Artificial Intelligence and Computational Biology in Kinase-Targeted Drug Discovery for Pancreatic Ductal Adenocarcinoma: A Survey

Abstract

This survey paper explores the interdisciplinary approach of utilizing artificial intelligence (AI) and computational biology to enhance drug discovery efforts targeting kinases in the treatment of pancreatic ductal adenocarcinoma (PDAC), a highly aggressive form of pancreatic cancer. The integration of AI in drug discovery has revolutionized the identification of molecular patterns, improving the efficiency and precision of therapeutic interventions. This paper highlights the transformative potential of AI techniques, including machine learning and deep learning, in predicting drug-kinase interactions and designing novel inhibitors. Computational biology complements these advancements by providing robust frameworks for understanding the structural dynamics of biological systems, essential for the rational design of kinase inhibitors. Despite significant progress, challenges such as data quality, model interpretability, and computational costs persist. The paper underscores the importance of interdisciplinary collaboration and the integration of diverse data types to overcome these challenges. It also emphasizes the need for further research to develop selective kinase inhibitors that can address drug resistance and improve patient outcomes. The survey concludes by highlighting the potential of AI and computational biology to significantly advance drug discovery processes for PDAC, paving the way for more effective and targeted therapeutic interventions.

1 Introduction

1.1 Significance of PDAC as a Major Health Challenge

Pancreatic ductal adenocarcinoma (PDAC) poses a significant challenge in oncology, accounting for approximately 90% of pancreatic cancer cases and contributing substantially to global cancer mortality [1]. The disease is frequently diagnosed at an advanced stage, complicating treatment and resulting in a dismal 5-year survival rate of less than 3% [2]. The prevalence of oncogenic KRAS mutations in PDAC further drives uncontrolled cellular proliferation, exacerbating the poor prognosis [2].

Early detection is critical for improving survival rates, yet it remains elusive due to subtle early symptoms and rapid disease progression [3]. The inadequacy of current imaging techniques to predict overall survival underscores the complexity of PDAC, as even experienced clinicians struggle to observe the disease [4]. This situation illustrates the substantial public health burden imposed by PDAC, which not only reflects its high prevalence but also the limited efficacy of existing treatment modalities [5].

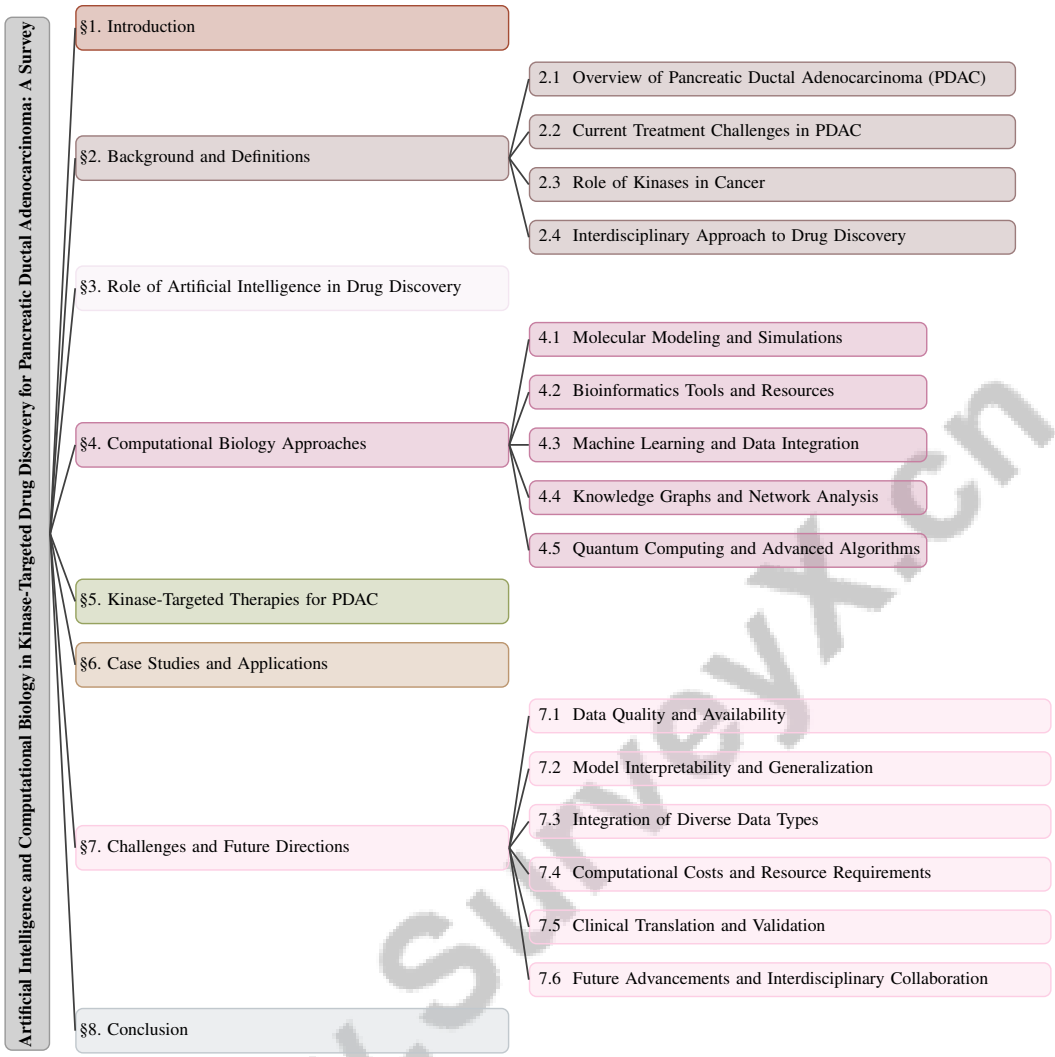


Figure 1: chapter structure

1.2 Potential of AI and Computational Biology in Drug Discovery

The incorporation of artificial intelligence (AI) and computational biology into drug discovery marks a significant advancement, particularly for complex diseases like PDAC. AI's capacity to analyze extensive datasets has transformed the identification of molecular patterns, enhancing drug discovery efficiency [6]. This is particularly vital for PDAC, where traditional methods face high failure rates and lengthy timelines [7]. The application of machine learning (ML) and deep learning (DL) technologies further emphasizes AI's potential in healthcare and drug discovery, addressing critical knowledge gaps [8].

AI techniques, including personalized sequential decision-making, are increasingly utilized in the pharmaceutical sector to customize drug development processes, thereby improving treatment precision and effectiveness [9]. The integration of genetic algorithms with deep learning models has also demonstrated the ability to generate novel kinase inhibitors targeting tyrosine kinases, enhancing bioactivity predictions [10].

Computational biology complements these advancements by providing a framework to understand the structural and functional dynamics of biological systems, essential for the rational design of kinase inhibitors [11]. The incorporation of graph machine learning (GML) into the drug development pipeline enhances target identification, small molecule design, and drug repurposing, offering a comprehensive view of molecular interactions within the kinome [12]. However, challenges such as

AI-generated, for reference only.

the scarcity of AI-ready datasets and standardized knowledge representations persist, necessitating methodological refinements to sustain AI's pivotal role in advancing drug discovery for PDAC [11].

Quantum machine learning frameworks have been proposed to tackle challenges in ligand-based virtual screening for emerging diseases, suggesting the potential of combining quantum computing with classical machine learning methods to improve drug candidate identification [13]. Additionally, computer-aided drug discovery (CADD) initiatives aim to develop efficient computational methods to reduce the time and cost of new drug development, which currently averages around $2.6 billion with a success rate below 10\%$[14]. $These advancements illustrate the transformative potential of AI and compu$

1.3 Structure of the Survey

This survey is structured to comprehensively explore the role of artificial intelligence (AI) and computational biology in kinase-targeted drug discovery for pancreatic ductal adenocarcinoma (PDAC). The introduction emphasizes the critical health challenge posed by PDAC, a cancer known for its late diagnosis and poor prognosis. It discusses AI and computational biology's transformative potential in enhancing drug discovery processes, particularly in PDAC management, including applications in patient risk stratification, early detection, and treatment outcome prediction, while addressing the challenges of integrating AI into clinical practice [15, 16, 17, 18]. The introduction concludes with an outline of the survey structure.

The second section provides background and definitions, including an overview of PDAC, current treatment challenges, and the pivotal role of kinases in cancer biology. It defines key concepts such as AI, computational biology, and kinase inhibitors, highlighting the interdisciplinary approach essential for advancing drug discovery.

Subsequent sections focus on the specific roles of AI and computational biology in drug discovery. Section three details the transformative impact of AI, including various techniques and applications in kinase-targeted drug discovery, while section four discusses computational biology approaches, such as molecular modeling, simulations, bioinformatics tools, and the integration of machine learning techniques.

The survey reviews the current state of kinase-targeted therapies for PDAC in section five, identifying challenges and opportunities for innovation. Section six presents case studies and real-world applications, illustrating the successful integration of AI and computational biology in drug discovery efforts.

Finally, section seven addresses challenges and future directions, discussing data quality, model interpretability, integration of diverse data types, and the clinical translation of AI-driven discoveries. The survey concludes by summarizing key points, emphasizing the potential impact of AI and computational biology on drug discovery for PDAC, and highlighting the importance of continued interdisciplinary collaboration.The following sections are organized as shown in Figure 1.

2 Background and Definitions

2.1 Overview of Pancreatic Ductal Adenocarcinoma (PDAC)

Pancreatic ductal adenocarcinoma (PDAC) is one of the most lethal cancers, primarily due to oncogenic KRAS mutations that drive tumorigenesis, contributing to its poor prognosis and high mortality rates [19, 2]. The absence of effective biomarkers complicates early detection and diagnosis [17]. The tumor microenvironment, notably its interaction with the peripheral nervous system (PNS), plays a critical role in PDAC progression [20]. Despite advancements, early detection remains elusive due to subtle early-stage PDAC features on imaging [4], and traditional radiomics often fail to capture medical image nuances, limiting their predictive power [5]. Early diagnosis is crucial as surgical resection remains the most effective treatment, albeit with variable outcomes [21]. The scarcity of clinical data for deep learning model training further hinders reliable diagnostic tool development [22]. Gene microarray datasets have identified potential targets, yet effective drug candidates with minimal adverse effects are still needed [1, 23]. The aggressive nature of PDAC necessitates accurate differential diagnosis of pancreatic cysts to prevent progression to PDAC [24].

2.2 Current Treatment Challenges in PDAC

PDAC treatment is fraught with challenges due to late-stage diagnosis and its aggressive nature. Traditional drug discovery is inefficient and costly, especially during early virtual screening for target protein-binding compounds [13]. The small size and indistinct tumor boundaries further complicate detection [25]. Resistance to receptor tyrosine kinase inhibitors (RTKIs) often leads to treatment discontinuation due to toxicity, highlighting the need for selective agents with minimal off-target effects [26, 10]. Traditional drug discovery often fails to ensure in vivo efficacy, necessitating integrative approaches [27]. The resource-intensive drug discovery process, taking over a decade and costing billions, often relies on biased datasets, limiting generalization [11, 14]. The poor predictive performance of ADMET systems for novel compounds is a critical issue due to limited labeled datasets [28]. Current imaging techniques, such as IVIM, face challenges like poor image quality and lengthy fitting times [29], and automated methods struggle with CE-CT imaging for accurate prognosis [21]. These challenges necessitate innovative AI and computational biology approaches to improve diagnostic and therapeutic outcomes in PDAC.

2.3 Role of Kinases in Cancer

Kinases are vital in cancer biology, regulating growth, differentiation, and apoptosis [30]. They mediate signal transduction through ATP-dependent phosphorylation, often dysregulated in cancer, leading to uncontrolled proliferation [31]. Kinases are thus prime targets for cancer therapy, with numerous inhibitors developed [32]. Designing selective inhibitors is challenging due to the conserved ATP-binding site across kinases, often causing off-target effects [33]. Resistance through kinase domain mutations, such as in EGFR, necessitates next-generation inhibitors [34]. The complexity of kinase signaling, with compensatory mechanisms and network interactions, complicates therapeutic outcome predictions [35]. Kinase inhibitors are classified by binding characteristics, aiding in rational design for specificity and efficacy [26]. Innovations in computational methods, like genetic algorithms integrated with deep learning, enhance kinase inhibitor activity prediction and evaluation [10]. Despite progress, many kinases remain underexplored, offering potential for novel therapies. Tyrosine kinase inhibitors (TKIs) underscore the importance of these enzymes, though development and application face hurdles [36]. Innovative approaches, including AI and computational biology, are crucial for advancing kinase-targeted cancer therapies.

2.4 Interdisciplinary Approach to Drug Discovery

The integration of AI, computational biology, and traditional biology is transforming drug discovery for complex diseases like PDAC. This interdisciplinary approach leverages each field's strengths to address drug development challenges. AI enhances drug discovery by combining deep learning with radiomics to improve prognostic accuracy for PDAC [3]. Massively multitask neural networks provide a robust framework for evaluating models in drug discovery, showcasing AI's capability in managing complex data [7]. Computational biology offers integrated modeling pipelines, streamlining the drug discovery process [11]. The complexity of biological systems and the need for data integration highlight computational biology's role in elucidating molecular interactions [12]. Advances in computational methods for kinase-inhibitor binding kinetics further emphasize computational approaches' significance [37]. Traditional biology provides essential insights into disease mechanisms and targets. Collaboration among computational scientists, biologists, and chemists is crucial for translating predictions into therapeutic strategies. This is exemplified by advanced ADMET systems integrating self-supervised and multi-task learning for robust drug discovery [28]. Prototype-based methods, like Conditional Diversity Networks (CDN), generate diverse, novel drug-like molecules [38]. This interdisciplinary approach is vital for advancing therapies for challenging diseases like PDAC, integrating AI, computational biology, and traditional insights to deepen understanding of biological interactions and pave the way for innovative interventions. AI-driven methodologies, such as multimodal deep learning and explainable machine learning, facilitate transitioning novel candidates from models to clinical applications, accelerating effective treatment development [15, 39, 18, 40].

3 Role of Artificial Intelligence in Drug Discovery

Artificial intelligence (AI) has become a pivotal force in drug discovery, addressing the complexities of biological systems and expansive chemical spaces through sophisticated computational techniques.

Category	Feature	Method
AI Techniques in Drug Discovery	3D Data Processing	HS[14]
	Generative Techniques	CDN[38]
AI in Kinase-Targeted Drug Discovery	Predictive Analysis	CE-ConvLSTM[21]
	Targeted Drug Design	DPPSG[41]
Data-Driven Predictions and Modeling	Multimodal Approaches	FN[4]
	Data Imbalance Handling	WkNNIR[42]
Challenges and Innovations in AI Applications	Interactive Learning Approaches	AHD[43]
	Training Strategies	H-ADMET[28]
	Quantum Computing Integration	QSVC[13]
	Attention-Based Techniques	IAG-Net[25]

Table 1: This table provides a comprehensive overview of various artificial intelligence (AI) techniques and methodologies employed in drug discovery. It categorizes these techniques into specific domains such as AI techniques in drug discovery, AI in kinase-targeted drug discovery, data-driven predictions and modeling, and challenges and innovations in AI applications, highlighting key features and methods associated with each category. The table serves to elucidate the diverse applications of AI in enhancing drug discovery processes and overcoming associated challenges.

This section explores AI methodologies in drug discovery, highlighting their contributions to improved predictions, process optimization, and novel therapeutic agent identification. Table 1 presents a detailed categorization of AI methodologies utilized in drug discovery, underscoring their significance in advancing therapeutic development and addressing existing challenges. Additionally, Table 4 provides a detailed comparison of AI methodologies in drug discovery, illustrating their diverse applications and advantages in advancing therapeutic development. **??** illustrates the hierarchical structure of AI's role in drug discovery, emphasizing key techniques and applications, particularly in kinase-targeted therapies. The figure further delineates data-driven predictions alongside the challenges and innovations associated with AI applications. By outlining the integration of predictive and generative models, AI-driven platforms, and advanced modeling techniques, it underscores AI's transformative potential in enhancing drug discovery processes.

3.1 AI Techniques in Drug Discovery

Method Name	Modeling Techniques	Technological Integration	Application Outcomes
HS[14]	3D Convolutional Neural	Docking And Machine	Improved Prediction Accuracy
CDN[38]	Variational Autoencoder	Machine Learning	Novel Drug-like
QSVC[13]	Quantum Kernel Methods	Quantum Integrated Workflow	Improved Classification Accuracy

Table 2: Summary of AI-based methodologies in drug discovery, detailing the modeling techniques, technological integration, and application outcomes. The table highlights the effectiveness of various AI approaches, including 3D convolutional neural networks, variational autoencoders, and quantum kernel methods, in improving prediction and classification accuracy in drug development.

AI significantly enhances drug discovery by employing diverse techniques to predict molecular properties, generate drug candidates, and understand drug-target interactions. The integration of genetic algorithms and deep learning models has facilitated the creation of novel tyrosine kinase inhibitors, improving bioactivity predictions [10, 11]. Graph machine learning (GML) models biological interactions, leveraging structural and relational data to enhance drug discovery predictions [12]. The HydraScreen framework, utilizing a 3D convolutional neural network, effectively represents molecular structures and interactions, offering a comprehensive approach [14].

Prototype-based generative models, such as Conditional Diversity Networks (CDN), expedite drug discovery by producing diverse candidate molecules that share desired characteristics with prototype drugs [38]. Quantum computing techniques, including the Quantum Support Vector Classifier (QSVC), improve compound classification by merging classical data processing with quantum capabilities [13]. These AI techniques exemplify AI's transformative potential in the pharmaceutical industry, enhancing the efficiency and accuracy of identifying novel therapeutic agents while deepening our understanding of complex molecular interactions [18, 44].

Table 2 provides a comprehensive overview of the AI techniques applied in drug discovery, illustrating the diverse modeling strategies and technological integrations employed to enhance application outcomes. Figure 2 illustrates the hierarchical categorization of AI techniques in drug discovery, emphasizing molecular prediction, drug candidate generation, and understanding interactions. The

first image depicts neural networks determining the relevance of input features in predicting drug outcomes, while the second image provides a schematic of a neural network's architecture, highlighting data flow and layer connections. These examples underscore AI's potential to revolutionize drug discovery by enabling complex dataset analysis and key molecular target identification with unprecedented precision and speed [45, 46]. Through these advanced computational techniques, AI is poised to significantly enhance drug discovery processes.

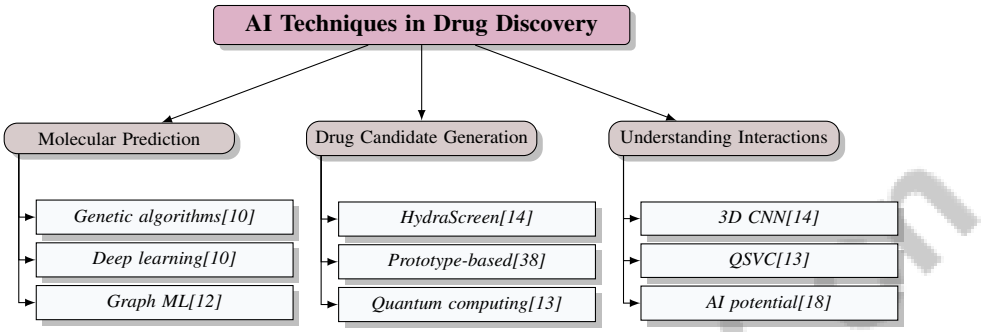


Figure 2: This figure illustrates the hierarchical categorization of AI techniques in drug discovery, emphasizing molecular prediction, drug candidate generation, and understanding interactions. Key AI methods include genetic algorithms, deep learning, and graph machine learning, along with advanced techniques like HydraScreen, prototype-based models, and quantum computing, highlighting AI's transformative role in enhancing drug discovery processes.

3.2 AI in Kinase-Targeted Drug Discovery

AI is integral to discovering kinase-targeted therapies, enhancing drug-kinase interaction prediction and novel inhibitor design. The E3Inv-ARaint diffusion model integrates pocket-specific information into peptide generation, crucial for interaction modeling [41]. Advanced deep learning models like CE-ConvLSTM improve predictive accuracy by learning tumor attenuation patterns from CE-CT imaging, informing PDAC therapeutic strategies [21].

Generative AI techniques generate novel molecules with high binding affinity and optimal ADME/PK properties, significantly increasing high-affinity molecule proportions compared to traditional methods [47, 18, 44]. Platforms like iMiner streamline drug discovery by ensuring generated molecules meet pharmacological criteria. AI-driven platforms like AlphaFold accelerate kinase inhibitor identification by predicting protein structures, exemplifying AI's role in enhancing drug discovery processes [15, 18, 40, 48, 16].

3.3 Data-Driven Predictions and Modeling

AI-driven data analysis and modeling advance drug discovery by providing sophisticated methodologies for predicting drug-target interactions (DTI) and elucidating complex biological systems. Large, standardized datasets enhance DTI prediction performance, improving drug discovery efforts [49]. Techniques like WkNNIR predict interactions by estimating the influence of similar drugs and targets, allowing predictions for novel drugs and targets [42]. FusionNet exemplifies AI's ability to improve detection accuracy by integrating multimodal data [4]. The DTIAM framework enhances predictions by learning from unlabeled data, improving drug-target interaction understanding [50]. Massively multitask networks significantly enhance predictive accuracy by promoting data sharing [7]. These advancements illustrate AI's transformative potential in drug discovery, enabling efficient exploration of chemical spaces and improving therapeutic prediction accuracy [18, 6].

3.4 Challenges and Innovations in AI Applications

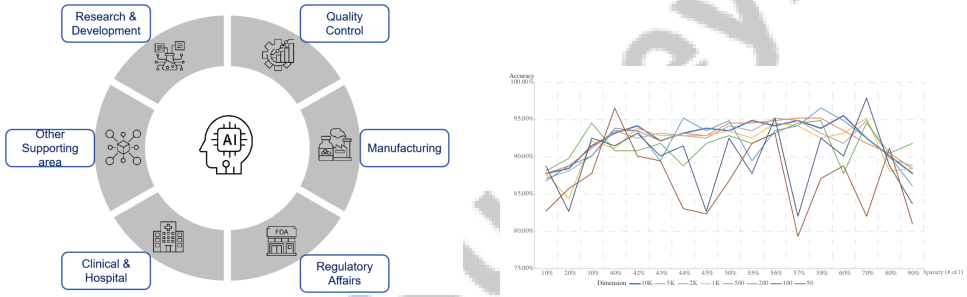
Integrating AI into drug discovery presents challenges, particularly computational demands for accurate predictions and variability in results due to differing methods and parameters [37]. Data scarcity for various ADMET endpoints often leads to overfitting and poor generalization [28]. Innovations like IAG-Net, which employs attention mechanisms within a multiple instance learning

Method Name	Computational Demands	Data Scarcity	Innovative Methodologies
H-ADMET[28]	-	Labelled Data Scarcity	Self-supervised Learning
IAG-Net[25]	-	Limited Voxel-level	Attention Mechanisms
QSVC[13]	Faster Training Times	Smaller Datasets	Quantum Kernel Methods
AHD[43]	Extensive Computational Resources	-	Reinforcement Learning

Table 3: Comparison of AI Methods Addressing Computational and Data Challenges in Drug Discovery. This table outlines the computational demands, data scarcity issues, and innovative methodologies of various AI methods, such as H-ADMET, IAG-Net, QSVC, and AHD, which are applied in drug discovery to enhance prediction accuracy and efficiency.

framework, improve classification and segmentation performance [25]. Quantum machine learning approaches, such as the QSVC, offer improved classification accuracy and faster training times [13].

These advancements highlight AI's transformative potential in drug discovery, enabling precise therapeutic interventions and significantly enhancing drug development efficiency. AI technologies streamline drug discovery, achieving up to a ten-fold reduction in the time required for drug molecule identification and facilitating safe and effective therapy development [15, 16, 51, 18]. Addressing computational limitations and data scarcity through innovative methodologies, AI continues to drive significant advancements in drug discovery. Table 3 provides a comparative analysis of various AI methods employed in drug discovery, highlighting their computational requirements, challenges related to data scarcity, and the innovative methodologies they incorporate.



(a) AI in Healthcare: A Comprehensive Overview[51]

(b) Accuracy[43]

Figure 3: Examples of Challenges and Innovations in AI Applications

Figure 3 demonstrates AI's critical role in drug discovery, presenting groundbreaking innovations alongside significant challenges. Figure 1(a) provides an overview of AI's impact across healthcare sectors, emphasizing AI's integration with human cognition and its role in accelerating drug discovery processes. Figure 1(b) focuses on AI application challenges, particularly model accuracy across varying dataset sparsity levels, essential for successful AI implementation in drug discovery [51, 43].

Feature	AI Techniques in Drug Discovery	AI in Kinase-Targeted Drug Discovery	Data-Driven Predictions and Modeling
Method Type	Predictive Models	Generative Models	Data Analysis
Application Focus	Molecular Properties	Kinase Inhibitors	Drug-target Interactions
Key Advantage	Improved Bioactivity Predictions	High Binding Affinity	Enhanced Prediction Accuracy

Table 4: Comparison of AI Techniques in Drug Discovery: This table categorizes various AI methodologies utilized in the field of drug discovery, highlighting their method types, application focuses, and key advantages. The comparison underscores the significant role of AI in improving bioactivity predictions, enhancing binding affinity, and increasing prediction accuracy across different drug discovery applications.

4 Computational Biology Approaches

Computational biology has emerged as a cornerstone in drug discovery, employing sophisticated methodologies to address complex biological challenges. This section delves into various approaches,

beginning with molecular modeling and simulations, which are pivotal for understanding kinase-inhibitor interactions and informing therapeutic strategies.

4.1 Molecular Modeling and Simulations

Molecular modeling and simulations are indispensable for elucidating kinase interactions, offering insights into structural and dynamic processes influencing drug-target affinity. These techniques employ advanced algorithms to predict and analyze interactions, particularly in diseases like pancreatic ductal adenocarcinoma (PDAC). The IMPECCABLE pipeline, integrating machine learning with physics-based simulations, exemplifies frameworks that enhance drug discovery by evaluating inhibitor binding interactions with kinases [11].

Recent advancements, including AI and microfluidics, augment molecular modeling capabilities [26]. AI platforms like KekuleScope utilize convolutional neural networks for bioactivity modeling from compound images, while frameworks like PGraphDTA enhance binding affinity predictions using protein language models and contact maps. In PDAC research, deep learning frameworks such as CE-ConvLSTM improve survival predictions by analyzing CE-CT imaging data [21]. Attention-guided frameworks like IAG-Net further refine predictions and tumor segmentation using dual-stream architectures [25].

Ongoing advancements in molecular modeling, particularly through AI integration, significantly streamline therapeutic compound identification and optimization, improving pharmacological property predictions. These innovations play an increasingly crucial role in developing effective therapies, addressing industry challenges, and enhancing patient outcomes [52, 48, 16, 53, 47].

4.2 Bioinformatics Tools and Resources

Bioinformatics tools are vital for kinase-targeted drug discovery, providing platforms for analyzing complex biological data and identifying therapeutic targets. As illustrated in Figure 4, these tools are hierarchically categorized to highlight the various computational methods, hybrid approaches, and AI-driven biomarker discovery utilized in the field. These tools employ various computational methods for predicting drug-drug interactions (DDI), including chemical structure-based, network-based, NLP-based, and hybrid approaches [54]. Structure-based methods assess compound similarities, while network-based approaches leverage graph theory to model biological interactions, facilitating drug target identification [55, 56, 57, 27, 58]. NLP-based methods extract insights from biomedical literature, identifying novel interactions and therapeutic opportunities.

Hybrid methods integrate diverse approaches for comprehensive DDI analysis, enhancing prediction accuracy. By capturing complex relationships among drugs, cell lines, and diseases, these methods leverage advanced modeling techniques to provide nuanced insights into therapeutic synergies, improving drug discovery and repurposing [15, 59, 60]. Bioinformatics tools are crucial for navigating chemical space and biological complexity, enabling efficient candidate identification and prioritization.

The ongoing advancement and integration of AI, ML, and bioinformatics tools significantly enhance kinase interaction understanding and biomarker identification, facilitating improved patient stratification and early detection in PDAC. By analyzing gene expression and protein interactions, researchers uncover critical pathways and hub genes as potential diagnostic and therapeutic targets, addressing PDAC management complexities effectively [1, 17].

4.3 Machine Learning and Data Integration

Machine learning (ML) integration with computational biology has advanced drug discovery by enhancing molecular interaction understanding and optimizing drug design. The ImDrug framework exemplifies ML's potential, offering novel evaluation metrics and benchmarking for deep imbalanced learning, ensuring model robustness in drug discovery [61]. Graph-based methods like GraphIX leverage biopharmaceutical knowledge graphs for disease-drug association predictions, emphasizing diverse data integration for enhanced predictive capabilities [62].

Innovations in NLP methods, emphasizing the Extract-Fuse-Predict workflow, enhance model creation and interpretation, underscoring data fusion's role in predictive accuracy [63]. Advanced ML

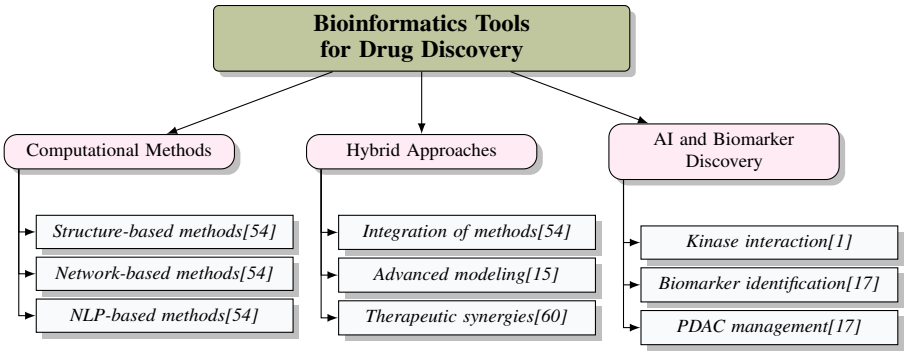


Figure 4: This figure illustrates the hierarchical categorization of bioinformatics tools and methods used in kinase-targeted drug discovery, highlighting computational methods, hybrid approaches, and AI-driven biomarker discovery.

techniques integrating pharmacokinetic data generation illustrate AI's application in optimizing drug discovery [64]. Iterative chemical structure generation, as described by Filella-Merce et al., merges molecular modeling with ML to optimize drug design, enhancing viable candidate production [47].

Causal inference techniques distinguish spurious from non-spurious associations in biological data, improving drug-target interaction identification [65]. Adaptive Sampling Optimization (ASO) enhances ML model efficiency by focusing resources on promising chemical space areas [66]. Conformal prediction (CP) integration and Physical formula Enhanced Multitask Learning (PEMAL) models further demonstrate ML's flexibility and domain-specific knowledge incorporation in drug discovery [67, 68].

Recent ML advancements in computational biology underscore their transformative potential in drug discovery, enhancing data-driven decision-making across development stages. Despite challenges like data quality and result interpretability, integrating these technologies drives therapeutic innovation and reduces discovery failure rates [69, 60, 70].

4.4 Knowledge Graphs and Network Analysis

Knowledge graphs (KGs) and network analysis are crucial for drug discovery, particularly in identifying targets and elucidating biological interactions. KGs enhance drug-target interaction predictions by modeling gene-disease and drug-target relationships, providing a comprehensive framework for understanding biological networks [53, 71].

KGs categorize research and datasets based on task relevance, aiding systematic biological data exploration and enhancing novel target identification and drug repurposing [72]. Network analysis complements KGs by visualizing interactions within biological systems, identifying key nodes and pathways as potential drug targets.

Integrating network analysis with KGs represents a transformative strategy for deciphering disease mechanisms and therapeutic responses. By leveraging KGs, researchers gain a comprehensive understanding of molecular pathways, streamlining drug discovery processes like target identification and toxicity prediction. This approach addresses critical biomedical research challenges, improving patient healthcare outcomes [53, 72].

KG and network analysis integration in drug discovery empowers researchers to navigate biological interactions effectively, facilitating tasks like repurposing, toxicity prediction, and target prioritization. Knowledge graph embeddings (KGEs) enhance predictive performance in target identification and decision-making, leading to efficient lab-based experiments and improved patient outcomes. By leveraging public datasets and understanding KGE performance factors, researchers construct high-quality KGs, addressing discovery challenges and paving the way for innovative biomedical solutions [72, 53].

4.5 Quantum Computing and Advanced Algorithms

Quantum computing and advanced algorithms promise to revolutionize computational biology by enhancing complex calculation efficiency and accuracy in drug discovery. The hedQM method exemplifies this by utilizing parallel density functional theory (DFT) algorithms for efficient quantum mechanical calculations on protein-ligand complexes [73].

Integrating quantum simulations with machine learning, as demonstrated by QMLS, showcases quantum computing's potential in drug candidate generation and optimization [74]. Quantum Circuit Born Machine (QCBM) and classical LSTM networks illustrate quantum computing's ability to handle complex data distributions, enabling precise drug-target interaction modeling [75]. The Quantum Protein Similarity Algorithm (QPSA) enhances protein sequence analysis, contributing to drug discovery [76].

Quantum computing's prospects in computational biology are underscored by its ability to perform previously infeasible computations, accelerating methods and opening new research avenues [77]. Hybrid quantum computing pipelines, as detailed by Li et al., emphasize practical applications in drug design, bridging theoretical advancements with real-world challenges [78].

A novel framework categorizing quantum machine learning research highlights quantum computing's potential applications in drug discovery, facilitating systematic exploration of quantum algorithms [79]. Quantum computing integration in computational biology signifies transformative advancement, unlocking unprecedented capabilities for drug discovery. Its ability to process vast datasets and execute complex simulations enhances molecular property predictions and novel compound design. Hybrid quantum-classical approaches address real-world drug design challenges, demonstrating potential to streamline workflows and produce viable candidates, as evidenced by effective KRAS inhibitor identification. While promising, awareness of quantum computing's challenges and limitations remains essential [75, 77, 76, 78, 79]. By harnessing quantum mechanics, researchers achieve detailed molecular interaction understanding, paving the way for targeted therapy development.

5 Kinase-Targeted Therapies for PDAC

5.1 Current State of Kinase-Targeted Therapies

Kinase-targeted therapies offer a promising strategy for treating pancreatic ductal adenocarcinoma (PDAC) by inhibiting specific kinases to disrupt tumorigenic pathways [30]. The development of selective kinase inhibitors, such as those in the Kinase Chemogenomic Set (KCGS), underscores their potential to address PDAC's molecular heterogeneity [32]. FDA-approved inhibitors, categorized into Type I and Type II, exhibit diverse binding modes and mechanisms, enabling therapies tailored to PDAC's unique molecular aberrations [80, 31].

Advancements in computational biology and AI have accelerated kinase-targeted therapy development. Genetic algorithms combined with deep learning have improved drug discovery efficiency, facilitating the design of novel tyrosine kinase inhibitors and predicting their bioactivity [10]. AI methodologies, such as the 3DGAUnet model, enhance the synthesis of clinical tumor CT images, improving diagnostic accuracy and informing treatment strategies [22]. However, resistance to receptor tyrosine kinase inhibitors (RTKIs) remains a significant challenge, necessitating innovative strategies to enhance long-term efficacy [26].

Diagnostic advancements, including IVIM-NET, offer superior accuracy in parameter estimates compared to traditional imaging, while improved methodologies for pancreatic cyst diagnosis highlight potential for early detection and intervention [29, 24]. The current landscape of kinase-targeted therapies for PDAC reflects substantial advancements driven by AI and machine learning, emphasizing the need for personalized treatment strategies to address tumor heterogeneity and drug resistance, ultimately improving outcomes for patients with this aggressive cancer [36, 17, 31, 26, 30].

5.2 Challenges in Developing Selective Kinase Inhibitors

Developing selective kinase inhibitors for PDAC involves navigating challenges such as the vast combinatorial space of drug-like molecules to identify viable candidates [11]. Limitations in computational methods, including force field selection and molecular dynamics sampling, impact the accuracy

of predicted interaction rates [37]. Predicting complex pharmacokinetic endpoints, particularly oral bioavailability, remains challenging due to multifaceted physiological processes [28].

Predictive model development is hindered by inconsistent benchmarks, plagued by issues like chemical representation and noisy data. The WelQrate benchmark aims to address these deficiencies by providing reliable standards for model evaluation [81]. The cold start problem persists, as models struggle to predict interactions for novel drugs or targets due to insufficient training data, requiring extensive dataset expansion at considerable costs.

The pharmaceutical industry's reluctance to share proprietary data limits collaboration necessary for robust predictive model development [65]. This impedes progress in multi-target drug design, crucial for effective kinase inhibitor development. Innovative approaches like HydraScreen, integrating pose and affinity predictions, enhance scoring and docking capabilities while providing practical applications [14]. Nonetheless, challenges in sampling chemical space and predicting complex endpoints highlight the need for continued methodological advancements and interdisciplinary collaboration to refine selective kinase inhibitor development for PDAC.

5.3 Opportunities for Innovation

AI and computational biology integration presents significant opportunities for innovation in kinase-targeted therapies, particularly in drug discovery and development. AI-driven frameworks, such as AIAltMed, have advanced the identification of compounds with therapeutic potential, enhancing drug discovery and repurposing efforts [15]. Zielinski's work emphasizes AI's transformative role in addressing the productivity crisis in pharmaceutical research [16].

Graph-based approaches, exemplified by GraphIX, demonstrate impressive prediction accuracy and explainability in drug repositioning applications, refining kinase-targeted drug discovery [62]. These methods leverage structural and relational data inherent in biological systems to enhance therapeutic development precision. The Tx-LLM model signifies a major advancement in using AI for therapeutic development, offering a versatile model capable of addressing multiple aspects of the drug discovery pipeline [82].

ML techniques show promise in enhancing the design process for PROTACs, potentially leading to drug-like properties and targeting previously undruggable proteins [83]. This innovation broadens kinase-targeted therapies' scope, addressing a wider range of therapeutic targets. The hybrid quantum computing pipeline demonstrated by Li et al. illustrates quantum models' potential for comparable or improved accuracy in energy calculations, enhancing computational efficiency and accuracy in drug discovery [78].

Federated learning frameworks, such as Pejo's, are crucial for identifying and mitigating privacy risks, facilitating collaborative drug discovery while ensuring data security [84]. This approach enhances data sharing across institutions, accelerating kinase inhibitor discovery. Liu et al.'s docking-based virtual screening method emphasizes innovation opportunities by leveraging existing target knowledge to improve predictive performance for new targets, enhancing kinase-targeted drug discovery efficiency [85].

Knowledge-augmented graph machine learning (KaGML) techniques, as described by Zhong, enhance interpretability and demonstrate the ability to work with limited training data, suggesting practical applications in drug discovery [57]. Quantum computing's potential, as surveyed by Outeiral et al., offers exponential speedups for specific tasks, such as matrix inversion and optimization, indicating its revolutionary potential in computational biology [77].

6 Case Studies and Applications

6.1 Case Studies and Real-World Applications

The integration of artificial intelligence (AI) and computational biology in kinase-targeted drug discovery is demonstrated by several case studies illustrating their transformative potential in therapeutic development. Federated learning frameworks, for instance, enable collaborative drug discovery across institutions while maintaining data privacy, though privacy risks remain a concern [84]. In pancreatic ductal adenocarcinoma (PDAC), the IAG-Net framework significantly improves segmentation accuracy, as evidenced by a more than 5% increase in the Dice Similarity Coefficient (DSC) over

existing methods [25]. This advancement highlights the framework's potential in clinical settings where precise tumor segmentation is critical for effective treatment planning.

These case studies underscore the profound impact of AI and computational biology on kinase-targeted drug discovery. By employing advanced methodologies, such as AI-driven platforms and specialized large language models for oncology, researchers can enhance therapeutic precision and efficiency. Such innovations are pivotal in developing safer, more effective treatments, addressing pharmaceutical industry challenges, and improving outcomes for patients with previously untreatable conditions [15, 16, 86]. Continued exploration and application of these technologies are essential for overcoming drug discovery challenges and enhancing patient outcomes.

6.2 Multi-phase CT Scans for PDAC Detection

Advanced imaging techniques, particularly multi-phase computed tomography (CT) scans, play a crucial role in detecting pancreatic ductal adenocarcinoma (PDAC). These modalities provide vital insights into the structural and functional characteristics of pancreatic tumors, facilitating early and accurate diagnosis [19]. By capturing detailed images across various contrast enhancement phases, multi-phase CT scans improve the detection and characterization of PDAC lesions. Their effectiveness lies in revealing subtle tissue attenuation variations, essential indicators of malignant transformation. Utilizing both arterial and venous phases enhances PDAC detection and segmentation, thereby improving clinical outcomes through advanced imaging and deep learning techniques [19, 87].

Integrating multi-phase CT imaging with advanced computational techniques, such as machine learning algorithms, further augments diagnostic accuracy. These algorithms analyze complex imaging data to identify patterns associated with PDAC, enhancing clinical decision-making. The combination of multi-phase CT scans with advanced analysis techniques, including alignment ensemble strategies and deep learning frameworks, improves PDAC detection and segmentation. This approach facilitates early diagnosis and enhances the accuracy of tumor measurements through automated segmentation, demonstrating significant performance improvements via self-learning frameworks that utilize both annotated and unannotated images [19, 87].

6.3 Self-learning Segmentation Methods

Self-learning segmentation methods are pivotal in pancreatic ductal adenocarcinoma (PDAC) research, offering enhanced accuracy and efficiency in tumor delineation. By leveraging advanced machine learning algorithms, these methods autonomously learn from data, significantly reducing reliance on manual annotations and expert input [25]. The integration of self-learning techniques into frameworks like IAG-Net has shown substantial improvements in segmentation performance, particularly in metrics such as the Dice Similarity Coefficient (DSC) [25]. These methods are particularly beneficial given the complex and heterogeneous nature of pancreatic tumors. They utilize iterative learning processes to refine segmentation accuracy, adapting to the diverse morphological characteristics of PDAC lesions. By incorporating feedback mechanisms and extensive datasets, these models iteratively enhance their predictive accuracy, improving tumor boundary delineation, especially in challenging cases like PDAC [25, 87, 88].

The synergy between self-learning segmentation methods and multi-phase CT imaging further enhances their effectiveness. The rich, multi-dimensional data provided by these imaging techniques enables comprehensive tumor feature analysis, allowing self-learning models to achieve higher accuracy levels in detecting and segmenting PDAC [19]. This integration represents a significant advancement in PDAC research, facilitating early diagnosis and informed treatment planning.

6.4 Quantum Machine Learning in Drug Discovery

Quantum machine learning (QML) is transforming drug discovery by leveraging quantum computing principles to enhance the efficiency and accuracy of predictive models. The synergy between quantum computing and machine learning offers substantial computational advantages, transforming drug discovery processes by facilitating precise exploration of complex molecular interactions and optimizing drug candidates through quantum algorithms and hybrid quantum-classical approaches. Quantum computing can efficiently process vast datasets and run sophisticated simulations, enabling accurate predictions of molecular properties and interactions. Recent studies demonstrate that

this innovative approach accelerates the identification of potential drug candidates and improves drug design simulations, showcasing its transformative potential in the pharmaceutical industry [75, 77, 13, 78, 79].

The Quantum Circuit Born Machine (QCBM), in conjunction with classical Long Short-Term Memory (LSTM) networks, exemplifies QML's capability to handle complex data distributions and sequential chemical data, enabling precise modeling of drug-target interactions [75]. Additionally, the Quantum Protein Similarity Algorithm (QPSA) enhances the efficiency and accuracy of protein sequence generation and similarity analysis, contributing to a deeper understanding of protein functions and interactions crucial for drug discovery [76]. The hedQM method, utilizing parallel density functional theory (DFT) algorithms optimized for cloud computing, illustrates quantum computing's potential to accelerate molecular interaction exploration, providing a robust framework for understanding kinase-targeted therapies [73]. Similarly, the integration of quantum simulations with machine learning, as demonstrated by QMLS, underscores quantum computing's transformative capabilities in generating and optimizing potential drug candidates [74].

These advancements highlight QML's transformative potential in drug discovery, enabling efficient exploration of chemical spaces and improving therapeutic predictions' accuracy. By applying quantum mechanics principles, researchers can significantly enhance their understanding of molecular interactions, leading to the development of more precise and effective therapies. Hybrid quantum-classical models have successfully identified promising inhibitors for critical cancer targets like KRAS, demonstrating practical applications of quantum computing in accelerating drug discovery and improving therapeutic outcomes [75, 77, 13, 73, 79].

6.5 Knowledge Graphs in Drug Repurposing and Target Identification

Knowledge graphs (KGs) have emerged as powerful tools in drug discovery, particularly for repurposing existing drugs and identifying novel therapeutic targets. By structuring complex biological data into interconnected nodes and edges, KGs facilitate the visualization and analysis of relationships between drugs, targets, and diseases [72]. This capability is crucial in drug repurposing, where the aim is to identify new therapeutic uses for existing drugs, thereby accelerating the drug development process and reducing costs. The utility of KGs in drug repurposing is exemplified by their ability to integrate diverse datasets, including genomic, proteomic, and pharmacological information, uncovering hidden relationships that may not be apparent through traditional methods [53]. This integration allows researchers to systematically explore the potential of existing drugs to interact with new targets, providing a robust framework for hypothesis generation and validation.

In target identification, KGs offer a comprehensive approach to mapping complex interactions within biological systems. By leveraging graph-theoretic algorithms, researchers can identify key nodes and pathways that may serve as potential drug targets, streamlining the drug discovery process [71]. This approach enhances target identification precision and provides insights into the mechanistic pathways underlying disease pathology. Furthermore, applying KGs in combination with advanced computational techniques, such as machine learning and network analysis, enhances their predictive capabilities. These hybrid approaches enable the identification of novel drug-target interactions with greater accuracy, facilitating new therapeutic opportunities [53]. By offering a holistic view of the molecular landscape, KGs play a crucial role in advancing drug repurposing and target identification, ultimately contributing to developing more effective and targeted therapies.

7 Challenges and Future Directions

Addressing the complexities of drug discovery requires a nuanced understanding of factors affecting AI applications. This section explores critical challenges including data quality and availability, which are pivotal for refining AI-driven models and enhancing their integration into drug discovery processes.

7.1 Data Quality and Availability

High-quality, annotated datasets are crucial for AI-driven drug discovery, significantly impacting model predictive accuracy. The scarcity of such datasets often necessitates extensive wet-lab experiments, creating bottlenecks in the discovery pipeline [11]. This is particularly challenging in cold

start scenarios, where models must predict interactions for new drugs or targets with limited prior data. The DTIAM framework addresses this through self-supervised learning, leveraging unlabeled data to improve predictions [50].

Dataset imbalances further complicate drug-target interaction predictions. Effective handling of these imbalances is crucial, as they can lead to overfitting, especially in small datasets [7]. Integrating genetic algorithms with deep learning models underscores the need for diverse datasets to ensure robustness and generalizability [10].

Frameworks like HydraScreen advance structure-based deep learning but face challenges related to computational costs and biases in training datasets, affecting prediction reliability [14]. The vast chemical space and method limitations complicate efficient exploration of drug-like compounds without introducing human intuition biases [38].

The H-ADMET system exemplifies advancements in data quality, showing a 4% improvement in predicting properties for molecules with unobserved scaffolds [28]. Future research should enhance kinetic prediction accuracy and leverage large datasets to improve machine learning models [37]. Expanding evaluation frameworks to include ADMET property metrics is crucial for model robustness and applicability, refining therapeutic strategies in drug discovery.

7.2 Model Interpretability and Generalization

Model interpretability and generalization are critical for AI applications in drug discovery, impacting their clinical utility. The complexity of ML models often results in opaque decision-making, complicating interpretation [89]. This opacity hinders validation, as clinicians require clear insights into prediction derivations to trust these tools.

Robust validation frameworks are necessary to assess model performance across diverse datasets. Support vector machines (SVMs) achieve competitive performance with deep learning models in bioactivity prediction, prompting a reevaluation of field metrics [90]. This underscores the importance of selecting appropriate validation metrics for meaningful model reliability insights.

Generalization challenges arise when models are applied to novel compounds outside the training data. Probabilistic programming models (PPMs) require careful calibration of uncertainty estimates for accurate predictions [91]. Enhancing model interpretability through explainable AI techniques and improving generalization capabilities by integrating diverse datasets are essential for robust performance across drug discovery contexts. Addressing data quality, ethical considerations, and AI limitations will improve AI solution reliability in the pharmaceutical industry, expediting drug discovery and enhancing therapeutic efficacy [18, 16].

7.3 Integration of Diverse Data Types

Integrating diverse data types in computational biology and AI is crucial for advancing drug discovery, especially for kinase-targeted therapies. Future research should focus on hybrid models that incorporate various data types, enhancing interpretability and addressing ethical considerations [8]. Autonomous knowledge updating and expanding applicability to more pharmaceutical tasks can improve AI-driven drug repurposing accuracy [39].

Incorporating multi-omics data and enhancing network analysis tools will facilitate a comprehensive understanding of molecular networks, promoting collaborative data-sharing frameworks [56]. Future work should explore integrating 3D representations to refine computational models for accurate molecular interaction insights [92].

Knowledge graphs (KGs) are vital in this integration process, yet challenges persist in optimally incorporating diverse data sources and representing uncertainty [72]. Addressing these challenges will enhance KGs' robustness and applicability in drug discovery.

The Adaptive Active Learning (AAL) approach, involving adding and deleting data points, offers a strategy for adapting to changing data distributions, improving integration of diverse data types in computational biology and AI [93]. This adaptability is essential for maintaining predictive model relevance and accuracy in dynamic research environments.

Incorporating protein 3D structure information and enhancing model interpretability are critical for improving drug-target interaction modeling accuracy [50]. Establishing standard datasets and performance metrics, alongside exploring alternative featurization methods, will bolster model performance and facilitate diverse data type integration [7].

Successful data type integration in computational biology and AI requires advanced methodologies, refinement of existing tools, and exploration of new applications in drug discovery [12]. Addressing these challenges can enhance therapeutic intervention precision and efficacy, paving the way for more effective oncology treatments.

7.4 Computational Costs and Resource Requirements

Computational costs and resource requirements for AI and computational biology applications pose significant challenges in drug discovery. The complexity of these technologies demands substantial computational resources, hindering scalability and practical implementation. For example, the architecture search process in AutoHDC highlights high computational costs in optimizing network architectures [43]. Molecular dynamics simulations, essential for understanding binding processes, are resource-intensive and may not capture all dynamic aspects due to computational demands [34].

Quantum computing offers potential solutions, enabling calculations beyond classical supercomputers' capabilities [77]. However, limitations like the restricted number of qubits and reliance on active space approximations can affect quantum machine learning systems' accuracy and practicality in real-world applications. Nevertheless, advancements in quantum algorithms have reduced computational time and enhanced accuracy, particularly in protein similarity assessments [76].

Innovative approaches like the Adaptive Sampling Optimization (ASO) method mitigate computational costs by reducing evaluation numbers, enhancing efficiency by focusing on promising search areas [66]. The MTL-DS method leverages existing data to lower virtual screening computational costs, addressing high resource demands [85].

7.5 Clinical Translation and Validation

Clinical translation and validation of AI-driven discoveries in drug discovery, especially for pancreatic ductal adenocarcinoma (PDAC), present challenges requiring strategic approaches to ensure efficacy and safety. Clinical trials are essential for validating identified compounds' efficacy and safety in real-world applications [2]. These trials bridge laboratory findings with clinical practice, ensuring AI-driven solutions are effective and safe for patients.

Addressing drug resistance is crucial in clinical translation, often undermining kinase inhibitors' effectiveness. Developing selective inhibitors that circumvent resistance mechanisms is necessary for sustained therapeutic benefits [30]. Integrating mathematical models with clinical data enhances predictive accuracy and clinical relevance [20].

Expanding datasets and exploring feature fusion methods that combine traditional radiomic features with transfer learning features are critical for improving prognostic accuracy and facilitating clinical translation [5]. This approach enhances AI model robustness, making them more applicable to diverse clinical scenarios.

Optimizing load balancing algorithms for distributed computing environments is vital for managing computational demands associated with clinical AI model implementation [60]. Addressing these computational challenges will enable efficient clinical data processing and AI-driven solution adoption.

Current studies face limitations related to data availability, robust AI model validation, and regulatory acceptance challenges [51]. Expanding training datasets to include diverse patient data and improving automatic anatomical segmentations' accuracy are crucial steps toward enhancing clinical applicability [94].

7.6 Future Advancements and Interdisciplinary Collaboration

The future of kinase-targeted drug discovery for pancreatic ductal adenocarcinoma (PDAC) will be transformed by computational technique advancements and interdisciplinary collaboration. AI and

computational biology integration is expected to drive significant improvements in designing novel receptor tyrosine kinase inhibitors (RTKIs), enhancing therapeutic efficacy while minimizing toxicity [26]. Future research should focus on validating computational predictions through experimental testing, refining models with new data, and applying methodologies to drug targets beyond tyrosine kinases [10]. This approach ensures computational models' accuracy and applicability across broader drug targets.

Optimizing computational pipelines, such as integrating machine learning with experimental validation, is crucial for advancing drug discovery. Future research should explore additional algorithms and improve method integration to enhance predictive accuracy and applicability [11]. Expanding quantum machine learning frameworks' scalability, like the Quantum Support Vector Classifier (QSVC), to accommodate larger datasets and exploring complex molecular descriptors will further realize quantum advantage potential in drug discovery [13].

Interdisciplinary collaboration is vital for fostering innovation in this field. Integrating high-quality data sources and advancing sophisticated data integration techniques can significantly enhance drug discovery algorithms' predictive capabilities. Systems like AGATHA, leveraging deep learning to generate data-driven hypotheses and prioritize research directions, exemplify this approach, enhancing efficiency and accuracy in identifying potential drug targets while addressing traditional methods' challenges by utilizing publicly available scientific information [60, 18]. Incorporating demographic information and non-radiological factors into predictive models can further enhance classification accuracy and applicability, particularly for specific cyst types.

Emerging trends highlight the importance of refining segmentation processes and exploring fusion strategies to improve detection accuracy and applicability in PDAC diagnosis [4]. Advancing conformal prediction (CP) models through alternative algorithms and integration methods can enhance their efficiency and applicability in drug discovery [67]. Additionally, applying physical formula-enhanced multitask learning (PEMAL) to other scenarios and refining physical constraints offer opportunities for model enhancement [68].

8 Conclusion

The survey underscores the pivotal role of artificial intelligence (AI) and computational biology in advancing kinase-targeted drug discovery for pancreatic ductal adenocarcinoma (PDAC). By leveraging AI-driven methodologies, particularly in enhancing docking score predictions through multi-task learning, the precision and efficiency of drug discovery processes are significantly improved. Computational logic provides a comprehensive framework for modeling and verifying biological systems, which is critical for the development of innovative therapeutic strategies. The ongoing pursuit of novel kinase inhibitors aims to overcome drug resistance and enhance treatment efficacy, highlighting AI's transformative potential in revolutionizing drug discovery and personalizing medicine. Additionally, the exploration of natural compounds offers a promising alternative to synthetic drugs, potentially reducing side effects and improving patient outcomes. Addressing data-centric challenges is crucial for the practical application of AI models, where incorporating causal reasoning can refine the identification of therapeutic targets. Despite the abundance of publicly available data for constructing drug discovery knowledge graphs, challenges remain in effectively integrating this data and ensuring the reliability of the resultant graphs. Advancements in PDAC detection, such as using multi-phase CT scans, further emphasize the potential of AI and computational biology to enhance drug discovery. The integration of federated learning with cryptographic techniques presents a promising avenue for ensuring privacy in biomedical applications, thereby fostering interdisciplinary collaboration and data sharing.

References

[1] Atefeh Akbarnia Dafrazi, Tahmineh Mehrabi, and Fatemeh Malekinejad. A bioinformatics study for recognition of hub genes and pathways in pancreatic ductal adenocarcinoma, 2023.

[2] Marsha Mariya Kappan and Joby George. In silico pharmacokinetic and molecular docking studies of natural plants against essential protein kras for treatment of pancreatic cancer, 2024.

[3] Yucheng Zhang, Edrise M. Lobo-Mueller, Paul Karanicolas, Steven Gallinger, Masoom A. Haider, and Farzad Khalvati. Improving prognostic performance in resectable pancreatic ductal adenocarcinoma using radiomics and deep learning features fusion in ct images, 2019.

[4] Fengze Liu, Yuyin Zhou, Elliot Fishman, and Alan Yuille. Fusionnet: Incorporating shape and texture for abnormality detection in 3d abdominal ct scans, 2019.

[5] Yucheng Zhang, Edrise M. Lobo-Mueller, Paul Karanicolas, Steven Gallinger, Masoom A. Haider, and Farzad Khalvati. Prognostic value of transfer learning based features in resectable pancreatic ductal adenocarcinoma, 2019.

[6] Catrin Hasselgren and Tudor I. Oprea. Artificial intelligence for drug discovery: Are we there yet?, 2023.

[7] Bharath Ramsundar, Steven Kearnes, Patrick Riley, Dale Webster, David Konerding, and Vijay Pande. Massively multitask networks for drug discovery, 2015.

[8] G Jignesh Chowdary, Suganya G, Premalatha M, Asnath Victy Phamila Y, and Karunamurthy K. Machine learning and deep learning methods for building intelligent systems in medicine and drug discovery: A comprehensive survey, 2021.

[9] Yuhan Li, Hongtao Zhang, Keaven Anderson, Songzi Li, and Ruoqing Zhu. Ai in pharma for personalized sequential decision-making: Methods, applications and opportunities, 2023.

[10] Ricardo Romero. Integration of genetic algorithms and deep learning for the generation and bioactivity prediction of novel tyrosine kinase inhibitors, 2024.

[11] Aymen Al Saadi, Dario Alfe, Yadu Babuji, Agastya Bhati, Ben Blaiszik, Thomas Brettin, Kyle Chard, Ryan Chard, Peter Coveney, Anda Trifan, Alex Brace, Austin Clyde, Ian Foster, Tom Gibbs, Shantenu Jha, Kristopher Keipert, Thorsten Kurth, Dieter Kranzlmüller, Hyungro Lee, Zhuozhao Li, Heng Ma, Andre Merzky, Gerald Mathias, Alexander Partin, Junqi Yin, Arvind Ramanathan, Ashka Shah, Abraham Stern, Rick Stevens, Li Tan, Mikhail Titov, Aristeidis Tsaris, Matteo Turilli, Huub Van Dam, Shunzhou Wan, and David Wifling. Impeccable: Integrated modeling pipeline for covid cure by assessing better leads, 2020.

[12] Thomas Gaudelet, Ben Day, Arian R. Jamasb, Jyothish Soman, Cristian Regep, Gertrude Liu, Jeremy B. R. Hayter, Richard Vickers, Charles Roberts, Jian Tang, David Roblin, Tom L. Blundell, Michael M. Bronstein, and Jake P. Taylor-King. Utilising graph machine learning within drug discovery and development, 2021.

[13] Stefano Mensa, Emre Sahin, Francesco Tacchino, Panagiotis Kl. Barkoutsos, and Ivano Tavernelli. Quantum machine learning framework for virtual screening in drug discovery: a prospective quantum advantage, 2022.

[14] Alvaro Prat, Hisham Abdel Aty, Gintautas Kamuntavičius, Tanya Paquet, Povilas Norvaišas, Piero Gasparotto, and Roy Tal. Hydrascreen: A generalizable structure-based deep learning approach to drug discovery, 2023.

[15] Oleksandr Bilokon, Nataliya Bilokon, and Paul Bilokon. Ai-driven alternative medicine: A novel approach to drug discovery and repurposing, 2024.

[16] Adam Zielinski. Ai and the future of pharmaceutical research, 2021.

[17] Ashwin Mukund, Muhammad Ali Afridi, Aleksandra Karolak, Margaret A Park, Jennifer B Permuth, and Ghulam Rasool. Pancreatic ductal adenocarcinoma (pdac): a review of recent advancements enabled by artificial intelligence. *Cancers*, 16(12):2240, 2024.

[18] Alexandre Blanco-Gonzalez, Alfonso Cabezon, Alejandro Seco-Gonzalez, Daniel Conde-Torres, Paula Antelo-Riveiro, Angel Pineiro, and Rebeca Garcia-Fandino. The role of ai in drug discovery: Challenges, opportunities, and strategies, 2022.

[19] Yingda Xia, Qihang Yu, Wei Shen, Yuyin Zhou, Elliot K. Fishman, and Alan L. Yuille. Detecting pancreatic ductal adenocarcinoma in multi-phase ct scans via alignment ensemble, 2020.

[20] Sophie Chauvet, Florence Hubert, Fanny Mann, and Mathieu Mezache. Tumorigenesis and axons regulation for the pancreatic cancer: a mathematical approach, 2022.

[21] Jiawen Yao, Yu Shi, Le Lu, Jing Xiao, and Ling Zhang. Deepprognosis: Preoperative prediction of pancreatic cancer survival and surgical margin via contrast-enhanced ct imaging, 2020.

[22] Yu Shi, Hannah Tang, Michael Baine, Michael A. Hollingsworth, Huijing Du, Dandan Zheng, Chi Zhang, and Hongfeng Yu. 3dgaunet: 3d generative adversarial networks with a 3d u-net based generator to achieve the accurate and effective synthesis of clinical tumor image data for pancreatic cancer, 2023.

[23] Nicholas Ekow Thomford, Dimakatso Alice Senthebane, Arielle Rowe, Daniella Munro, Palesa Seele, Alfred Maroyi, and Kevin Dzobo. Natural products for drug discovery in the 21st century: innovations for novel drug discovery. *International journal of molecular sciences*, 19(6):1578, 2018.

[24] Hongwei Li, Kanru Lin, Maximilian Reichert, Lina Xu, Rickmer Braren, Deliang Fu, Roland Schmid, Ji Li, Bjoern Menze, and Kuangyu Shi. Differential diagnosis for pancreatic cysts in ct scans using densely-connected convolutional networks, 2018.

[25] Yan Wang, Peng Tang, Yuyin Zhou, Wei Shen, Elliot K. Fishman, and Alan L. Yuille. Learning inductive attention guidance for partially supervised pancreatic ductal adenocarcinoma prediction, 2021.

[26] 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.

[27] Fan Hu, Dongqi Wang, Huazhen Huang, Yishen Hu, and Peng Yin. Bridging the gap between target-based and cell-based drug discovery with a graph generative multi-task model, 2022.

[28] Shanzhuo Zhang, Zhiyuan Yan, Yueyang Huang, Lihang Liu, Donglong He, Wei Wang, Xiaomin Fang, Xiaonan Zhang, Fan Wang, Hua Wu, and Haifeng Wang. Helixadmet: a robust and endpoint extensible admet system incorporating self-supervised knowledge transfer, 2022.

[29] Misha P. T. Kaandorp, Sebastiano Barbieri, Remy Klaassen, Hanneke W. M. van Laarhoven, Hans Crezee, Peter T. While, Aart J. Nederveen, and Oliver J. Gurney-Champion. Improved unsupervised physics-informed deep learning for intravoxel incoherent motion modeling and evaluation in pancreatic cancer patients, 2021.

[30] Khushwant S Bhullar, Naiara Orrego Lagarón, Eileen M McGowan, Indu Parmar, Amitabh Jha, Basil P Hubbard, and HP Vasantha Rupasinghe. Kinase-targeted cancer therapies: progress, challenges and future directions. *Molecular cancer*, 17:1–20, 2018.

[31] Radhamani Kannaiyan and Daruka Mahadevan. A comprehensive review of protein kinase inhibitors for cancer therapy. *Expert review of anticancer therapy*, 18(12):1249–1270, 2018.

[32] Carrow I Wells, Hassan Al-Ali, David M Andrews, Christopher RM Asquith, Alison D Axtman, Ivan Dikic, Daniel Ebner, Peter Ettmayer, Christian Fischer, Mathias Frederiksen, et al. The kinase chemogenomic set (kcgs): an open science resource for kinase vulnerability identification. *International journal of molecular sciences*, 22(2):566, 2021.

[33] Zheng Zhao and Philip E. Bourne. Using the structural kinome to systematize kinase drug discovery, 2021.

[34] Marawan Ahmed, Maiada M. Sadek, Khaled A. Abouzid, and Feng Wang. In silico design, extended molecular dynamic simulations and binding energy calculations for a new series of dually acting inhibitors against egfr and her2, 2013.

[35] George A. Elder and Conrad Bessant. Inferring compensatory kinase networks in yeast using prolog, 2023.

[36] Liling Huang, Shiyu Jiang, and Yuankai Shi. Tyrosine kinase inhibitors for solid tumors in the past 20 years (2001–2020). *Journal of hematology & oncology*, 13:1–23, 2020.

[37] Farzin Sohraby and Ariane Nunes-Alves. Recent advances in computational methods for studying ligand binding kinetics, 2022.

[38] Shahar Harel and Kira Radinsky. Accelerating prototype-based drug discovery using conditional diversity networks, 2018.

[39] Yoshitaka Inoue, Tianci Song, and Tianfan Fu. Drugagent: Explainable drug repurposing agent with large language model-based reasoning, 2024.

[40] Yizhen Luo, Xing Yi Liu, Kai Yang, Kui Huang, Massimo Hong, Jiahuan Zhang, Yushuai Wu, and Zaiqing Nie. Towards unified ai drug discovery with multiple knowledge modalities, 2023.

[41] Po-Yu Liang and Jun Bai. E(3)-invaraint diffusion model for pocket-aware peptide generation, 2024.

[42] Bin Liu, Konstantinos Pliakos, Celine Vens, and Grigorios Tsoumakas. Drug-target interaction prediction via an ensemble of weighted nearest neighbors with interaction recovery, 2021.

[43] Junhuan Yang, Yi Sheng, Sizhe Zhang, Ruixuan Wang, Kenneth Foreman, Mikell Paige, Xun Jiao, Weiwen Jiang, and Lei Yang. Automated architecture search for brain-inspired hyperdimensional computing, 2022.

[44] Jianyuan Deng, Zhibo Yang, Iwao Ojima, Dimitris Samaras, and Fusheng Wang. Artificial intelligence in drug discovery: Applications and techniques, 2021.

[45] José Jiménez-Luna, Francesca Grisoni, and Gisbert Schneider. Drug discovery with explainable artificial intelligence. *Nature Machine Intelligence*, 2(10):573–584, 2020.

[46] Yankang Jing, Yuemin Bian, Ziheng Hu, Lirong Wang, and Xiang-Qun Sean Xie. Deep learning for drug design: an artificial intelligence paradigm for drug discovery in the big data era. *The AAPS journal*, 20:1–10, 2018.

[47] Isaac Filella-Merce, Alexis Molina, Marek Orzechowski, Lucía Díaz, Yang Ming Zhu, Julia Vilalta Mor, Laura Malo, Ajay S Yekkirala, Soumya Ray, and Victor Guallar. Optimizing drug design by merging generative ai with active learning frameworks, 2023.

[48] Tri Minh Nguyen, Thin Nguyen, and Truyen Tran. Learning to discover medicines, 2022.

[49] Alex Golts, Vadim Ratner, Yoel Shoshan, Moshe Raboh, Sagi Polaczek, Michal Ozery-Flato, Daniel Shats, Liam Hazan, Sivan Ravid, and Efrat Hexter. A large dataset curation and benchmark for drug target interaction, 2024.

[50] Zhangli Lu, Chuqi Lei, Kaili Wang, Libo Qin, Jing Tang, and Min Li. Dtiam: A unified framework for predicting drug-target interactions, binding affinities and activation/inhibition mechanisms, 2023.

[51] Yu Han and Jingwen Tao. Revolutionizing pharma: Unveiling the ai and llm trends in the pharmaceutical industry, 2024.

[52] Tristan Aumentado-Armstrong. Latent molecular optimization for targeted therapeutic design, 2018.

[53] Stephen Bonner, Ian P Barrett, Cheng Ye, Rowan Swiers, Ola Engkvist, Charles Tapley Hoyt, and William L Hamilton. Understanding the performance of knowledge graph embeddings in drug discovery, 2022.

[54] Xuan Lin, Lichang Dai, Yafang Zhou, Zu-Guo Yu, Wen Zhang, Jian-Yu Shi, Dong-Sheng Cao, Li Zeng, Haowen Chen, Bosheng Song, Philip S. Yu, and Xiangxiang Zeng. Comprehensive evaluation of deep and graph learning on drug-drug interactions prediction, 2023.

[55] A knowledge graph representation learning approach to predict novel kinase-substrate interactions.

[56] Peter Csermely, Tamas Korcsmaros, Huba J. M. Kiss, Gabor London, and Ruth Nussinov. Structure and dynamics of molecular networks: A novel paradigm of drug discovery. a comprehensive review, 2013.

[57] Zhiqiang Zhong, Anastasia Barkova, and Davide Mottin. Knowledge-augmented graph machine learning for drug discovery: A survey, 2024.

[58] F. Pellicani, D. Dal Ben, A. Perali, and S. Pilati. Machine learning scoring functions for drug discoveries from experimental and computer-generated protein-ligand structures: Towards per-target scoring functions, 2023.

[59] Jiawei Wu, Jun Wen, Mingyuan Yan, Anqi Dong, Shuai Gao, Ren Wang, and Can Chen. Heterogeneous entity representation for medicinal synergy prediction, 2024.

[60] Justin Sybrandt, Ilya Tyagin, Michael Shtutman, and Ilya Safro. Agatha: Automatic graph-mining and transformer based hypothesis generation approach, 2020.

[61] Lanqing Li, Liang Zeng, Ziqi Gao, Shen Yuan, Yatao Bian, Bingzhe Wu, Hengtong Zhang, Yang Yu, Chan Lu, Zhipeng Zhou, Hongteng Xu, Jia Li, Peilin Zhao, and Pheng-Ann Heng. Imdrug: A benchmark for deep imbalanced learning in ai-aided drug discovery, 2022.

[62] Atsuko Takagi, Mayumi Kamada, Eri Hamatani, Ryosuke Kojima, and Yasushi Okuno. Graphix: Graph-based in silico xai(explainable artificial intelligence) for drug repositioning from biopharmaceutical network, 2022.

[63] James Michels, Ramya Bandarupalli, Amin Ahangar Akbari, Thai Le, Hong Xiao, Jing Li, and Erik F. Y. Hom. Natural language processing methods for the study of protein-ligand interactions, 2024.

[64] Bing Hu, Anita Layton, and Helen Chen. Drug discovery smiles-to-pharmacokinetics diffusion models with deep molecular understanding, 2024.

[65] Kieran Didi and Matej Zečević. On how ai needs to change to advance the science of drug discovery, 2022.

[66] Marylens Hernandez, Alexander Lachmann, Shan Zhao, Kunhong Xiao, and Avi Ma'ayan. Inferring the sign of kinase-substrate interactions by combining quantitative phosphoproteomics with a literature-based mammalian kinome network, 2010.

[67] Isidro Cortés-Ciriano and Andreas Bender. Concepts and applications of conformal prediction in computational drug discovery, 2019.

[68] Ruifeng Li, Dongzhan Zhou, Ancheng Shen, Ao Zhang, Mao Su, Mingqian Li, Hongyang Chen, Gang Chen, Yin Zhang, Shufei Zhang, Yuqiang Li, and Wanli Ouyang. Physical formula enhanced multi-task learning for pharmacokinetics prediction, 2024.

[69] Jessica Vamathevan, Dominic Clark, Paul Czodrowski, Ian Dunham, Edgardo Ferran, George Lee, Bin Li, Anant Madabhushi, Parantu Shah, Michaela Spitzer, et al. Applications of machine learning in drug discovery and development. *Nature reviews Drug discovery*, 18(6):463–477, 2019.

[70] Khartik Uppalapati, Eeshan Dandamudi, S. Nick Ice, Gaurav Chandra, Kirsten Bischof, Christian L. Lorson, and Kamal Singh. A comprehensive guide to enhancing antibiotic discovery using machine learning derived bio-computation, 2024.

[71] Jian-Ping Mei, Chee-Keong Kwoh, Peng Yang, and Xiao-Li Li. Classification and its applications for drug-target interaction identification, 2015.

[72] Stephen Bonner, Ian P Barrett, Cheng Ye, Rowan Swiers, Ola Engkvist, Andreas Bender, Charles Tapley Hoyt, and William L Hamilton. A review of biomedical datasets relating to drug discovery: A knowledge graph perspective, 2021.

[73] Narbe Mardirossian, Yuhang Wang, David A. Pearlman, Garnet Kin-Lic Chan, and Toru Shiozaki. Novel algorithms and high-performance cloud computing enable efficient fully quantum mechanical protein-ligand scoring, 2020.

[74] Yifan Zhou, Yan Shing Liang, Yew Kee Wong, Haichuan Qiu, Yu Xi Wu, and Bin He. Implementation of the future of drug discovery: Quantumbased machine learning simulation (qmls), 2024.

[75] Mohammad Ghazi Vakili, Christoph Gorgulla, AkshatKumar Nigam, Dmitry Bezrukov, Daniel Varoli, Alex Aliper, Daniil Polykovsky, Krishna M. Padmanabha Das, Jamie Snider, Anna Lyakisheva, Ardalan Hosseini Mansob, Zhong Yao, Lela Bitar, Eugene Radchenko, Xiao Ding, Jinxin Liu, Fanye Meng, Feng Ren, Yudong Cao, Igor Stagljar, Alán Aspuru-Guzik, and Alex Zhavoronkov. Quantum computing-enhanced algorithm unveils novel inhibitors for kras, 2024.

[76] Anthony Chagneau, Yousra Massaoudi, Imene Derbali, and Linda Yahiaoui. Quantum algorithm for bioinformatics to compute the similarity between proteins, 2024.

[77] Carlos Outeiral, Martin Strahm, Jiye Shi, Garrett M. Morris, Simon C. Benjamin, and Charlotte M. Deane. The prospects of quantum computing in computational molecular biology, 2020.

[78] Weitang Li, Zhi Yin, Xiaoran Li, Dongqiang Ma, Shuang Yi, Zhenxing Zhang, Chenji Zou, Kunliang Bu, Maochun Dai, Jie Yue, Yuzong Chen, Xiaojin Zhang, and Shengyu Zhang. A hybrid quantum computing pipeline for real world drug discovery, 2024.

[79] Anthony M. Smaldone, Yu Shee, Gregory W. Kyro, Chuzhi Xu, Nam P. Vu, Rishab Dutta, Marwa H. Farag, Alexey Galda, Sandeep Kumar, Elica Kyoseva, and Victor S. Batista. Quantum machine learning in drug discovery: Applications in academia and pharmaceutical industries, 2024.

[80] Zheng Zhao and Philip E. Bourne. Overview of current type i/ii kinase inhibitors, 2018.

[81] Yunchao, Liu, Ha Dong, Xin Wang, Rocco Moretti, Yu Wang, Zhaoqian Su, Jiawei Gu, Bobby Bodenheimer, Charles David Weaver, Jens Meiler, and Tyler Derr. Welqrate: Defining the gold standard in small molecule drug discovery benchmarking, 2024.

[82] Juan Manuel Zambrano Chaves, Eric Wang, Tao Tu, Eeshit Dhaval Vaishnav, Byron Lee, S. Sara Mahdavi, Christopher Semturs, David Fleet, Vivek Natarajan, and Shekoofeh Azizi. Tx-llm: A large language model for therapeutics, 2024.

[83] Yossra Gharbi and Rocío Mercado. A comprehensive review of emerging approaches in machine learning for de novo protac design, 2024.

[84] Balazs Pejo, Mina Remeli, Adam Arany, Mathieu Galtier, and Gergely Acs. Collaborative drug discovery: Inference-level data protection perspective, 2022.

[85] Zijing Liu, Xianbin Ye, Xiaomin Fang, Fan Wang, Hua Wu, and Haifeng Wang. Docking-based virtual screening with multi-task learning, 2021.

[86] Fujian Jia, Xin Liu, Lixi Deng, Jiwen Gu, Chunchao Pu, Tunan Bai, Mengjiang Huang, Yuanzhi Lu, and Kang Liu. Oncogpt: A medical conversational model tailored with oncology domain expertise on a large language model meta-ai (llama), 2024.

[87] Ling Zhang, Yu Shi, Jiawen Yao, Yun Bian, Kai Cao, Dakai Jin, Jing Xiao, and Le Lu. Robust pancreatic ductal adenocarcinoma segmentation with multi-institutional multi-phase partially-annotated ct scans, 2020.

[88] Alexander Ziller, Ayhan Can Erdur, Friederike Jungmann, Daniel Rueckert, Rickmer Braren, and Georgios Kaissis. Exploiting segmentation labels and representation learning to forecast therapy response of pdac patients, 2023.

[89] Suresh Dara, Swetha Dhamercherla, Surender Singh Jadav, CH Madhu Babu, and Mohamed Jawed Ahsan. Machine learning in drug discovery: a review. *Artificial intelligence review*, 55(3):1947–1999, 2022.

[90] Matthew C. Robinson, Robert C. Glen, and Alpha A. Lee. Validating the validation: Reanalyzing a large-scale comparison of deep learning and machine learning models for bioactivity prediction, 2019.

[91] Stanley E. Lazic and Dominic P. Williams. Quantifying sources of uncertainty in drug discovery predictions with probabilistic models, 2021.

[92] Isidro Cortes Ciriano and Andreas Bender. Kekulescope: prediction of cancer cell line sensitivity and compound potency using convolutional neural networks trained on compound images, 2019.

[93] Jimin Tan, Jianan Yang, Sai Wu, Gang Chen, and Jake Zhao. A critical look at the current train/test split in machine learning, 2021.

[94] Natália Alves, Megan Schuurmans, Geke Litjens, Joeran S. Bosma, John Hermans, and Henkjan Huisman. Fully automatic deep learning framework for pancreatic ductal adenocarcinoma detection on computed tomography, 2021.

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.