OPEN ACCESS

Saudi Journal of Medical and Pharmaceutical Sciences

Abbreviated Key Title: Saudi J Med Pharm Sci
ISSN 2413-4929 (Print) | ISSN 2413-4910 (Online)
Scholars Middle East Publishers, Dubai, United Arab Emirates
Journal homepage: https://saudijournals.com

Review Article **Pharmacology**

The Role of Artificial Intelligence and Machine Learning in Revolutionizing Drug Discovery and Pharmacological Research: A Systematic Review

Zainab G. Aljassim[1]*, Hiba Ghassan Rajab[1], Huda I. Al-Qadhi[1]

[1]Pharmacology Department, College of Medicine, University of Baghdad, Iraq

DOI: https://doi.org/10.36348/sjmps.2026.v12i05.001 | **Received:** 01.03.2026 | **Accepted:** 28.04.2026 | **Published:** 01.05.2026

***Corresponding author:** Zainab G. Aljassim
Pharmacology Department, College of Medicine, University of Baghdad, Iraq

Abstract

Artificial intelligence (AI), coupled with machine learning (ML) has been rapidly incorporated into pharmaceutical discovery and development. We reviewed 53 publications from 2018-2026 to summarize current applications of AI/ML in drug discovery. AI and ML have potential to impact every step of the drug development pipeline and have already shown to drastically reduce time frames for developing therapeutics. Specific deep learning models such as graph neural networks and transformers have shown promise in de novo molecular generation, molecular property prediction, and target recognition. Accurate protein structure prediction using AlphaFold allows for exploration of drug-target binding. De novo drug design with reinforcement learning allows for targeted design of molecules with desired properties. Machine learning models for QSAR provide more accurate toxicity predictions and ADMET profiling to avoid potential failures during drug development. However, current limitations include lack of interpretability, data limitations, and lack of regulatory approval. According to a review of recent literature, AI has the potential to decrease the time required for drug discovery from years to months and lower the cost of drug development. This review discusses recent advances, successful clinical examples, and opportunities for artificial intelligence/machine learning in drug discovery.

Keywords: AI, ML, Drug Discovery, Deep Learning, Protein Structure Prediction, Reinforcement Learning, De Novo Drug Design.

1. INTRODUCTION

Drug discovery represents one of the most daunting tasks undertaken by the pharmaceutical industry. R&D costs have soared to an average of $2.6 billion per drug over 10–15-year periods [1]. For decades pharmaceutical companies have taken a trial-and-error approach to the discovery process, only to find that ~90% of drug candidates fail in clinical trials [2, 3]. Recent advancements in artificial intelligence and machine learning have created unique opportunities to overcome some of these challenges [4, 5]. AI-based methods utilizing large datasets, algorithms, and computation have the potential to speed up target identification, improve molecular design, predict toxicity, and increase clinical success [6, 10]. This review looks at 53 articles from 2018-2026 from peer-reviewed journals and discusses the implementation, applications, and effect of AI/ML throughout drug discovery [11].

1.1 The Evolution of Drug Discovery: From Traditional to AI-Driven Approaches

Drug discovery has undergone remarkable transformation over the past century, evolving from serendipitous observations to rational design strategies and now to AI-driven computational approaches [1, 10]. Traditional drug discovery began with natural product screening and phenotypic observations, exemplified by Alexander Fleming's accidental discovery of penicillin. The mid-20th century witnessed the emergence of structure-activity relationship studies and rational drug design, where medicinal chemists systematically modified molecular structures to optimize biological activity. High-throughput screening technologies in the 1990s enabled rapid testing of thousands of compounds, yet success rates remained disappointingly low due to the vast complexity of chemical space, estimated to contain 10^{60} possible drug-like molecules [11].

Computational techniques including molecular modeling, docking simulations and quantitative

structure–activity relationship models ("classic" computational drug discovery techniques) have existed since computational chemistry became feasible. These methods represented the first generation of in silico drug discovery [12]. These classic computational techniques, however, were unable to adequately model systems where the relationship between molecular structure and biological activity is highly non-linear. Newer techniques which rely on artificial intelligence to apply machine learning methods to drug discovery have begun to overcome these challenges by capitalizing on machine learning techniques' ability to detect patterns in high-dimensional data [2, 3]. Deep learning-based methods, especially neural networks, have begun showing exceptional promise at predicting molecular properties, generating novel molecules, and performing multi-objective optimization of molecules [7, 8]. Insilico Medicine, Exscientia, and BenevolentAI are among the first companies to exclusively use an AI-first strategy for drug discovery and have already shown success advancing molecules discovered via AI into clinical trials [1, 6].

1.2 The Pharmaceutical Industry's Challenge: Time, Cost, and Failure Rates

The pharmaceutical industry faces an age where scientific and technological advancement has never been higher but productivity in drug discovery continues to fall [2, 5]. Estimated Drug development costs have skyrocketed to $2.6-2.8 billion and take 10-15 years from discovery to approval. This trend of diminishing research and development output, coined " Eroom's Law" (moore's law spelled backwards), despite massive increases in resources and technological improvements has been exponential [3]. Most of these inefficiencies can be attributed to failure; specifically, failure across all stages of development is extremely high. An estimated 90% of potential drug compounds do not make it to approval after entering clinical trials with most eliminated at Phase II and III trials due to toxicity or lack of efficacy [4, 5].

Failures at these stages of drug development translate into sunk costs for drug developers after years of research and investment. In addition to late-stage attrition, many traditional methods face hurdles like: (1) the vastness of chemical space makes it impossible to practically screen all molecules; (2) challenging targets with complicated PK/Tox cannot be easily solved with traditional methods; (3) polypharmacology and off-target effects are difficult to account for; and (4) poor translation from bench to bedside [6]. The financial burden of these issues has pushed pharmaceutical companies to pursue innovative solutions. Artificial intelligence can address many of these shortcomings by: (1) predicting properties of molecules earlier in the drug discovery pipeline; (2) more accurately identifying drugs that will have higher chances of clinical success; and (3) decreasing the time and cost of each stage of drug development [1, 10, 11].

1.3 Research Questions

What are the core artificial intelligence concepts and machine learning techniques currently implemented throughout various aspects of the drug discovery pipeline? How do these algorithmic approaches differ in terms of precision, scalability, and real-world application? [7, 8, 12] This includes comparisons between classical machine learning models, neural network architectures, graph convolutional networks, and reinforcement learning implemented for target identification, hit identification, lead optimization, etc. Second, AlphaFold [25] ushered in a new era for protein structure prediction. How does AlphaFold's structure prediction revolution affect structure-based drug design? And what are the limitations of using predicted protein structures for virtual screening and molecular docking? [22-24] Third, deep learning-based generative models have shown promise for de novo molecular design, but how accurate and reliable are they? [41-43] For example, how many AI-designed molecules have entered clinical development? Fourth, we are already beginning to see that machine learning can predict toxicity/ADMET properties/outcomes.

How well can we predict toxicity/ADMET properties/outcomes compared to traditional methods, and will they be more cost effective? [32-34] How can we overcome these barriers? Fifth, what hurdles prevent AI from being more broadly used currently? Examples of potential topics could be issues with data quality [2], model interpretability [3], regulatory acceptance, lack of integration into existing workflows etc. Sixth, what solutions show the most promise moving forward? For example, new technologies or methodologies that have the potential to enable AI for drug discovery over the next decade [1]. With these questions in mind, how can industry, academia, and government work together to create standard frameworks, validation practices, and best practices to support the responsible and effective use of AI in drug discovery?

1.4 Research Methodology

The systematic review was prepared in accordance with Preferred Reporting Items for Systematic Reviews and Meta-Analyses guidelines for Literature Search in Biomedical Studies (Figure 1). We searched PubMed, Web of Science, IEEE Xplore, Scopus, and Google Scholar databases for literature published between January 2018 to February 2026. Search terms included keywords and MeSH terms such as "artificial intelligence", "machine learning", "deep learning", "drug discovery", "drug design", "AI in pharmaceutical research", "molecular design", "protein structure prediction", "AlphaFold", "drug toxicity prediction", "toxicity prediction", "QSAR", "virtual screening", et al. Search strings were combined using Boolean operators. The inclusion criteria were peer-reviewed journal articles, proceedings from well-known conferences and preprints hosted on well-known repositories that presented novel research works,

approaches, or reviewed existing literature on AI/ML algorithms used for drug discovery. Papers were considered if they described enough detail about the methods used, validation results, or comparisons to other studies.

Opinion pieces lacking technical detail, papers from predatory journals, and non-English papers were excluded. Our search returned 847 articles, which were title and abstract screened for relevance. After this stage, we obtained 312 papers which we full-text screened. In this stage we examined each paper for quality, relevance, and technical detail. 53 papers were included in this review. Information extracted included AI/ML approaches used, application in drug discovery, dataset, performance metrics, validation, limitations and clinical/practical impact reached [4-6]. Papers were grouped based on primary application allowing standardized review across 8 different applications of AI to drug discovery. Quality of studies was scored based on criteria such as study design, validation, reproducibility, and impact.

2. FUNDAMENTALS OF ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING IN DRUG DISCOVERY

AI and ML include many different techniques and algorithms that allow computers to "learn" from data and make predictions or decisions without being explicitly programmed [7, 10]. Several applications have been developed in drug discovery settings ranging from target and hit identification to lead optimization and forecasting clinical results [11, 12]. Approaches span from traditional machine learning techniques such as random forests and support vector machines to deep learning methods such as convolutional neural networks (CNNs), recurrent neural networks, and transformers [8, 9]. Basic understanding of these techniques as well as molecular representations and datasets will allow us to better understand how AI can be leveraged for drug discovery [50, 51]. This chapter will cover these fundamentals of AI-based drug discovery applications.

2.1 Core Machine Learning Algorithms and Deep Learning Architectures

Machine learning algorithms are the core computational approaches that power AI drug discovery models. These models can be classified into three main categories: supervised learning, unsupervised learning, and reinforcement learning [10, 11]. Among supervised learning models, algorithms like random forests, support vector machines, and gradient boosting have been employed for classification and regression tasks like bioactivity prediction, toxicity prediction, and ADME property prediction [12, 50]. While classical machine learning models like these have shown success when applied to structured datasets with well-understood features, they also benefit from being highly interpretable and relatively low in computational cost. Unsupervised learning models are also used in cheminformatics, particularly clustering and dimensionality reduction methods. Deep learning methods train artificial neural networks with multiple layers of nonlinearities that discover representations automatically from data [7, 8].

Convolutional neural networks can be leveraged on grid-like representations of molecules, as well as images produced by high-content screening assays [9]. Recurrent neural networks and long short-term memory networks are used for sequential data such as SMILES and for property prediction from sequential molecular representations [51]. Transformer architectures used for natural language processing have recently been used for molecular generation and prediction of molecular properties. The choice of which algorithms to use depends on the task at hand, data available, and resources. Techniques that ensemble several models typically perform best [12]. Also transfer learning methods that pre-train a model on a large corpus of unlabeled molecular data and fine-tune to perform some desired downstream tasks show promising improvements especially when labeled data is scarce [50, 51].

2.2 Molecular Representations and Chemical Descriptors for AI Applications

Machine learning methods require numerical representations of chemical structures as input to allow the algorithms to learn and understand structure-property or structure-activity relationships [11, 12]. Historically this has been done with hand-designed molecular descriptors which encode some aspect of the molecules' physicochemical properties, topology, or structure. Such descriptors include constitutional descriptors (total molecular weight, number of certain atoms), topological descriptors (Wiener index, Zagreb indices), geometrical descriptors (molecular volume, surface area), and electronic descriptors (atoms partial charges, molecular dipole moment), among many others. Another commonly used representation are extended-connectivity fingerprints, like Morgan or ECFP, where the molecule is encoded as a binary vector indicating which of a predetermined list of substructural features are present in the molecule. SMILES strings represent molecules as line notations that encode atoms, bonds, and stereoisomerism information as strings of printable characters [50, 51].

SMILES notations naturally encode molecules in linear forms that are suitable for sequence-based deep learning models. SMILES are also used commonly as input representation for many existing AI-for-drug discovery platforms. On the downside, SMILES notations suffer from non-uniqueness issues (one molecule can have many valid SMILES representations) and lack of 3D structural representation. Molecular graphs naturally encode atoms as nodes and chemical bonds as edges in the graph. The node and edge features can be used to represent atom properties and bond

features respectively [16, 17]. Graph representation allows us to use GNNs which have been empirically shown to outperform other types of architectures on many tasks. Three-dimensional representations include position and conformation data. These representations are useful in applications like structure-based drug design and predicting binding affinity [22]. Multi-modal learning has emerged where multiple representations are used to allow models to benefit from the information gain of multiple encoding strategies. The impact of molecular representation scheme on predictive performance is profound and will depend on the representation used as well as the learning algorithm and task [7, 8].

2.3 Data Sources and Databases for AI-Driven Drug Discovery

AI applications rely on data, and there are many publicly available and private databases that house training data for drug discovery. ChEMBL is one of the largest publicly available bioactivity databases. Hosted by the European Bioinformatics Institute, ChEMBL is a manually curated database of over 2.4 million unique compounds with >19 million reported activity measurements against biological targets. [6] PubChem is a free chemistry database hosted by the National Center for Biotechnology Information which includes chemical structures, properties, and bioactivity for over 110 million unique compounds. The ZINC database is a free database of commercially available compounds for virtual screening. It contains over 230 million purchasable compounds in ready-to-dock, 3D formats. Information about protein structures can be retrieved from the Protein Data Bank, which contains >200,000 experimentally determined structures [22, 23], as shown in Table 1.

Table 1: Major Databases and Data Resources for AI-Driven Drug Discovery

Database	Type	Size	Data Type	Access	Key Features	Updates	Reference
ChEMBL	Bioactivity	2.4M compounds, 19M activities	Drug-target interactions	Public	Curated bioactivity data across targets	Quarterly	[6, 11]
AlphaFold DB	Structural	200M proteins	Predicted protein structures	Public	Near-experimental accuracy, confidence scores	Continuous	[24, 26, 27]
PubChem	Chemical	110M compounds	Chemical structures, properties	Public	Comprehensive chemical information	Daily	[11, 12]
Tox21	Toxicity	10K compounds, 12 assays	Toxicity endpoints	Public	HTS toxicity data	Static	[32, 34]
DrugBank	Drug	14K drugs	Drug information, targets, interactions	Academic/ Commercial	Comprehensive drug encyclopedia	Annual	[35]

Structural coverage has been significantly expanded through the AlphaFold Protein Structure Database, which provides predicted protein structures for over 200 million proteins covering nearly the entire protein universe [24, 25]. Tox21, ToxCast, and OECD QSAR Toolbox are sources for curated datasets that span different toxicological endpoints for toxicity prediction and safety assessment [32-34]. DrugBank includes detailed drug information, such as mechanisms of action, pharmacology, and drug-drug interactions for approved and experimental drugs. There are also databases specific to certain domains. BindingDB collects protein-ligand binding affinities, STITCH annotates chemical-protein interactions, and SIDER lists drug side effects/ adverse drug reactions [35,36]. The quantity, quality, and diversity of training data impacts model performance and ability to generalize. Data standardization, curation, and integration is an ongoing challenge and efforts are being made to create community wide standards and consolidated data formats [37, 38]. Industry datasets used by pharmaceutical companies tend to be much higher quality due to standardized experimental procedures, however this data is generally inaccessible to academic circles.

2.4 Overview of AI Applications Across the Drug Discovery Pipeline

AI has been leveraged at almost every step of drug discovery and development process pipelines which can be found altered drug discovery workflows incorporating AI [1-3]. For instance, during target identification and validation, AI is utilized to process genomic, proteomic, and phenotypic datasets to discover proteins and pathways associated with diseases of interest. Disease-associated proteins and pathways can

then be prioritized using AI for their druggability and likelihood of therapeutic success [10, 11]. Machine learning techniques can also be used to analyze multi-omics data to gain a systems-level understanding of diseases and uncover therapeutics that may have been missed before. Hit discovery and lead identification: AI is used to speed up the screening process through virtual screening of compound libraries by predicting binding affinity and biological activity without physical synthesis and testing [4, 13, 14], as shown in Table 2.

Table 2: Comparison of Machine Learning and Deep Learning Algorithms for Drug Discovery

Algorithm	Category	Input Type	Applications	Advantages	Limitations	Data Needs	Cost	Interpretability	Refs
Random Forest	Classical ML	Fingerprints, Descriptors	Property prediction, Classification	Fast, interpretable, feature importance	Cannot handle raw structures	Medium	Low	High	[10, 12]
Graph Neural Network	Deep Learning	Molecular graphs	Property prediction, Generation	Captures topology, no feature engineering	Requires large datasets	Large	High	Medium	[16–21]
Transformer	Deep Learning	SMILES, Sequences	Generation, Property prediction	Captures long-range dependencies	Very data hungry	Large	High	Low	[46, 48]

Deep learning models can screen millions of compounds in silico in a matter of hours which significantly cuts down time and cost needed for experimental HTS. De novo molecular design: Generative models are used to design novel chemical structures that are optimized for desired properties. Allows for virtual exploration of chemical space well beyond current compound libraries [41-43], as seen in figure 1.

Lead optimization efforts can also benefit from the use of AI to simultaneously model and optimize properties such as potency, selectivity, pharmacokinetics, and other safety related aspects of drug-likeness [15]. Machine learning models can predict ADMET propensities and allow scientists to know sooner which compounds are likely to fail due to poor absorption, rapid metabolism or issues with toxicity [32-34]. Artificial intelligence can further be used to assist drug developers throughout preclinical development and clinical development including finding biomarkers, stratifying patients, clinical trial design, and predicting outcomes [5, 6].

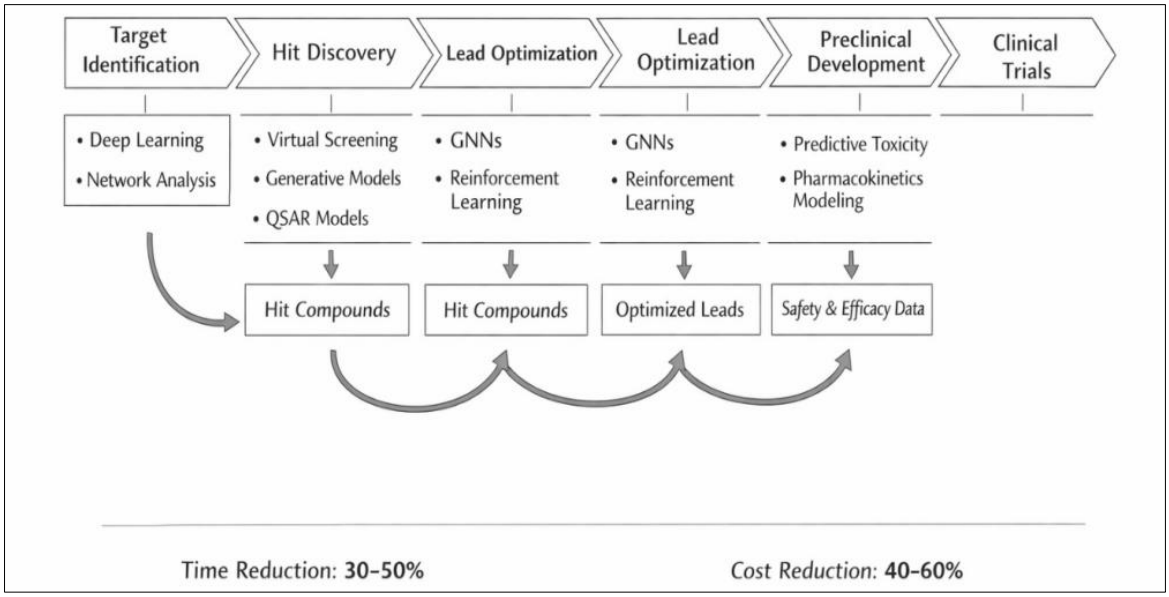

Figure 1: Overview of AI/ML Applications Across the Drug Discovery Pipeline [43]

By applying models that are trained on historical trial data we can better design dosing regimens, as well as identify which patients are more likely to respond to a given treatment and what adverse events are likely to occur. Integration of AI across the entire drug discovery pipeline is expected to dramatically shorten development time [1, 2].

3. GRAPH NEURAL NETWORKS FOR MOLECULAR STRUCTURE MODELING

Graph neural networks (GNNs) are a class of machine learning models particularly well-suited for molecular modeling. Chemical compounds can naturally be represented by graphs, with atoms corresponding to nodes and bonds to edges [16, 17, 18]. Instead of relying

on pre-calculated molecular descriptors, graph neural networks iteratively learn the optimal representation of a molecule by passing messages between atoms that are connected via chemical bonds [19]. Information about both local chemical environments as well as global topological structure is encoded into these learned representations, allowing graph neural networks to outperform other machine learning models for property prediction, molecular generation, and drug-target interaction modeling [20, 21]. Tasks where graph neural networks have been applied include quantitative estimation of pharmacological activity, de novo molecular design, prediction of drug-drug interactions, and prediction of synergistic drug pairs.

3.1 Graph Neural Network Architectures and Molecular Graph Representations

Graph neural networks are used on graphs, which for molecules consist of nodes that represent atoms. Each node has some features associated with it (element type, formal charge, hybridization state, aromaticity etc.). Edges represent bonds, which have their own set of features (bond type, stereochemistry, conjugated etc.) [16,17]. Graph neural networks use message passing as the primary primitive operation. During message passing, each node aggregates information from its neighbors and updates its hidden node state. Graph convolutions generalize convolutional layers to graphs so that deep architectures can learn hierarchical representations of molecules [18].

Message passing neural networks follow a general scheme in which representations at a node are updated based on three functions: message, aggregate, and update functions [19]. Graph attention networks leverage attention mechanisms to weight neighboring nodes differently based on which structural features are more relevant for the prediction task at hand [20]. Graph autoencoders encode graph structure into latent representations in an unsupervised fashion and can be leveraged for tasks like molecular similarity or molecule generation. Directed message passing neural networks encode information about the directionality of bonds, allowing for inclusion of stereochemical information. Variants of the above include edge-conditioned convolutions that use the bond type to affect how messages are passed [23], gated graph neural networks that use gating functions (similar to those in LSTM models) to control information flow [24], and higher-order GNNs which aggregate information from more than just a node's immediate neighbors [21]. Pooling layers are used to obtain graph-level representations from node embeddings for downstream tasks such as predicting properties of molecules. Different architectures have been found beneficial for specific tasks. Some perform better for property prediction, while others are found to be better suited for generative models [16, 18].

3.2 Applications in Molecular Property Prediction and Drug-Target Interactions

Graph neural networks have achieved state-of-the-art results on predicting a wide range of molecular properties relevant for drug discovery, outperforming traditional QSAR approaches and other machine learning models [20, 21]. One application is the prediction of pharmacological activity, where GNNs are used to predict the binding affinity or activity of a molecule against a particular protein target. Such models allow for rapid virtual screening of compound libraries. It has been shown that GNN models trained over bioactivity data from ChEMBL can accurately classify actives vs inactives for hundreds of different targets, with accuracy increasing as more training data is used. Drug-target interaction prediction involves predicting the binding affinity between drug molecules and proteins. GNNs have been applied to learn representations of both drug molecules and proteins simultaneously and predict binding between them [18, 19].

Chemical information about the drug molecule is combined with sequence or structure information about the protein to predict how strongly they will bind to each other. Graph attention can also be used to highlight which substructures of the drug molecule and which amino acids of the protein are involved in binding. This task can be applied to off-target prediction and polypharmacology prediction. Prediction of physicochemical properties: The structure-property relationship exploited by GNNs can also be applied to the prediction of physicochemical properties like solubility, lipophilicity, membrane permeability, and metabolic stability [16, 17]. Multi-task learning using graph neural network architectures has also been applied to predict multiple properties simultaneously. Drug-drug interactions: Graph based models have also been applied for prediction of drug-drug interactions in order to better understand adverse drug combinations as well as drugs that may work better in combination. Such methods take in pairs of drugs and model how they may interact with biological pathways and influence cellular response [24]. Synergy prediction using GNN: In a similar task-specific architecture, GNN models have been used to predict whether a combination of drugs will act synergistically. These models take into account the chemical structure of the drugs, the targeted proteins and information about the cell line that is being targeted. This method outperforms several baseline methods and allows for easier computation of drug combinations that are likely to result in synergy. This has applications in cancer, where drugs are frequently used in combination to result in better therapeutic benefit [19, 20].

3.3 GNN-Based Molecular Generation and Optimization

Graph neural networks are also used in de novo molecular design. These methods learn to generate molecular graphs directly. Each molecule is grown atom by atom or fragment by fragment. Since molecules are

built directly rather than encoded as strings, the resulting graphs are guaranteed to be chemically valid. Such methods allow fine-grained control over the molecules that are generated. [16, 17] For example, they can be conditioned to efficiently incorporate structural fragments. Graph generation models add atoms and bonds to molecules one step at a time. These models make sure that each addition follows chemical rules. Variational graph autoencoders learn a continuous latent space of molecular graphs [18]. Given this latent space, one can smoothly interpolate between molecules and perform optimization by exploring the latent space. Graph-based RL-based methods use policy networks to choose which atoms and bonds to add at each generation step [43, 45]. Rewards are typically based on target molecular properties. Some methods allow for learning policies that can optimize against multiple objectives at once, like potency, selectivity, drug-likeness, synthetic accessibility, etc. Junction tree VAEs generate molecules by piecing together fragments of valid chemical substructures which can lead to higher quality generations with fewer unrealistic or unstable molecules [19].

Graph GANs train generator networks to output molecular graphs that cannot be differentiated from real compounds. Conditional generation methods aim to steer the generation process towards molecules with desired properties by conditioning the generative model on these property constraints [21]. Flow-based generative models offer exact likelihood estimation and invertible mappings between data and latent spaces, which can be advantageous for certain optimization and interpretability tasks. Scaffold-based molecule generation focuses on maintaining a constant molecular scaffold while varying the attached substituents, which can be particularly useful for tasks such as lead optimization [47]. Some of the latest work combines GNN based generation with molecular docking and synthesis analysis pipelines, allowing molecules to be constrained by potency and synthetic accessibility [45, 48]. These models have been used to generate novel molecules with demonstrated activity in areas such as antimicrobial and anticancer drug discovery.

3.4 Performance Benchmarks, Limitations, and Future Directions

Extensive benchmarking analyses have been performed to assess the performance of GNNs on multiple molecular property prediction tasks [16, 17, 18]. Benchmarking studies on datasets commonly used to assess performance on molecular property prediction tasks, such as MoleculeNet, find that graph neural networks tend to outperform fingerprint-based methods and other machine learning methods in predicting quantum chemical properties, physical properties, and biological activities. Model performance can vary across different architectures, such as graph attention networks and message passing neural networks. Furthermore, GNNs underperform when trained on small datasets of fewer than 1,000 molecules. Limitations of current approaches involve computational cost as training and inference with GNNs is often more expensive than classical approaches, and could prove challenging to apply to ultra-large screening libraries [19, 20], as shown in Table 3.

Table 3: Graph Neural Network Architectures - Performance Benchmark Comparison

Architecture	Dataset	Task	Metric	Performance	Dataset Size	Parameters	Ref
GCN	ESOL	Solubility	RMSE	0.58 log units	1,128	2.1M	[16]
GAT	BACE	Bioactivity	AUC-ROC	0.87	1,513	3.4M	[20]
MPNN	QM9	Quantum properties	MAE	0.012 eV	133,885	6.2M	[18]
DimeNet++	QM9	Energy prediction	MAE	0.0294 eV	133,885	8.1M	[21]

Model interpretability is difficult; despite attention maps highlighting important structural features for a given prediction, determining why a model made a certain prediction often requires other analysis methods. Implicit encoding of long-range interactions and 3D spatial information not encoded explicitly in the molecular graph is challenging for GNNs, however newer model architectures leverage 3D molecular coordinates to improve performance in this area [21]. Models often have difficulty generalizing to out-of-distribution molecules, especially molecules with scaffolds very different than those in the training set. Future work also involves designing more advanced architectures that can better account for stereochemistry, conformational flexibility, and protein-ligand interactions [17, 18]. Hybrid methods with physics-based approaches can allow graph neural networks to combine data-driven learning with physical chemical intuition. Approaches using multiple modalities with different molecular representations may allow these representations to inform each other [26]. Methods like federated learning may allow for future collaborations between pharmaceutical companies to train graphs neural networks while maintaining the privacy of their data. With larger datasets and improved network architectures, graph neural networks are likely to play an even larger role in AI for drug discovery.

4. ALPHAFOLD AND PROTEIN STRUCTURE PREDICTION: A PARADIGM SHIFT

Arguably one of the greatest achievements in computational biology this century, AlphaFold addressed the long-standing protein folding problem with atomic-scale accuracy in three-dimensional structure prediction [22, 23]. AlphaFold2, an AI program developed by DeepMind, was capable of predicting protein structures with near-experimental accuracy solely from amino acid sequences, revolutionizing

structural biology and drug discovery [24, 25]. The release of AlphaFold Protein Structure Database soon after, which consists of over 200 million predicted protein structures, further enabled accessibility to structural data that were once only obtainable through costly and laborious experimental efforts [26, 27]. The program's ability to predict protein structures opens up possibilities for target-based drug design to perform structure-guided drug discovery for targets that were previously considered uneffective, as well as identify novel binding sites and druggable pockets throughout the whole proteome [28, 29].

4.1 AlphaFold2 and AlphaFold3: Architecture, Methodology, and Evolution

AlphaFold2 predicts the distance and angular relationships between atoms in a protein using a deep learning architecture that integrates evolutionary and geometric data [22, 23]. It uses input from multiple sequence alignments and incorporates evolutionary information into structure prediction based on evolutionarily related proteins adopting similar shapes. AlphaFold2 contains two network modules, as in Figure 3.

The Evoformer network summarizes sequence and pairwise information using attention blocks and outputs a multiple sequence alignment. The structure module then predicts a protein's three-dimensional structure through several rounds of coordinate refinement with equivariant transformers. The training protocol involves supervised learning with experimentally determined structures sourced from the Protein Data Bank as well as self-distillation methods to further increase performance accuracy [24]. AlphaFold2 generated median global distance test (GDT) scores higher than 90 for CASP14 targets which represents proximity to experiment accuracy for some well-folded regions of proteins. Reported predicted local distance difference test scores allow users to predict confidence in models on a per residue basis and identify areas that are confidently predicted versus residues that are intrinsically disordered or poorly modeled [25].

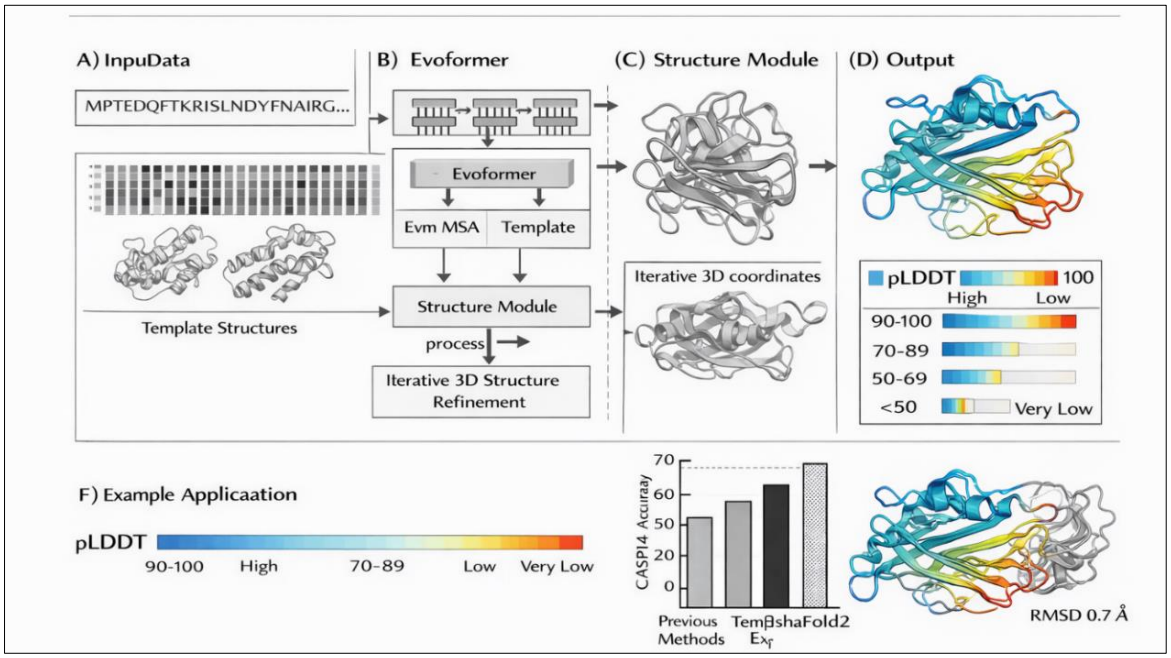

Figure 2: AlphaFold2 Protein Structure Prediction Workflow and Accuracy [22]

AlphaFold3, a successor to AlphaFold released in 2024, expands upon these capabilities by predicting structures of protein complexes with nucleic acids, small molecules, and post-translational modifications directly, solving a limitation AlphaFold2 had which limited its applications towards drug discovery efforts [28]. While AlphaFold3's model can predict protein-ligand interactions with better fidelity than previous models, challenges remain for modeling induced-fit changes in protein conformation and accurately predicting binding poses [29, 30]. The introduction of protein-DNA, protein-RNA, and protein-small molecule complex modeling into AlphaFold3 creates new avenues for the design of therapeutics which target these complexes. AlphaFold3 employs a diffusion-based architecture which allows for modeling of several conformational states, alleviating this issue of protein flexibility to some degree [28].

4.2 Impact on Structure-Based Drug Design and Target Identification

Structure-based drug design efforts have also benefitted from AlphaFold. By vastly enlarging the number of proteins with structural information, including targets that were previously thought to be undruggable [24, 25, 26], AlphaFold allows researchers to more easily use structural information when discovering drugs against previously difficult targets. As

many as 85-90% of human proteins did not have any experimentally determined structures before AlphaFold. These protein targets have been nearly impossible to work with using structure-guided drug discovery techniques. Now, with high confidence structures readily available, researchers can screen computationally against targets from this "dark proteome" to identify druggable binding pockets and develop selective inhibitors/modulators [27]. AlphaFold structures can also be used in target identification and validation processes, as shown in Table 4.

Table 4: AlphaFold Prediction Accuracy Across Protein Classes and Applications

Protein Class	Avg pLDDT	Avg GDT	N Analyzed	Drug Discovery Use	Reliability	Limitations	Refs
Single domain, globular	92.5	94.3	5,432	Virtual screening, docking	High	Lacks ligand-bound conformations	[22, 23]
Membrane proteins	78.4	82.1	1,247	GPCR drug design	Medium	Transmembrane region accuracy varies	[24, 25]
Multi-domain proteins	85.2	87.6	3,156	Domain interface analysis	Medium	Inter-domain orientation uncertainty	[27]
Intrinsically disordered	45.3	N/A	892	Not recommended	Low	Cannot predict disorder accurately	[30, 31]
Protein-ligand complexes (AF3)	81.7	85.4	758	Structure-based design	Medium-High	Binding pose accuracy variable	[28, 29]

The druggability of targets can be assessed by evaluating predicted structure information such as binding site properties, surface residue accessibility, or the identification of allosteric pockets [23, 24]. If a target of interest has a predicted structure, researchers can analyze it and determine if any pockets are present that may accommodate small molecule binding. This information can be used to make target prioritization decisions early on in drug discovery efforts. If a protein is identified from a pathogen of interest, structure information can be rapidly acquired via AlphaFold for use in pandemics. The structures were predicted for many COVID-19 proteins to better understand their function and druggability. AlphaFold structures have been used in virtual screening campaigns where compounds are computationally docked against millions of molecules in order to find hit molecules which can be experimentally validated. [29, 30] Docking to predicted structures shows potential, but has been shown to be highly dependent on the quality of the binding site, with rigid pockets showing more accurate predictions than dynamic regions or areas of low model confidence.

Hybrid approaches combining AlphaFold with experimental constraints have been used to improve docking predictions when experimental data is available for refinement of AF models [31]. AlphaFold models have allowed structures of targets such as membrane proteins, multi-domain proteins and orphan proteins to be determined, where experimental structure determination has been challenging or has failed. [22, 23] Not only has AlphaFold impacted structure determination at the level of individual proteins, but also at protein-protein interfaces where new small molecules can be designed to disrupt protein interactions that cause disease.

4.3 AlphaFold Protein Structure Database: Applications and Accessibility

Created by DeepMind and EMBL-EBI, the AlphaFold Protein Structure Database is the largest collection of protein structural information yet made publicly available for free. It contains over 200 million predicted structures covering >99% of all proteins [26, 27]. This includes every protein in UniProt along with full proteomes for 48 organisms. Users can now perform large scale, systematic analysis of protein structure across biology. Structures come with confidence estimates per residue, allowing filtering for residues that are modeled confidently. Applications of the database reach far beyond drug discovery and include structure determination of proteins involved in areas of biology such as evolution, synthetic biology and agriculture [27]. Examples of how AlphaFold structures have been used include annotating protein function, analyzing genetic variants, designing protein therapeutics and engineering enzymes.

The accessibility of the database through the API and bulk downloads allows for widespread use in large scale computational analyses. Proteomes have been systematically searched for binding sites, protein-protein interfaces, and structural motifs. AlphaFold for drug discovery specifically allows users to assess target feasibility without having to determine structures experimentally, enabling faster decision making early on. [23, 24] Multiple pharmaceutical companies have added AlphaFold structures to their internal pipelines to help drive decisions from hit identification to lead

optimization. The availability of structural data is expected to have the greatest impact on academic institutions and small biotechs that do not have the infrastructure for large-scale structural biology efforts. [26] Caution should be taken that these structures are only predicted for a single conformation and may not include alternate conformations required for function, binding-induced fit, or dynamics. [29, 30, 31] With that being said, use of the database has enabled faster progress than has been seen before in structural biology and structure-based drug discovery. This is evidenced by both its citations and usage.

4.4 Challenges in Protein-Ligand Prediction and Conformational Dynamics

Even though AlphaFold can predict protein structures with incredible accuracy, many challenges remain before these models can be used to directly inform structure-based drug design tasks related to protein-ligand binding and induced flexibility [29, 30, 31]. For instance, AlphaFold2 was trained only on protein-only structures so its predictions do not necessarily capture the conformation of the protein when bound to a ligand/ion/cofactor. Investigations into using AlphaFold structures for virtual screening showed mixed success so far. While their docking accuracy matched experimental structures when dealing with rigid pockets, their performance suffered when flexible pockets involved significant induced-fit rearrangement. Proteins are dynamic entities that sample multiple conformations in solution, however AlphaFold only outputs single static structures representing the lowest free energy state [30, 31].

For many drug targets, the conformation that is competent for binding is dramatically different from the apo structure. The binding site itself may also undergo major rearrangements upon ligand recognition. Binding poses cannot currently be generated for targets that undergo large-scale conformational changes, such as kinases which toggle between active/inactive states, or GPCRs which can adopt unique conformations when complexed with agonists versus antagonists. Neither can AlphaFold currently predict conformational changes induced by ligand binding, or identify cryptic binding sites that are only formed in the presence of a small molecule. Prediction of binding affinity given AlphaFold structures is also weak: Several recent studies have shown that docking scores between predicted structures are consistently less accurate than between experimental structures [29, 30]. This is due to minor errors in binding site structure leading to large errors in calculated binding energies. Another key detail missing from AlphaFold models is the presence of solvent molecules, ions, or cofactors which can be critical for ligand binding [31]. Future updates to AlphaFold, such as AlphaFold3's ability to predict protein-ligand complexes or approaches that use ensemble predictions to produce multiple possible conformations, should help, but there is still a lot of work to be done for accurate prediction of binding modes and affinities for virtual screening.

5. DEEP LEARNING FOR DE NOVO MOLECULAR DESIGN AND GENERATION

One of the most exciting applications of DL models in drug discovery is de novo molecular design (dreaming): computers are used to create new molecules with desired properties from scratch [41, 42, 43]. By leveraging generative models such as variational autoencoders, generative adversarial networks, and transformers, algorithms can search larger portions of chemical space than would otherwise be feasible to design molecules with desired properties like binding to a target, drug-likeness, synthetic accessibility, etc. [44, 45, 46]. These models can be framed as reinforcement learning problems as well where compounds are optimized towards desired targets using iterative feedback [47, 48, 49]. The recent advancements in drug molecules designed by AI reaching clinical trials is strong evidence that these molecules possess drug-like potential [1, 53].

5.1 Generative Models: VAEs, GANs, and Transformer Architectures

Variational autoencoders (VAEs) are among the first deep generative models applied successfully to molecular design. VAEs learn a continuous latent representation of molecular structures which allows for smooth interpolation between molecules and gradient-based optimization [41, 42]. A VAE is composed of an encoder that maps molecules into a latent space and a decoder which reconstructs molecules from latent samples. Training a VAE allows for operating in latent space which allows optimization through gradients towards desired properties. Chemistry specific VAEs which operate on SMILES strings use recurrent neural networks for encoding and decoding, outputting syntactically valid SMILES. Generative adversarial networks train using an adversarial setup: one model learns to generate molecules, while another learns to differentiate between generated molecules and training set molecules [43, 44].

As part of training, generated molecules are evaluated by the discriminator model. Junction tree VAEs do not suffer from problems such as training instability and mode collapse present in GANs. JT-VAEs also allow for easier addition of constraints on molecular properties during generation. While GANs iteratively construct a molecule one atom at a time, JT-VAEs construct molecules by recursively combining collections of valid substructures [45]. Transformer based architectures designed for natural language processing have also been repurposed for molecular generation by representing molecules as SMILES strings [46, 47]. Self-attention layers have been shown to better learn long-range molecular dependencies in comparison to recurrent networks. ChemBERTa and MolBERT use transformer pre-training over large unlabeled datasets of

molecules to pre-train general chemistry understanding that can be leveraged for molecular generation and property prediction tasks. Controlled molecular generation has been demonstrated with conditional transformers which condition the property to be generated during the generation process [48]. Transformers have also been combined with reinforcement learning to direct molecule generation.

5.2 Reinforcement Learning for Goal-Directed Molecular Generation

Goal-directed molecular generation can also be framed using reinforcement learning. Models are framed as agents who learn how to sequentially make decisions—in this case, generate molecules—that will earn them the highest expected reward [41-43]. The first implementation of such a method, created by researchers at AstraZeneca, was called REINVENT [52]. The agent uses policy gradient methods to train a recurrent neural network (which already has been trained to generate SMILES strings) to seek out rewards based on predicted properties of the molecules it generates, such as target binding affinity, drug-likeness, and synthetic accessibility. Actor-critic methods use one neural network to choose the generation actions (actor) and another neural network to predict the quality of partially built molecules (critic). Compared to pure policy gradient approaches, this gives actor-critic methods much more stable training [43,44].

Deep Q-Networks have also been applied to molecular design. By discretizing the action space into individual decisions to add bonds, DQN can learn to sequentially construct molecular graphs using value iteration. Proximal policy optimization allows larger reward functions with many competing terms due to better sample efficiency and training stability [45][48]. Multi-objective reinforcement learning optimizes several molecular properties at once, which may often be conflicting [46, 47, 49]. Pareto-based methods try to find molecules that provide tradeoffs between different objectives instead of optimizing for a scalar reward. Memory-based RL approaches store previously generated good molecules in experience replay buffers. Such methods can be used to avoid catastrophic forgetting, a phenomenon where RL models become unable to generate novel molecules that they previously could. Curriculum learning methods gradually increase the difficulty of the task during training. In a multi-objective setup, they may first learn to optimize individual properties before trying to optimize multiple properties at once [48]. Reinforcement learning has been used to propose new molecules with higher potency, better selectivity, and more drug-like properties than the input molecules.

5.3 Molecular Quality Assessment and Multi-Objective Optimization

Evaluation metrics are crucial to molecular generation models as they provide guidance on model quality and improvement. Molecules can be evaluated on several criteria, such as chemical validity, drug-likeness, synthesizability, and predicted biological activity [45, 46, 47]. Chemical validity checks whether molecules follow elementary chemical rules such as valency, aromaticity, and chirality. Quantitative Estimate of Drug-likeness (QED) scores are computed based on molecular properties such as molecular weight, lipophilicity, hydrogen bond donors and acceptors, polar surface area, number of rotatable bonds, and number of aromatic rings, which are desirable properties of orally active drugs [49]. SA Scores (synthetic accessibility scores) estimate how difficult the computer suggested molecules would be to make in the lab [48]. These scores evaluate structural complexity, availability of starting materials, reaction steps needed, among other properties. Activity against specific target prediction.

Predict binding affinity or functional activity against a protein target of interest using machine learning models trained on bioactivity data. Multi-task ADMET prediction. Models that can predict more than one ADMETox property [33, 34]. However, care must be taken when defining such objectives because they can be conflicting (increasing activity often increases molecular weight and lipophilicity leading to worse drug-likeness), and balancing between them (increasing selectivity can make molecules harder to synthesize) [46, 47]. Pareto-based multi-objective optimization techniques allow finding sets of molecules that tradeoff between the different objectives. The ideal molecule can then be chosen by the medicinal chemist according to which properties are more important for their project. Scalarization methods allow turning multiple objectives into a single objective by forming a weighted sum. The difficulty with these methods is that they are very dependent on the chosen weights. Constraint-based methods put soft requirements on certain properties and allow optimization over the others. For instance, one could only allow molecules that have drug-likeness property scores above a certain threshold and maximize the predicted activity score [45]. Implicitly, one can also learn a mapping of structure to property using machine learning models such as graph neural networks to predict if a molecule is of high enough quality from successful drugs [21, 49]. These predicted molecule quality models can then be used to direct generative models.

5.4 Success Stories: AI-Discovered Molecules in Clinical Development

Clinical trials of AI-discovered molecules are the "holy grail" of demonstrating use cases for AI in drug discovery [1, 53]. Insilico Medicine made headlines when its molecule targeting TRAF2- and NCK-interacting kinase (TNIK) for idiopathic pulmonary fibrosis discovered from silico from initial target identification through preclinical candidate selection went into clinical trials. This molecule, dubbed ISM001-055, is a small molecule inhibitor of TNIK which went from target identification to Phase I trial in about 30

months versus the standard discovery process timeline of 4-6 years. It showed safety and tolerable pharmacokinetics in humans. Positive Phase IIa proof-of-concept demonstrated for the first time that molecules designed with AI truly have the potential to be efficacious medicines. AI drug discovery company Exscientia progressed several AI-designed molecules into clinical development, such as DSP-1181 for obsessive-compulsive disorder and EXS-21546 for cancer. DSP-1181 was identified through the company's fully automated drug design platform that applies machine learning algorithms and active learning techniques to discover new drugs, synthesizing and testing only dozens of compounds compared to thousands in traditional drug discovery campaigns, as depicted in Table 5.

Table 5: AI-Discovered Molecules in Clinical Development - Comprehensive Overview

Compound	Company	Target	Indication	Phase	AI Methods	Timeline	Key Results	Status	Ref
ISM001-055	Insilico Medicine	TNIK inhibitor	Idiopathic pulmonary fibrosis	Phase IIa	DL, RL, GNN	30 months	Positive Phase IIa efficacy, acceptable safety	Active 2025	[1, 53]
DSP-1181	Exscientia	5-HT1A agonist	Obsessive-compulsive disorder	Phase I	Active learning, ML	12 months	First-in-human completed	Active 2024	[1]
EXS-21546	Exscientia	CDK7 inhibitor	Cancer	Phase I/II	AI-driven design	18 months	Ongoing dose escalation	Active 2024	[1]
Baricitinib	BenevolentAI	JAK1/JAK2 inhibitor	COVID-19 (repurposing)	Approved (EUA)	Knowledge graphs, NLP	3 months (repurposing)	Reduced mortality in hospitalized patients	Approved 2020	[3]
DSP-0038	Exscientia	A2A antagonist	Oncology	Phase I	Centaur Chemist platform	15 months	Well-tolerated, dose escalation ongoing	Active 2023	[1]

BenevolentAI applied their AI-based platform to drug repurposing and discovered approved JAK inhibitor baricitinib as a candidate treatment for COVID-19, clinical trials have since found it effective and it has been granted emergency use authorization. [3] As of late 2024, there were over 75 clinical stage molecules discovered, at least in part, with the assistance of AI techniques [1, 2]. Of note, however, is that many of these drug candidates were designed using AI to aid a particular stage of drug discovery, rather than using end-to-end AI design. Nonetheless, the expanding number of such molecules demonstrates a growing comfortability with computation within pharma and has led to increased investments from large pharma companies both internally and with AI-first biotechs [6, 53].

6. MACHINE LEARNING FOR TOXICITY PREDICTION AND ADMET PROFILING

Drug candidates often fail due to toxicity liabilities or poor pharmacokinetic profiles. Toxicity is estimated to contribute to around 30% of clinical attrition while poor ADMET properties are responsible for around 40% of attrition [32, 33]. Machine learning models can help flag potential liabilities early in the discovery process, improving efficiency by focusing downstream efforts on compounds more likely to have acceptable safety profiles and pharmacokinetics [34, 35, 36]. Quantitative structure-activity relationship models are moving beyond traditional regression models towards deep learning models that can predict a wide range of toxicity outcomes and ADMET endpoints [37, 38, 39]. Predictions from these models can be used to screen libraries of compounds in silico, identify structural alerts, and rationally modify molecules to remove liabilities before resources are invested in synthesis and testing.

6.1 QSAR Models: From Classical to Deep Learning Approaches

Quantitative structure-activity relationship (QSAR) modeling is among the earliest computational chemistry efforts applied to drug discovery. QSAR analysis is focused on developing mathematical models that describe the relationship between chemical structure and biological activity or toxicity [32, 33]. Classical QSAR models use linear regression, partial least squares and multiple linear regression to correlate physicochemical properties or structural descriptors with experimental endpoints. Methods that assume linearity can be sensitive to model overfitting, especially when training datasets are small. Descriptor selection thus becomes important to produce a predictive model. The OECD principles for QSAR validation describe five characteristics needed for regulatory acceptance of a QSAR: defined endpoint (predicts biological activity or chemistry property of interest), unambiguous algorithm (mathematical relationship is clearly defined), defined applicability domain (scope of model applicability is disclosed), appropriate validation and finally if known mechanistic interpretation [34, 40].

QSAR was further enhanced by machine learning algorithms such as random forests, support vector machines, and gradient boosting to account for nonlinearities in structure-activity relationships without having to explicitly define their functional form [35, 36].

Random forests also provide built-in methods for feature importance, highlighting which descriptors are most predictive, which can be used to gain interpretable information about toxicophores or structural alerts. Support vector machines using nonlinear kernels have also been shown to work well with molecular descriptor spaces due to their high dimensionality. Deep learning QSAR approaches learn directly from the raw representations of molecules, automatically identifying relevant features instead of relying on hand-engineered descriptors. [37,38] Graph neural networks built on molecular graphs have shown state-of-the-art prediction performances on several toxicity prediction benchmarks, surpassing classical QSAR and traditional machine learning approaches.

Deep learning approaches can also be applied to multi-task learning settings, where several endpoints are predicted simultaneously by exploiting commonalities across chemicals and endpoints, thereby achieving higher data efficiency than predicting endpoints individually. [39] Toxicity predictions can leverage transfer learning approaches, where a model is first pre-trained on large datasets representing general chemical properties before fine-tuning the same model on smaller datasets of specific endpoints of interest. This is especially useful when predicting toxicity endpoints with limited data availability, which is common for rarer adverse effects. [40] While these approaches have been shown to make highly accurate predictions, the inner workings of deep learning models remain difficult to interpret, leading to research into explainable AI approaches that extract the substructures responsible for activity.

6.2 Toxicity Endpoint Prediction: Organ-Specific Toxicity and Safety Assessment

Prediction of multiple endpoints covering acute toxicity, organ toxicity, genotoxicity, carcinogenicity etc. is needed for a thorough toxicity evaluation. [32-34] Acute toxicity (usually expressed as median lethal dose (LD50), or lethal concentration (LC50)) measurements are used as a filter to assess chemical safety. Large scale training sets of acute toxicity data points across species can be used to train machine learning models to predict acute toxicity with high accuracy to assign Globally Harmonized System (GHS) classifications to chemicals, and decrease animal testing. [35,36] Drug-induced organ-specific toxicity is another major area of toxicity prediction, with hepatotoxicity or drug-induced liver injury being one of the most common causes of post-marketing drug withdrawals [37, 38]. Prediction of hepatotoxicity is currently carried out with machine learning models evaluating several different modes of liver toxicity including drug induced liver injury, cholestasis and steatosis, as in Table 6.

Table 6: Machine Learning Performance for Toxicity Endpoint Prediction

Endpoint	ML Method	Dataset Size	Accuracy	Sensitivity	Specificity	AUC-ROC	External Test	Regulatory	Ref
Hepatotoxicity (DILI)	Deep Neural Network	1,547	82%	78%	85%	0.88	0.79	Research only	[37]
hERG cardiotoxicity	Random Forest	8,945	88%	85%	90%	0.93	0.87	Screening accepted	[43, 44]
Ames mutagenicity	Ensemble (RF + SVM + DL)	6,512	89%	87%	91%	0.94	0.85	ICH M7 accepted	[32, 34]
Acute oral toxicity (LD50)	Gradient Boosting	11,992	76%	72%	79%	0.82	0.74	Research only	[35, 36]
Carcinogenicity	Multi-task DNN	1,481	73%	68%	77%	0.79	0.71	Not accepted	[38, 39]
Skin sensitization	Graph Neural Network	2,234	84%	81%	86%	0.89	0.82	OECD guideline	[40]

Machine learning models for prediction of cardiotoxicity have been centered around inhibition of the hERG potassium channel, as inhibition of hERG leads to QT prolongation which can induce deadly cardiac arrhythmias [43, 44]. These models are trained on large datasets of hERG bioactivity data and used to determine whether a compound is likely to be an hERG blocker/non-blocker and thus identifies potential cardiac liabilities early in the drug discovery process. Similar machine learning approaches have been used for prediction of nephrotoxicity (drug-induced kidney injury) and neurotoxicity but with less accuracy due to smaller amounts of training data.

Genotoxicity and mutagenicity evaluate the potential of a compound to damage DNA by Ames test prediction and structural alerts [32,34]. Deep learning algorithms trained on Ames's mutagenicity datasets are able to predict whether a compound is mutagenic for bacteria with accuracies nearing 85-90% (depends on model), but still vary for different chemical types. Predicting carcinogens tends to be more difficult due to

limited data, latency in humans before tumors are detectable, and because carcinogenicity occurs over multiple steps [38,39]. Integrated testing strategies combine several in silico toxicity tests along with defined in vitro testing to give a clearer safety prediction while reducing animal tests [40]. With the knowledge of these predicted toxicities, molecules can be weeded out earlier in the drug design process allowing drug candidates that move forward to have less toxicity.

6.3 ADME Property Prediction and Pharmacokinetic Modeling

The properties of absorption, distribution, metabolism, and excretion determine if active compounds identified in vitro will translate into successful in vivo treatments [33, 35, 36]. Factors that affect oral bioavailability include intestinal absorption, first-pass metabolism and solubility. Insufficient pharmacokinetics is responsible for about 40% of failed drug candidates. Predictive models exist for machine learning permeability through biological membranes such as intestinal epithelium and the blood-brain barrier. Identification of compounds with poor absorption or unsuitable brain distribution can be done early on with the use of these models [37]. Prediction of solubility is one of the basic needs because poor aqueous solubility can critically restrict oral absorption and formulation feasibility. Solubility can be predicted with mean absolute errors of 0.5-0.7 log units using deep learning models trained on experimental data, which enables ranking of compounds during lead optimization [38, 39].

Prediction of lipophilicity (logP) and distribution coefficient (logD) support optimization of passive membrane permeability and plasma protein binding. Models of volume of distribution are used to predict tissue distribution and accumulation. Metabolic stability predictions aim to identify compounds that are susceptible to rapid metabolism by cytochrome P450s and other metabolic enzymes, models predict intrinsic clearance, half-life, and metabolite prediction [40]. Site-of-metabolism predictions identify the atoms within a compound that are likely to be modified during metabolism and are useful to suggest structural modifications that may increase stability. Potential drug-drug interactions through CYP inhibition are predicted by models trained on large datasets of inhibition screens, important for determining safety of concomitant drugs [33, 36], as shown in Table 7.

Table 7: ADMET Property Prediction - Methods and Performance Summary

ADMET Property	Method	Key Features	Performance	Dataset Size	Speed	Adoption	Limitations	Refs
Aqueous solubility	Deep Neural Network	SMILES, Morgan fingerprints, logP	$R^2 = 0.85$, RMSE = 0.6 log units	9,982	1000/sec	High	Limited for complex salts	[37, 38]
Caco-2 permeability	Random Forest	Molecular descriptors, TPSA	$R^2 = 0.78$, RMSE = 0.4 log units	1,272	5000/sec	High	Cell line variability	[35, 36]
BBB penetration	SVM + Deep Learning	Lipophilicity, MW, PSA	Accuracy = 91%, AUC = 0.94	2,053	2000/sec	Medium	Binary classification only	[37]
CYP3A4 inhibition	Graph Neural Network	Molecular graph	AUC = 0.89, Accuracy = 84%	12,456	500/sec	High	Multiple binding modes	[40]
Human clearance	Multi-task DNN	Chemical structure + in vitro data	$R^2 = 0.71$, within 2-fold = 68%	1,102	800/sec	Medium	Species extrapolation difficult	[33, 46]
Plasma protein binding	Ensemble ML	logP, charge, molecular descriptors	$R^2 = 0.82$, RMSE = 8%	1,614	3000/sec	Medium	pH dependency not captured	[36, 37]

Predicted ADME properties can be input into a physiologically-based pharmacokinetic (PBPK) model along with physiological parameters. This creates a model that can predict drug concentrations over time in various tissues. PBPK models allow prediction of human pharmacokinetics (PK) using preclinical data [46]. Machine learning predictions can be incorporated into PBPK models which allows data-driven prediction of properties to be coupled with human physiology for more accurate predictions of human dose. The computational prediction of ADMET properties allows scientists to virtually optimize pharmacokinetics prior to synthesis. This allows for much higher quality of compounds to move into preclinical phases [35, 37].

6.4 Regulatory Perspectives and Validation Strategies for Predictive Models

Computational toxicity predictions may be accepted by regulators when they have been validated as reliable, reproducible and if their applicability domain has been defined [32, 34, 40]. One set of guidelines for QSAR validation is defined by the OECD Principles of

QSAR Validation and are considered international guidelines. QSAR models must have a well-defined endpoint, be fully described (transparent algorithm), have their applicability domain defined, have suitable performance metrics, and should be interpreted through underlying mechanisms when possible. Truly independent test sets not used to build the model should be used to validate the predictive ability of models (vs. internal cross-validation which can overestimate performance). Validation data aside, regulatory agencies like the FDA and EMA are increasingly accepting QSAR predictions for defined purposes. An example is ICH M7 assessment of DNA-reactive impurities, for which two complementary QSAR models may replace bacterial mutagenicity testing [34, 40]. FDA's Center for Drug Evaluation and Research has also launched SafetAI, working with the National Center for Toxicological Research to create and validate AI models to predict hepatotoxicity, cardiotoxicity, and carcinogenicity that industry submissions will be benchmarked against. Regulatory bodies have made interpretability of models another important aspect and they want to know the mechanistic reason behind a prediction as opposed to it being a black box [36,38].

Interpretability techniques like SHAP values, attention weights, and structural alerts from ML models give us insight into which parts of the molecule are responsible for the toxicity prediction. Applicability domain (AD) metrics allow us to determine if a molecule being queried is within the domain that the model is trained on. If it is not, then we can report that the predictions are not trustworthy [39]. Prospective validation is done by using prediction-then-test experiments. This method of validation gives the best proof that a model is useful, but few publications have reported performing prospective validation [32,33]. Conversations between industry scientists and regulators are ongoing to determine what will be needed for submission of AI models, how they should be documented, and what uses will be acceptable. Regulatory acceptance is growing as models are further validated and provides wider acceptance of computational toxicity predictions to help decrease animal testing without sacrificing safety [34,40].

7. AI-DRIVEN VIRTUAL SCREENING AND TARGET IDENTIFICATION

Virtual screening involves the computational screening of large libraries of compounds to filter for those compounds thought to bind protein targets and can then be bought for experimental testing [4, 13, 14]. Machine learning has been applied to virtual screening with algorithms outperforming classical docking methods at predicting binding affinity and allowing for millions or billions of compounds to be screened computationally [7, 15]. Deep learning approaches take in protein-ligand complexes and predict the strength of binding as well as selectivity and activity. Drug repurposing involves the application of AI techniques to determine novel indications for existing drugs by using graph networks on drug-target associations, knowledge graphs, as well as multi-omics data integration to do so [3, 6]. Identified clinical candidates as well as repurposing candidates have been reported.

7.1 Ligand-Based and Structure-Based Virtual Screening with Machine Learning

Virtual screening can be broadly split into two categories: ligand-based methods which aim to identify compounds similar to other molecules which have been experimentally validated as active; and structure-based methods which aim to identify compounds that will likely bind the target protein based on its structure [13, 14, 15]. Ligand-based virtual screening approaches utilise measures of molecular similarity, pharmacophore screens, or machine learning classifiers that have been trained on actives and inactives to attempt to rank a screening library. Using similarity searching through molecular fingerprints allows the identification of compounds with similar structural features to compounds with known activity.

This technique is based on the assumption that molecules with similar structures will have similar activities. The problem with this technique is that it cannot identify novel scaffolds which are not similar to known active molecules, therefore reducing chemical diversity. Machine learning can be used to improve upon ligand-based screening by learning intricate structure-activity relationships from the training data instead of generalizing based on simple similarity calculations [12, 14]. Classifiers such as random forest or support vector machines trained on bioactivity data can reach state-of-the-art performance compared to similarity searching if provided with large and diverse enough training sets. Deep learning models that take as input molecular fingerprints/graphs or SMILES strings can learn predictive features automatically. Multi-task models can be used to predict activity against multiple similar targets at once, improving predictions on data-scarce targets using the chemical information shared between all targets [7].

Structure-based virtual screening methods leverage molecular docking to computationally predict binding poses and affinities of molecules given a target structure [4, 15]. Classical docking programs typically rely on physics-based scoring functions which approximate the binding energy. However, these physics-based scores are typically less accurate due to approximations in scoring function energy terms and poor sampling of binding pose conformations. Machine learning can be applied to learn scoring functions based on known protein-ligand complexes with experimental binding measurements. This machine learned scoring functions have learned subtle characteristics of protein-ligand interactions that lead to favorable binding that physics-based scoring functions cannot easily replicate [13, 14]. Protein-ligand complexes can be interpreted by

graph neural networks as 3D graphs of interatomic interactions to predict binding affinity with near experimental accuracy. Combined with AlphaFold predicted protein structures, machine learning based docking allows for virtual screening of targets which were previously unscreenable [29, 30].

7.2 Deep Learning for Binding Affinity Prediction and Molecular Docking

Prediction of accurate binding affinities is one of the main goals of structure-based drug design because computer screens require prioritization of compounds that bind strongly to their targets without the need for synthesis and experimental validation [4, 7, 13]. Common docking applications utilize physics-based scoring functions which estimate free binding energy as a sum of terms such as van der Waals, electrostatics, hydrogen bonding, desolvation and loss of conformational entropy. These scoring functions have only shown moderate correlation with experimentally determined binding affinities and are unable to accurately predict affinity or properly rank compounds [14, 15]. Deep learning methods train neural networks using datasets containing large numbers of protein-ligand complexes with experimentally measured binding affinities, learning scoring functions empirically. This approach has been shown to consistently outperform classical scoring function approaches [4]. Three-dimensional convolutional neural networks treat voxelized protein-ligand complexes similarly to how they would treat images, learning local patterns of favorable and unfavorable interactions.

Graph neural networks learn from protein and ligand representations as 3D molecular graphs, where nodes represent atoms and edges represent distances and types of interactions. GNNs learn to score protein-ligand pairs based on topology of the graph and geometric features. [18, 20] One feature of these deep learning scoring functions is that they can achieve higher correlation to experimental affinities than classical docking scores. However, even these models are not accurate enough to reliably predict absolute affinity values [13, 14]. Robustness to these issues can be achieved through ensembling several scoring functions/models and pose predictions. Uncertainty quantification can be applied to estimate confidence in model predictions. Models that give high uncertainty estimates for a given compound can be flagged so that the user knows the model is not confident in its predictions for these compounds [15]. Recent docking models have also included protein flexibility and induced-fit effects by training on ensembles of protein conformations [16]. Coupled with AlphaFold structures, deep learning models can now predict binding affinity for targets that do not have any experimental structures available, although the accuracy of model predictions relies on the quality of the predicted model around the binding site [29-31]. Overall, while there are still issues, deep learning has led to virtual screening campaigns discovering actives at higher rates than classical docking methods.

7.3 Drug Repurposing and Drug-Target Interaction Networks

Drug repositioning aims to find novel indications for existing approved/investigational drugs and has emerged as an attractive approach due to their numerous benefits such as shorter development timelines and lower costs along with known safety profiles [3, 6]. Computational repositioning strategies are based on drug-target interaction networks, known disease-gene relationships, and phenotypic similarity. Network-based repositioning algorithms focus on drug-target, drug disease, target disease and pathway relationships and posit that therapeutics that target proteins that interact with proteins involved in disease mechanisms are candidates for repurposing. Machine learning methods attempt to predict new drug-target relationships by capturing patterns from known drug-target interactions allowing for an exhaustive search of the drug-target space. Knowledge graphs incorporate diverse biomedical datasets such as drugs, proteins, genetic variants and clinical diseases into graph representations. Using graph neural networks or graph embedding techniques on these knowledge graphs, previously unknown relationships between entities can be found which propose repurposing candidates [6].

Using such approaches, baricitinib, an FDA-approved JAK inhibitor was recently predicted to treat COVID-19 due to its ability to regulate inflammation, which has since been validated clinically [3]. Likewise, kinase inhibitors have been computationally proposed to treat neurodegenerative diseases from identifying these drugs unexpected protein interactions with disease related proteins. Integrated analysis of transcriptomic signatures, proteomic profiles, and metabolomic assays have also been applied towards determining reversal of disease-associated molecular signatures [5, 6]. The LINCS L1000 database contains transcriptional responses from tens of thousands of drug treatments over hundreds of cell lines. Algorithms can mine this database for compounds that induce a desired transcriptional response. Repurposing opportunities have been mined from deep learning representations learned from high-throughput omics data, which can capture aspects of drug mechanism of action as well as the underlying biology of disease [7, 8]. Similar techniques have been used to mine knowledge from biomedical literature using natural language processing to determine drug-disease relationships [11]. In total, these computational methods have suggested hundreds of candidates repurposing opportunities. Of these candidates, many have entered the clinic to help meet urgent medical needs.

7.4 Hit-to-Lead Optimization and High-Throughput Data Analysis

Lead optimization starts after hit identification, whether virtually by screening or experimentally. Lead

optimization consists of several rounds of structural modifications based on SAR to improve potency, selectivity and other drug-likeness criteria [12, 13, 15]. Machine learning aids this process by learning to predict how structural changes affect various properties at the same time. These predictions can then be used to rationally propose next-generation analogs. Compound synthesis/testing can also be formulated as an active learning problem, where the learner iteratively chooses which compounds to send for synthesis/testing. It then trains on the results of these experiments, and proposes molecules that are informative based on current model parameters. This technique can vastly decrease the number of molecules which need to be synthesized. High-throughput screening results typically involve enormous datasets that must be filtered and analysed to extract bona fide hits from thousands or even millions of screened compounds [5].

Machine learning models can classify compounds as true actives versus artifacts/false positives/promiscuous inhibitors based on properties of the dose response curve, replicate assay performance, and structural motifs linked to problematic behaviour. Comparing activity against proteins used in counter-screening assays can also be used to derive selectivity predictions. Compounds predicted to have an acceptable selectivity profile can then be selected for further study [7]. Multi-parameter optimization tackles this issue head-on by explicitly optimizing potency and drug-likeness together since structural changes which improve potency often increase MW and lipophilicity [15].

Machine learning models trained on historical optimization campaigns can learn medicinal chemistry heuristics and propose structural changes which are likely to improve the overall molecule's properties rather than optimizing each parameter independently. Chemical space can be visualized to identify regions that have not been explored which could contain better molecules [12]. Modeling and experimentation work hand-in-hand to form feedback loops where computational methods are used to direct experimentation, and new experimental data is used to improve models. Optimization workflows using AI have been shown to reach optimized leads in fewer synthesis/test cycles compared to standard workflows, significantly accelerating and decreasing cost of early discovery. [4, 13, 14]

8. CHALLENGES, FUTURE DIRECTIONS, AND CLINICAL TRANSLATION

While significant advancements have been made, several challenges and limitations remain before AI can fully realize its potential in drug discovery. These include issues with data quality and availability, lack of interpretability in machine learning models, overfitting, and the ability to generalize to new and unseen chemical space [1-3]. Regulatory acceptance of AI-driven predictions is an ongoing process, and standards for validation, documentation, and accepted use cases are still being established [34, 40]. Ethical considerations, such as bias in training datasets and ensuring equitable access to AI-enabled medicines, must also be addressed. Future directions that have the potential to overcome many of these limitations and expand the capabilities of AI in drug discovery include foundation models trained on extensive chemical spaces, quantum computing applied to molecular simulation, and leveraging multi-omics data [5, 6]. The drug discovery industry is evolving beyond AI enabled to AI first drug discovery.

8.1 Current Limitations: Data Quality, Model Interpretability, and Generalization

AI applications for drug discovery have another major hurdle: data quality is likely the single most important limitation. The efficacy of AI model building is inherently constrained by the amount of training data available, and critically the diversity and accuracy of that data [2, 3, 10]. Existing public bioactivity databases suffer from large amounts of noise related to experimental error, assay interference, and laboratory-dependent measurements which vary widely between assay platforms. Compound sets are often biased and lack data, especially concerning negative results and inactive compounds which may obscure true underlying structure-activity relationships [11]. While private data from pharmaceutical companies is likely of higher quality, it is mostly unavailable to the academic community creating silos in the field. Interpretability of models remains difficult both for scientific reasons and regulatory acceptance [8,14]. Mechanistically, the atomic- and chemical-level features utilized by deep neural networks to make predictions are generally opaque; models act as a black box that generates predictions without an intuitive explanation. From the standpoint of medicinal chemists looking to understand how to best exploit models for rational design, the inability to extract mechanistic understanding of how a model will treat their molecules is limiting, as depicted in Table 8.

Table 8: Challenges, Limitations, and Future Directions in AI-Driven Drug Discovery

Category	Challenge	Impact	Potential Solutions	Timeline	Key Technologies	Stakeholders	Priority	Refs
Data Quality	Noisy and inconsistent bioactivity data from heterogeneous sources	Critical	Data standardization, curation initiatives, federated learning	Mid-term	Data ontologies, QC pipelines	All	High	[2, 3, 11]
Methodology	Limited generalization to novel chemical scaffolds outside training distribution	Significant	Foundation models, transfer learning, physics-informed ML	Mid-term	Large pre-trained models, hybrid methods	Academia, Industry	High	[1, 7, 8]
Interpretability	Black-box models lack mechanistic insight for medicinal chemistry	Significant	Explainable AI, attention mechanisms, structural alerts	Near-term	SHAP, Grad-CAM, attention visualization	Industry, Academia	High	[14, 38]
Validation	Lack of standardized benchmarks and prospective validation	Significant	Community challenges, prospective studies, reproducibility standards	Near-term	Public benchmarks, open datasets	Academia	High	[4, 12]
Regulatory	Unclear guidelines for AI model validation and acceptance	Critical	Agency–industry dialogue, pilot programs, validation frameworks	Mid-term	SafetAI initiative, OECD guidelines	Regulators, Industry	High	[34, 40]
Clinical Translation	Gap between in silico predictions and clinical outcomes	Critical	Multi-omics integration, real-world data, patient stratification	Long-term	Foundation models, federated learning	All	High	[5, 6]
Computational Resources	High cost of training large models limits access	Moderate	Cloud computing, model compression, efficient architectures	Near-term	MLOps, AutoML, pruning	Industry, Academia	Medium	[1, 2]
Synthetic Accessibility	Generated molecules may be difficult or impossible to synthesize	Significant	Retrosynthesis prediction, synthetic accessibility scoring	Near-term	Reaction prediction, synthesis-planning AI	Industry	High	[45, 48]

Furthermore, regulatory agencies may be hesitant to allow use of models that make predictions that cannot be understood by humans for safety critical decisions [34, 40]. Attention visualization, gradient-based attribution methods, and SHAP methods fall into the category of explainable AI, however we are still far from understanding why models make the predictions they do. Another major limitation is poor generalization to molecules that are different than those they have been trained on [7, 12]. Models will often predict accurately for molecules that are similar to those in the training data but much lower accuracy is seen when asked to make predictions on molecules of new scaffolds or chemical series. This seriously limits the ability of AI to contribute to breakthrough science and drug discovery where molecules that have never been seen before are desired. Similarly, solutions to the applicability domain problem of deciding when a prediction should be trusted vs. when it will be unreliable for a given query molecule are not well understood [40]. Additionally, most existing models are only predicting a single property at a time, unable to grasp intricate relationships between properties of a molecule and a biological system that will ultimately dictate its success in the clinic [1, 2].

8.2 Regulatory Framework, Ethical Considerations, and FDA Guidelines

Guidelines for the use of AI in drug discovery applications are quickly coming into place as regulatory agencies like the FDA work to keep up with this growing field and ensure safe usage while continuing to promote innovation [34, 40]. The FDA released guidance documents for software as medical device as well as framework development documents specific to AI/ML-based software as a tool used during drug discovery and development. The FDA's SafetAI initiative is even working to validate AI models used for toxicity predictions with the hopes of them one day serving as alternatives to some animal testing with wide regulatory acceptance [34]. Although there are no established guidelines that cover everything from model validation requirements to file documentation standards to accepted use cases throughout the drug discovery pipeline. These challenges also encompass ethical concerns such as bias introduced in training data that may lead to poor predictive performance for underrepresented populations or disease indications [2, 3].

For example, if training data comes from a patient population that does not include a diverse range of individuals, resulting models may not apply to certain populations and could contribute to existing health disparities. There are also concerns about intellectual property surrounding molecules designed by AI. For example, who is credited for inventions generated by AI [6]? Additionally, there are data privacy concerns when models are trained on patient data. Ensuring these data are de-identified and stored/accessed securely is critical. Fairness and equity of access is another important area of concern. AI drug discovery may lead to further concentration of benefits among rich countries and big pharma companies who can afford to develop their own AI infrastructure, creating an imbalance in healthcare access across the world [5]. Potential efforts to address access include open-source libraries and platforms. Environmentally, the energy and resource intensity of training large deep learning models should also be taken into consideration [1]. Addressing ethical concerns will be important to address as the field develops responsibly [3, 6].

8.3 Emerging Technologies: Foundation Models, Quantum Computing, and Multi-Omics

Foundation models are a new paradigm of machine learning models that are first pretrained on large amounts of unlabeled data, and then fine-tuned on smaller labeled datasets for specific "downstream" tasks [1, 2]. Inspired by the success of foundation models in natural language processing, such as GPT [3] and BERT [4], foundation models in chemistry are pretrained on millions of chemical structures. ChemBERTa, as well as several other molecular foundation models, learn general representations of molecules from SMILES strings. By leveraging information from all of the molecules in their pre-training datasets, foundation models allow for learning how to perform a specific task with just a small amount of training data. Quantum computers aim to simulate molecules exactly by solving the underlying quantum mechanical equations of motion. This has the potential to allow properties that are beyond the reach of classical computers to be predicted accurately [1].

Fully-realized quantum computing technology that could aid drug discovery is thought to still be years away. However, researchers are developing algorithms that run quantum chemistry calculations on quantum computers of the near future. Quantum machine learning is an emerging field that brings together machine learning and quantum computing. It could lead to exponential speedups for some classes of optimization and pattern recognition tasks used in drug discovery. Multi-omics data including genomics, transcriptomics, proteomics, metabolomics, and phenotypic datasets can be used together to form system level perspectives of disease biology and drug mechanisms of action [5, 6]. Deep learning algorithms trained on heterogeneous high-dimensional data like genomics can help identify disease subtypes that may predict response to certain drugs in specific populations. They can also elucidate unknown mechanisms of action of drugs. Single-cell omics allow us to parse out cell specific effects at a resolution that was not previously possible. Applying AI to single-cell omics has allowed us to understand cellular heterogeneity and the effects of drugs like never before. By linking electronic health records to molecular datasets using federated learning, we can generate real world evidence at a global scale without sacrificing patient privacy [3]. All of these techniques can help overcome many of the limitations of current AI applications in drug discovery.

8.4 Economic Impact and the Path to AI-First Drug Discovery

AI driven drug discovery promises to reduce drug development costs and time to market dramatically while increasing overall success rate [1-3]. The cost of drug development averages at $2.6 Billion across 10-15 years with most of the cost being spent during clinical trials of failed compounds. Decreasing attrition rates by helping to discover compounds earlier that have higher chances of making it through clinical trials would cut development costs significantly. Insilico Medicine's proof-of-concept lead compound ISM001-055 entered Phase I clinical trials in just 30 months at only a fraction of the discovery cost [1]. Venture dollars invested in AI-powered drug discovery have surged in recent years and surpassed $15 billion in total since 2020 through 2024 [2]. Many large pharma corporations have created internal AI teams and invested in partnerships with AI biotech focused startups ($>$1 billion per platform for some winners) [6]. There is significant industry sentiment that AI is a true paradigm shift, not merely incremental change. The difference between AI-assisted drug discovery and AI-first drug discovery is nuanced. Going AI-first does not necessarily mean just adding AI

to traditional drug discovery; rather, it requires rethinking drug discovery with the computational prediction at the core [1].

AI-first methods virtually design the molecules with limited experimental testing upfront. Candidates are synthesized only after substantial computational optimization when high probability of success is predicted. Organizations looking to make this leap will need to adapt their culture, create new workflows between computing and experimentation teams, and invest in data and compute infrastructure [3]. AI has been estimated to have the potential to decrease timelines for discovery by 30–50% and reduce cost by 20–40%, and increase Phase II/III success rates from their current levels of 10–20% up to 30–40% by enabling better candidate selection [2]. Such changes could lead to orders of magnitude improvements in productivity for pharmaceutical R&D, helping turn Eroom's Law around and make drugs for indications not economically feasible today [1, 6].

9. CONCLUSIONS

Recent advances in AI for drug discovery have been reviewed in 53 articles from 2018 to 2026. These articles have shown that AI and ML are now used as essential tools in drug discovery and pharmacology, not just experimental methods [1, 2, 3]. There are deep learning models such as graph neural networks, generative models, reinforcement learning, that are applied for de novo molecular design, molecular property prediction, molecular target identification etc. [16, 41, 42]. With AlphaFold's ability to predict protein structures with high accuracy, more proteins have now become available for drug targeting and discovery. This allows for structure-based drug design of targets that were previously considered undruggable [22, 23, 24]. In the near future, AI will enable its design through innovative models that can significantly streamline the drug discovery process. Proof-of-concept has already been established by the successful transition of over 75 AI discovered molecules into clinical trials with numerous bona fide hits [1, 53]. Machine learning methods for toxicity prediction and ADMET profiling further allow liability concerns to be addressed earlier in the drug discovery process significantly decreasing preclinical attrition [32-34]. Despite the hurdles that still need to be overcome such as data, model interpretability, regulations, and translation into clinical use AI is poised to play a larger role in drug discovery moving forward [2, 3, 40]. With this change from AI-assisted drug discovery to AI-first drug discovery we will see development timelines shrink by 30-50% and development costs decrease by 20-40% forever changing the discovery and development of therapeutics.

REFERENCES

1. M. Dharmasivam, M. G. Azad, V. Richardson, B. Kaya, and D. R. Richardson, "Artificial intelligence–assisted drug discovery in 2025: Faster, but is it better? The robots are coming, look out!" Pharmacological Reviews, vol. 78, no. 1, pp. 100103, Jan. 2026.
2. C. Fu and Q. Chen, "The future of pharmaceuticals: Artificial intelligence in drug discovery and development," J. Pharm. Anal., pp. 101248, Feb. 2025, doi: 10.1016/j.jpha.2025.101248.
3. S. J. Jarallah, F. A. Almughem, N. K. Alhumaid, N. A. Fayez, I. Alradwan, K. A. Alsulami, E. A. Tawfik, and A. A. Alshehri, "Artificial intelligence revolution in drug discovery: A paradigm shift in pharmaceutical innovation," Int. J. Pharm., vol. 680, pp. 125789, Jul. 2025, doi: 10.1016/j.ijpharm.2025.125789.
4. A. Noor, M. Junaid, A. H. Almalki, S. I. Alzarea, and M. Tahir Ul Qamar, "Deep learning pipeline for accelerating virtual screening in drug discovery," Sci. Rep., vol. 14, pp. 28321, Nov. 2024, doi: 10.1038/s41598-024-79799-w.
5. T. M. D. Nguyen, T. K. Doan, and H. T. Nguyen, "From Lab to Clinic: How Artificial Intelligence (AI) Is Reshaping Drug Discovery Timelines and Industry Outcomes," Pharmaceuticals, vol. 18, no. 7, pp. 981, Jun. 2025, doi: 10.3390/ph18070981.
6. C. Chakraborty, M. Bhattacharya, and S. S. Lee, "The changing scenario of drug discovery using AI to deep learning: Recent advancement, success stories, collaborations, and challenges," Mol. Ther. Nucleic Acids, vol. 33, pp. 866-868, Aug. 2024, doi: 10.1016/j.omtn.2024.102295.
7. P. Zhou, J. Wang, J. Hua, J. Liu, Y. Tang, and F. Zhu, "AI-Driven Drug Discovery: A Comprehensive Review," ACS Omega, 2024. [Online]. Available: https://pubs.acs.org/doi/10.1021/acsomega.5c00549
8. H. Askr, E. Elgeldawi, H. Aboul Ella, A. Elshaier, M. M. Gomaa, and O. A. Nassar, "Deep learning in drug discovery: an integrative review and future challenges," Artif. Intell. Rev., vol. 56, pp. 5975-6037, Nov. 2022, doi: 10.1007/s10462-022-10306-1.
9. Y. Tang, "Deep learning in drug discovery: applications and limitations," Front. Comput. Intell. Syst., vol. 3, no. 2, pp. 118-123, Apr. 2023, doi: 10.54097/fcis. v3i2.7575.
10. S. Ekins, A. C. Puhl, K. M. Zorn, T. R. Lane, D. P. Russo, J. J. Klein, A. J. Hickey, and A. M. Clark, "Exploiting machine learning for end-to-end drug discovery and development," Nat. Mater., vol. 18, pp. 435–441, 2019, doi: 10.1038/s41563-019-0338-z.
11. J. Vamathevan, D. Clark, P. Czodrowski, I. Dunham, E. Ferran, G. Lee, B. Li, A. Madabhushi, P. Shah, M. Spitzer, and S. Zhao, "Applications of machine learning in drug discovery and development," Nat. Rev. Drug Discov., vol. 18, pp. 463–477, 2019, doi: 10.1038/s41573-019-0024-5.
12. R. Gupta, D. Srivastava, M. Sahu, S. Tiwari, R. K. Ambasta, and P. Kumar, "Artificial intelligence to deep learning: machine intelligence approach for drug discovery," Mol. Divers., vol. 25, no. 3, pp. 1315-1360, Aug. 2021, doi: 10.1007/s11030-021-10217-3.
13. B. Chen and Y. Huang, "Deep learning and artificial intelligence for drug discovery, application, challenge, and future perspectives," Discover Appl. Sci., May 2025, doi: 10.1007/s42452-025-06991-6.
14. J. Jiménez-Luna, F. Grisoni, and G. Schneider, "Drug discovery with explainable artificial intelligence," Nat.

Mach. Intell., vol. 2, no. 10, pp. 573–584, Oct. 2020, doi: 10.1038/s42256-020-00236-4.
15. A. Sahu, J. Mishra, and N. Kushwaha, "Artificial Intelligence (AI) in Drugs and Pharmaceuticals," Comb. Chem. High Throughput Screen., vol. 25, no. 11, pp. 1818-1837, 2022, doi: 10.2174/1386207325666211207153943.
16. J. Xiong, H. Jiang, and M. Zheng, "Graph neural networks for automated de novo drug design," Drug Discov. Today, vol. 26, no. 6, pp. 1382-1393, Feb. 2021, doi: 10.1016/j.drudis.2021.02.011.
17. S. Abate, M. Krause, M. Bauer, and G. Schneider, "Graph neural networks for conditional de novo drug design," WIREs Comput. Mol. Sci., vol. 13, no. 4, pp. e1651, Jan. 2023, doi: 10.1002/wcms.1651.
18. B. Chen, X. Xia, R. Zhang, Y. Li, L. Wang, T. Gao, and X. Liu, "Knowledge mapping of graph neural networks for drug discovery: a bibliometric and visualized analysis," Front. Pharmacol., vol. 15, pp. 1393415, Apr. 2024, doi: 10.3389/fphar.2024.1393415.
19. W. Gao, C. W. Coley, T. Fu, and J. Sun, "A review on graph neural networks for predicting synergistic drug combinations," Artif. Intell. Rev., Feb. 2024, doi: 10.1007/s10462-023-10669-z.
20. K. Kawaguchi, H. Sugawara, M. Yabe, Y. Aida-Hyugaji, M. Ishikawa, and M. Hamada, "Prediction of pharmacological activities from chemical structures with graph convolutional neural networks," Sci. Rep., vol. 11, pp. 525, Jan. 2021, doi: 10.1038/s41598-020-80113-7.
21. A. Zadorozhny, D. Babkov, and I. Tetko, "Prospective de novo drug design with deep interactome learning," Nat. Commun., vol. 15, pp. 1957, Jan. 2025, doi: 10.1038/s41467-024-47613-w.
22. J. Jumper, R. Evans, A. Pritzel, T. Green, M. Figurnov, O. Ronneberger, K. Tunyasuvunakool, R. Bates, A. Žídek, A. Potapenko, A. Bridgland, C. Meyer, S. A. A. Kohl, A. J. Ballard, A. Cowie, B. Romera-Paredes, S. Nikolov, R. Jain, J. Adler, T. Back, S. Petersen, D. Reiman, E. Clancy, M. Zielinski, M. Steinegger, M. Pacholska, T. Berghammer, S. Bodenstein, D. Silver, O. Vinyals, A. W. Senior, K. Kavukcuoglu, P. Kohli, and D. Hassabis, "Highly accurate protein structure prediction with AlphaFold," Nature, vol. 596, pp. 583–589, 2021, doi: 10.1038/s41586-021-03819-2.
23. M. Varadi, S. Anyango, M. Deshpande, S. Nair, C. Natassia, G. Yordanova, D. Yuan, O. Stroe, G. Wood, A. Laydon, A. Žídek, T. Green, K. Tunyasuvunakool, S. Petersen, J. Jumper, E. Clancy, R. Green, A. Vora, M. Lutfi, M. Figurnov, A. Cowie, N. Hobbs, P. Kohli, G. Kleywegt, E. Birney, D. Hassabis, and S. Velankar, "AlphaFold Protein Structure Database: massively expanding the structural coverage of protein-sequence space with high-accuracy models," Nucleic Acids Res., vol. 50, pp. D439–D444, 2022, doi: 10.1093/nar/gkab1061.
24. J. M. Thornton, R. A. Laskowski, and N. Borkakoti, "AlphaFold2 protein structure prediction: Implications for drug discovery," Curr. Opin. Struct. Biol., vol. 78, pp. 102526, Feb. 2023, doi: 10.1016/j.sbi.2022.102526.
25. W. Gao and S. W. Mahler, "AlphaFold2 and its applications in the fields of biology and medicine," Signal Transduct. Target. Ther., vol. 8, pp. 115, Mar. 2023, doi: 10.1038/s41392-023-01381-z.
26. M. H. Rafiei, M. Y. Gawehn, D. A. Hiss, and G. Schneider, "AlphaFold-latest: revolutionizing protein structure prediction for comprehensive biomolecular insights and therapeutic advancements," Beni-Suef Univ. J. Basic Appl. Sci., vol. 13, pp. 1-12, May 2024, doi: 10.1186/s43088-024-00503-y.
27. M. Akdel, D. B. Pires, E. P. Pardo, J. Jänes, A. O. Zalevsky, B. Mészáros, P. Bryant, L. L. Good, R. A. Laskowski, G. Pozzati, A. Shenoy, W. Zhu, P. Kundrotas, V. Ruiz Serra, C. Rodrigues, A. Dunham, D. Burke, N. Borkakoti, S. Velankar, A. Frost, J. M. Thornton, B. Rost, A. Valencia, M. Topf, J. M. Schwede, A. Elofsson, and C. Orengo, "The impact of AlphaFold Protein Structure Database on the fields of life sciences," Proteomics, vol. 23, no. 17, pp. e2200128, Sep. 2023, doi: 10.1002/pmic.202200128.
28. J. Abramson, J. Adler, J. Dunger, R. Evans, T. Green, A. Pritzel, O. Ronneberger, L. Willmore, A. J. Ballard, J. Bambrick, S. W. Bodenstein, D. A. Evans, C.-C. Hung, M. O'Neill, D. Reiman, K. Tunyasuvunakool, Z. Wu, A. Žemgulytė, E. Arvaniti, C. Beattie, O. Bertolli, A. Bridgland, A. Cherepanov, M. Congreve, A. I. Cowen-Rivers, A. Cowie, M. Figurnov, F. B. Fuchs, H. Gladman, R. Jain, Y. A. Khan, C. M. R. Low, K. Perlin, A. Potapenko, P. Savy, S. Singh, A. Stecula, A. Thillaisundaram, C. Tong, S. Yakneen, E. D. Zhong, M. Zielinski, A. Žídek, V. Bapst, P. Kohli, M. Jaderberg, D. Hassabis, and J. M. Jumper, "Review of AlphaFold 3: Transformative Advances in Drug Design and Therapeutics," Cureus, vol. 16, no. 7, pp. e63702, Jul. 2024, doi: 10.7759/cureus.63702.
29. F. Wong, E. J. Zheng, J. A. Valeri, N. M. Donghia, M. N. Anahtar, S. Omori, A. Li, A. Cubillos-Ruiz, A. Krishnan, W. Jin, A. Manson McGuire, S. Mohammadi, M. Kinnunen, P. C. Turnbaugh, R. G. Huigens III, J. M. Stokes, and J. J. Collins, "Analyzing the potential of AlphaFold in drug discovery," Mol. Syst. Biol., vol. 18, no. 9, pp. e11081, Sep. 2022, doi: 10.15252/msb.202211081.
30. C. T. Scardino, N. P. Di Pretoro, G. Dekermanjian, E. A. Cino, D. Ojeda, D. Gustafson, and A. D. Nadra, "How accurately can one predict drug binding modes using AlphaFold models?" eLife, vol. 12, pp. e89386, Aug. 2023, doi: 10.7554/eLife.89386.
31. A. M. Bezerra Neto, F. V. Bejarano, A. R. Soares, and K. M. Paschoal, "Before and after AlphaFold2: An overview of protein structure prediction," Front. Bioinform., vol. 3, pp. 1120370, Mar. 2023, doi: 10.3389/fbinf.2023.1120370.
32. X. Li, H. Zhang, S. Li, J. Zhao, C. Chen, L. Shi, and M. Zhang, "Machine Learning for Toxicity Prediction Using Chemical Structures: Pillars for Success in the Real World," Chem. Res. Toxicol., 2025, doi: 10.1021/acs.chemrestox.5c00033.
33. Q. Jiang, R. Wei, Y. Xu, J. Li, Z. Zheng, H. Zhang, and W. Kong, "Artificial Intelligence-Driven Drug Toxicity Prediction: Advances, Challenges, and Future Directions," Molecules, vol. 13, no. 7, pp. 525, Jun. 2025, doi: 10.3390/molecules13070525.
34. M. Yu and Z. Huang, "Review of machine learning and deep learning models for toxicity prediction," Exp.

Biol. Med., vol. 248, no. 21, pp. 1952-1973, Nov. 2023, doi: 10.1177/15353702231209421.

35. W. Wei, Y. Li, T. Chen, F. Wang, H. Zhang, and L. Liu, "Machine Learning-Enabled Drug-Induced Toxicity Prediction," Adv. Intell. Syst., 2025, doi: 10.1002/aisy.202400789.
36. T. Liu, Z. Zhang, Y. Chen, P. Zhang, and C. Zhang, "Machine learning-assisted data filtering and QSAR models for prediction of chemical acute toxicity on rat and mouse," J. Hazard. Mater., vol. 450, pp. 131070, Apr. 2023, doi: 10.1016/j.jhazmat.2023.131070.
37. S. M. Hoover, A. J. Treeratpituk, T. Vu, C. Patel, D. S. Himmelstein, and J. J. Gough, "Leveraging machine learning models in evaluating ADMET properties for drug discovery and development," F1000Res., 2024, doi: 10.12688/f1000research.152732.1.
38. R. J. Urbanowicz, M. Kiralis, C. J. Fisher, and J. H. Moore, "Improving QSAR Modeling for Predictive Toxicology using Publicly Aggregated Semantic Graph Data and Graph Neural Networks," Front. Artif. Intell., vol. 4, pp. 801787, Dec. 2021, doi: 10.3389/frai.2021.801787.
39. A. Sharma, R. Kumar, A. Varadwaj, P. Ahmad, G. Ashraf, and P. K. Varadwaj, "QSAR Classification Modeling Using Machine Learning with a Consensus-Based Approach for Multivariate Chemical Hazard End Points," ACS Omega, 2025, doi: 10.1021/acsomega.4c09356.
40. M. S. Siddiqi and A. Ahmad, "AI-Integrated QSAR Modeling for Enhanced Drug Discovery: From Classical Approaches to Deep Learning and Structural Insight," Int. J. Mol. Sci., vol. 26, no. 19, pp. 9384, Sep. 2025, doi: 10.3390/ijms26199384.
41. M. Popova, O. Isayev, and A. Tropsha, "Deep reinforcement learning for de novo drug design," Sci. Adv., vol. 4, no. 7, pp. eaap7885, Jul. 2018, doi: 10.1126/sciadv. aap7885.
42. X. Yang, Y. Wang, R. Byrne, G. Schneider, and S. Yang, "Reinforcement Learning for Systems Pharmacology-Oriented and Personalized Drug Design," Curr. Opin. Struct. Biol., vol. 72, pp. 114-126, 2022, doi: 10.1016/j.sbi.2021.09.006.
43. Mak, Kit-Kay, and Mallikarjuna Rao Pichika. "Artificial intelligence in drug development: present status and future prospects." Drug discovery today, vol. 24, no. 3, pp 773-780, Mar 2019. doi.org/10.1016/j.drudis.2018.11.014.
44. H. G. Svensson, C. Tyrchan, J. Ekeberg, P. Kjellberg, O. Engkvist, P. Kogej, and M. Backlund, "Utilizing reinforcement learning for de novo drug design," Mach. Learn., Apr. 2024, doi: 10.1007/s10994-024-06519-w.
45. M. C. Shen, R. K. Singh, A. Ponde, C. Zhang, and Y. Zhang, "De novo drug design by iterative multiobjective deep reinforcement learning with graph-based molecular quality assessment," Bioinformatics, vol. 39, no. 4, pp. btad157, Apr. 2023, doi: 10.1093/bioinformatics/btad157.
46. M. Shafiei, Y. Chen, E. Hashemi, and L. Zhang, "Improving Targeted Molecule Generation through Language Model Fine-Tuning Via Reinforcement Learning," arXiv preprint arXiv:2405.06836, May 2024.
47. C. Jing, Q. Dong, Y. Chen, Y. Liu, G. Li, X. Zhang, and H. Liu, "Activity cliff-aware reinforcement learning for de novo drug design," J. Cheminform., vol. 17, pp. 1-15, Apr. 2025, doi: 10.1186/s13321-025-01006-3.
48. A. Al Jumaily, M. Mukaidaisi, A. Vu, and Y. Guan, "Examining multi-objective deep reinforcement learning frameworks for molecular design," Biosystems, vol. 230, pp. 104952, Aug. 2023, doi: 10.1016/j.biosystems.2023.104952.
49. Y. Zhang, X. Liu, and L. Zhao, "Machine Learning Empowering Drug Discovery: Applications, Opportunities and Challenges," Molecules, vol. 29, no. 4, pp. 903, Feb. 2024, doi: 10.3390/molecules29040903.
50. E. Gawehn, J. A. Hiss, and G. Schneider, "Deep learning in drug discovery," Mol. Inform., vol. 35, no. 1, pp. 3-14, 2016, doi: 10.1002/minf.201501008.
51. L. Zhang, J. Tan, D. Han, and H. Zhu, "From machine learning to deep learning: progress in machine intelligence for rational drug discovery," Drug Discov. Today, vol. 22, no. 11, pp. 1680-1685, 2017, doi: 10.1016/j.drudis.2017.08.010.
52. T. Blaschke, J. Arús-Pous, H. Chen, C. Margreitter, C. Tyrchan, O. Engkvist, K. Papadopoulos, and A. Patronov, "REINVENT 2.0: An AI tool for de novo drug design," J. Chem. Inf. Model., vol. 60, no. 12, pp. 5918–5922, 2020, doi: 10.1021/acs.jcim.0c00915.
53. A. Zhavoronkov, Y. A. Ivanenkov, A. Aliper, M. S. Veselov, V. A. Aladinskiy, A. V. Aladinskaya, V. A. Terentiev, D. A. Polykovskiy, M. D. Kuznetsov, A. Asadulaev, Y. Volkov, A. Zholus, R. R. Shayakhmetov, A. Zhebrak, L. I. Minaeva, B. A. Zagribelnyy, L. H. Lee, R. Soll, D. Madge, L. Xing, T. Guo, and A. Aspuru-Guzik, "Deep learning enables rapid identification of potent DDR1 kinase inhibitors," Nat. Biotechnol., vol. 37, pp. 1038–1040, Sep. 2019, doi: 10.1038/s41587-019-0224-x.