.

[47] Tongxin Wang, Wei Shao, Zhi Huang, Haixu Tang, Jie Zhang, Zhengming Ding, and Kun Huang. Mogonet integrates multi-omics data using graph convolutional networks allowing patient classification and biomarker identification. *Nature communications*, 12(1):3445, 2021.

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

[49] Nicholas J Schork. Artificial intelligence and personalized medicine. *Precision medicine in Cancer therapy*, pages 265–283, 2019.

[50] Tianxi Cai, Tony Cai, and Zijian Guo. Optimal statistical inference for individualized treatment effects in high-dimensional models, 2020.

[51] Shuhan Xiao, Lukas Klein, Jens Petersen, Philipp Vollmuth, Paul F. Jaeger, and Klaus H. Maier-Hein. Enhancing predictive imaging biomarker discovery through treatment effect analysis, 2024.

[52] Luca Cattelani and Vittorio Fortino. Dual-stage optimizer for systematic overestimation adjustment applied to multi-objective genetic algorithms for biomarker selection, 2024.

[53] Xin Zhou, Nicole Mayer-Hamblett, Umer Khan, and Michael R. Kosorok. Residual weighted learning for estimating individualized treatment rules, 2015.

[54] 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.

[55] Xing Wu and Qiulian Fang. Stacked autoencoder based multi-omics data integration for cancer survival prediction, 2022.

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.