**International Journal of Biology and Life Sciences**
ISSN: 2957-9511 | Vol. 15, No. 3, 2026

# AI in Drug Discovery: Applications, Challenges and Future Prospects

**Yue Peng ***

College of Engineering, China Pharmaceutical University, Nanjing, Jiangsu, 211100, China
*** Corresponding author Email:** 13352032263@163.com

**Abstract:** Traditional pharmaceutical R&D is constrained by substantial financial investment, lengthy development cycles, and a high probability of failure. Artificial intelligence (AI) is now being incorporated into multiple stages of the pharmaceutical pipeline, including target identification, molecular design, synthesis planning, and clinical research. This paper reviews how machine learning, deep learning, natural language processing, and related computational methods are being applied across the drug discovery process. Particular attention is given to AlphaFold-based protein structure prediction, AI-supported virtual screening, generative chemistry, retrosynthetic planning, digital pathology, and the use of real-world clinical data. The review also considers limitations that are often hidden by strong computational performance, such as incomplete training data, limited interpretability, weak interoperability, uncertain external validity, and the continuing need for laboratory and clinical confirmation. In addition, several practical examples from industry and academic research are discussed to connect technical principles with their actual use in pharmaceutical development. Future progress is likely to depend on multimodal data integration, explainable models, robotic design-make-test-analyze cycles, privacy-preserving collaboration, and regulatory frameworks that evaluate both model performance and the quality of the evidence generated. AI should therefore be understood as an augmentation technology: it can prioritize hypotheses and accelerate iteration, but it cannot replace biological reasoning, experimental judgment, or clinical responsibility.

**Keywords:** Drug Discovery; AI Model; Deep Learning.

## 1. Introduction

Developing a new therapeutic is historically a slow, capital-intensive, and high-risk undertaking. The process usually begins with a biological hypothesis and continues through target validation, lead generation, optimization, preclinical testing, and clinical trials. At each stage, a large number of candidates are removed because of weak efficacy, toxicity, poor pharmacokinetics, manufacturing difficulty, or insufficient clinical benefit. Bringing a new chemical entity to market often requires more than a decade of research, while the aggregate investment can reach billions of dollars. The central difficulty is not simply the volume of experiments. It is also the need to connect observations produced at different biological scales, from molecular interactions to patient outcomes.

Recent developments in artificial intelligence (AI), especially machine learning (ML), deep learning (DL), and natural language processing (NLP), are changing how these problems are approached. AI systems can compare large numbers of molecular structures, infer patterns from genomic and proteomic datasets, extract evidence from scientific literature, and support the selection of patients for clinical trials. In practical terms, such systems help researchers decide which experiment should be conducted next. They do not eliminate uncertainty, but they may reduce the number of low-value experiments and make the overall workflow more focused.

This paper presents an overview of AI applications throughout the drug discovery pipeline. It examines representative technologies and company platforms, connects computational methods with practical cases, and discusses the limitations that influence their reliability. The analysis also considers common challenges shared by different stages and outlines future developments. The purpose is not to treat AI as an independent substitute for pharmaceutical science, but to examine how computational tools can be integrated with chemical, biological, and clinical expertise.

## 2. Applications of AI Across the Drug Discovery Pipeline

### 2.1. Target Identification and Validation

The first major task in drug discovery is to identify and validate disease-associated biomolecules, such as proteins, genes, RNA molecules, or signaling pathways. A promising target must be connected to disease biology, accessible to an intervention, and sufficiently safe when modulated. AI algorithms can examine heterogeneous biomedical data, including genomics, proteomics, scientific literature, and clinical registries. By linking these sources, they help researchers recognize relationships that may be difficult to detect through manual reading alone. Nevertheless, a statistical association is not equivalent to a causal mechanism. Target hypotheses still require experiments in cells, animals, and, ultimately, human studies.

#### 2.1.1. Applications in Target Identification

Tools such as AlphaFold, RoseTTAFold, and ESMFold have expanded the role of computational structural biology. AlphaFold2 uses deep learning to predict three-dimensional protein structures from amino acid sequences alone [1]. Its predicted structures have made structural information available for proteins that are difficult to study through crystallography or cryo-electron microscopy. The AlphaFold Protein Structure Database, which contains a very large number of predicted structures, has consequently become a useful starting point for structural investigation. Researchers have also used AlphaFold2 to predict structures covering a

substantial proportion of the human proteome [2].

The practical value of this technology can be seen in infectious disease research. In malaria drug development, AlphaFold-based predictions have been used to examine proteins involved in the invasion of red blood cells by Plasmodium falciparum. Predicted pockets can guide the selection of compounds for target-directed design, although the resulting hypotheses must be confirmed experimentally. Similar structural information may support work on protein families associated with neurodegeneration, particularly when conventional structural methods have not yet produced a complete model.

RoseTTAFold employs a three-track neural network that considers sequence, distance, and coordinate information to balance accuracy and speed [3]. It has been applied to protein structure and interaction analysis, including the study of the SARS-CoV-2 spike protein and its host-receptor interface. ESMFold uses a protein language model and can infer structures rapidly from sequence information [4]. Its large-scale application to metagenomic proteins has helped researchers inspect previously uncharacterized protein families. In parallel, companies such as BenevolentAI use knowledge graphs to connect biomedical publications, disease associations, and experimental findings. Such systems have been used to identify potential targets for complex diseases such as amyotrophic lateral sclerosis (ALS) [5].

#### 2.1.2. Limitations

Despite these advances, important limitations remain. AI models depend strongly on the quality and diversity of the data used for training. Their performance may decline for proteins with few known homologs, flexible conformations, membrane localization, or unusual post-translational modifications. A static structural model also cannot fully represent functional plasticity, transient protein-protein interactions, ligand-induced conformational changes, or allosteric dynamics [6]. Therefore, structural prediction should be treated as a research aid rather than a final description of biological reality. AI-generated target hypotheses require biochemical, cellular, and disease-model validation, which remains time-consuming and expensive.

## 2.2. Molecular Design and Optimization

After a target has been validated, researchers search for small molecules or biologics that can modulate it with adequate potency, selectivity, and safety. Traditional high-throughput screening can examine many compounds, but it requires substantial material, instrumentation, and time. AI-supported virtual screening and de novo design provide a complementary route. These methods rank existing compounds or generate new structures before synthesis, allowing experimental resources to be directed toward a smaller set of candidates.

Schrödinger combines physics-based calculations with machine learning in a drug development platform. The platform has supported the design of clinical-stage integrin inhibitors [7], including MORF-057 and SGR-1505. Its collaboration with Bristol Myers Squibb illustrates how computational chemistry can be incorporated into a broader industrial pipeline. Atomwise uses AtomNet, a convolutional neural network designed for molecular screening. The system screened millions of compounds and contributed to the identification of a potential OTUD7B inhibitor [8]. It has also been applied in antiviral and neurological disease programs. Exscientia developed CentaurAI for de novo molecular design and designed DSP-1181, a 5-HT1A receptor agonist for obsessive-compulsive disorder. The candidate became an important example of an AI-originated molecule entering human trials [9].

#### 2.2.1. Applications

AI technologies are mainly applied in three aspects of molecular design. First, virtual screening uses molecular descriptors, docking information, and graph neural networks (GNNs) to prioritize compounds from large libraries. Second, generative chemistry uses generative adversarial networks (GANs), variational autoencoders (VAEs), diffusion methods, or language-model approaches to propose structures with desired properties. Third, multi-parameter optimization uses reinforcement learning or Bayesian optimization to balance potency, selectivity, solubility, permeability, metabolic stability, and toxicity. In practice, the most useful workflow is usually iterative: the model proposes compounds, chemists assess synthetic feasibility, experiments generate new measurements, and the model is updated.

#### 2.2.2. Limitations

AI molecular design still faces several problems. Models may generate structures that resemble the training data, which limits exploration of chemical space. A structure can also appear attractive computationally but be unstable, difficult to synthesize, or unsuitable for formulation. Synthetic accessibility is therefore essential rather than an optional final check. Noisy biological measurements and inconsistent assay conditions further reduce the reliability of ADMET prediction [10]. In addition, deep learning models often provide limited explanations for their recommendations. Chemists may understand which molecule is ranked highly, but not always why a particular structural change is expected to improve performance.

## 2.3. Synthesis Planning and Route Optimization

#### 2.3.1. Applications

A promising molecule has limited value if it cannot be prepared safely, economically, and at a suitable scale. AI systems use repositories of chemical reactions to propose retrosynthetic pathways. In retrosynthetic analysis, a target molecule is repeatedly divided into simpler building blocks. Candidate routes can then be compared according to reaction yield, cost, number of steps, reagent availability, and safety. AI may also predict reaction outcomes and suggest reaction conditions.

Synthia, IBM RXN, and BenchSci are representative platforms. Synthia combines rule-based chemistry with machine learning for retrosynthetic analysis and has been used to design routes for natural products and pharmaceutical compounds. IBM RXN applies Transformer models to reaction prediction and retrosynthesis from SMILES strings [11]. In studies of Buchwald-Hartwig cross-coupling reactions, high top-three prediction accuracy has been reported for selected reaction matrices. BenchSci focuses on extracting experimental protocols and product information from the literature, thereby helping researchers locate relevant evidence more efficiently.

#### 2.3.2. Limitations

The main limitations of AI synthesis planning arise from incomplete data and simplified assumptions. Chemical databases contain many successful reactions, whereas failed

experiments are less likely to be published. This positive-reporting bias can make a reaction appear more reliable than it is in a real laboratory. AI-planned routes may also assume ideal mixing, unlimited time, standard equipment, and straightforward purification. In practice, scale-up introduces heat-transfer problems, mass-transfer limitations, equipment incompatibility, waste-handling requirements, and additional safety concerns [12]. Consequently, AI is effective at enumerating and ranking possible routes, but experienced chemists remain necessary to evaluate practicality and respond to unexpected experimental outcomes.

## 2.4. Preclinical and Clinical Research

The preclinical and clinical stages are especially costly because biological variability becomes more visible and failure can occur despite encouraging laboratory results. AI can analyze pathological images, electronic health records (EHRs), genomic profiles, and real-world data to support biomarker discovery, patient selection, trial design, and safety monitoring. PathAI applies deep learning to pathology images and has been studied for diagnostic assessment and treatment-response analysis[13]. Tempus integrates multimodal data for disease modeling and personalized medicine. Flatiron Health structures oncology EHR data to produce real-world evidence (RWE)[14]. NVIDIA Clara provides computational tools for medical imaging and genomics, including research on treatment-response prediction from standard pathology slides.

### 2.4.1. Applications

1) Biomarker Discovery: AI identifies molecular or imaging features associated with treatment response and may help distinguish patient subgroups.

2) Patient Stratification & Recruitment: EHR systems can locate patients who meet eligibility criteria, reducing manual screening and shortening recruitment periods.

3) Clinical Trial Simulation & Optimization: Digital twins and related models can simulate parts of trial progression and inform dose selection, endpoint definition, or enrollment strategy. Such simulations support planning but do not replace controlled clinical evidence.

4) Safety Monitoring: NLP can examine adverse-event reports and clinical notes to identify signals that deserve earlier investigation.

### 2.4.2. Limitations

Clinical AI faces major hurdles. Clinical data are fragmented, unstructured, and collected under different standards. Privacy regulations, including GDPR and HIPAA, restrict the use and transfer of sensitive patient information. Models trained on non-representative datasets may reproduce or intensify health disparities[15]. In addition, regulatory pathways for AI and machine-learning software as a medical device continue to develop. A model that performs well in one hospital may lose accuracy in another because of differences in patient populations, instruments, coding practices, or treatment patterns.

# 3. Overarching Challenges

1)Data silos and quality: standardized, well-annotated biomedical data remains limited, and organizations do not always have mechanisms for responsible sharing.

2)Black-box behavior: complex models can be difficult for domain experts to inspect, which weakens confidence in their recommendations.

3)Lack of standardization: different datasets and evaluation protocols make model comparisons and reproduction difficult.

4)Talent gap: effective deployment requires people who understand both computational methods and life-science workflows.

5)Computational cost: large models require substantial hardware and engineering support, which can limit access.

6)Regulatory uncertainty: developers need clearer expectations for validation, monitoring, documentation, and accountability.

# 4. Future Directions

Several developments are likely to shape the next stage of AI-enabled drug discovery.

1)Multimodal and foundation models: models may combine genomic, structural, imaging, and textual information so that evidence can be considered together rather than in isolated streams.

2)Explainable AI: explanations linked to molecular features, biological pathways, or uncertainty estimates could make model outputs easier to evaluate.

3)Automation and closed-loop systems: robotic synthesis and testing can connect design, making, measurement, and analysis in repeated cycles.

4)Simulation and digital twins: multiscale biological simulations may help compare candidate effects before some physical experiments are performed.

5)Privacy-preserving techniques: federated learning and related methods may allow institutions to collaborate without placing all sensitive data in one repository.

Regulatory science: regulators, industry, and academia should establish practical standards for evidence, model updates, performance monitoring, and post-deployment oversight.

# 5. Conclusion

**Table 1.** A conceptual summary derived from the reviewed literature

| Pipeline stage | Typical AI contribution | Human validation required | Main risk |
|---|---|---|---|
| Target research | Structure prediction and evidence integration | Biological assays and mechanism studies | Static or incomplete representations |
| Molecular design | Screening, generation, and property ranking | Synthesis and pharmacology testing | Poor novelty or ADMET reliability |
| Synthesis planning | Retrosynthesis and route comparison | Lab-scale feasibility and scale-up review | Positive-reporting bias in reaction data |
| Clinical research | Biomarkers, recruitment, and safety signals | Prospective clinical validation | Bias, privacy, and interoperability |

Artificial intelligence has initiated a significant change in pharmaceutical research. Its influence now extends from protein structure prediction and target discovery to molecular generation, synthesis planning, digital pathology, and clinical pipeline optimization. Nevertheless, AI is not an all-purpose solution. Data quality, interpretability, experimental validation, privacy, and regulation remain decisive constraints. The development of AI-enabled discovery is therefore unlikely to eliminate human expertise. Instead, it will strengthen collaborative workflows in which computational

systems prioritize hypotheses and researchers test, explain, and refine them. To realize this potential, the field must promote interdisciplinary training, fair data-sharing platforms, rigorous validation, and regulatory policies that can adapt to changing technologies. When computational predictions remain anchored in chemical and biological reality, AI may help deliver safer and more effective therapeutics with greater efficiency.

The Table 1 is a conceptual summary derived from the reviewed literature.

# References


[1] Jumper, J., Evans, R., Pritzel, A., et al. (2021). Highly accurate protein structure prediction with AlphaFold. *Nature*, *596*, 583–589. https://doi.org/10.1038/s41586-021-03819-2.

[2] Tunyasuvunakool, K., Adler, J., Wu, Z., et al. (2021). Highly accurate protein structure prediction for the human proteome. *Nature*, *596*, 590–596. https://doi.org/10.1038/s41586-021-03828-1.

[3] Baek, M., et al. (2021). Accurate prediction of protein structures and interactions using a three-track neural network. *Science*, *373*(6557), 871–876. https://doi.org/10. 1126/science. abj8754.

[4] Lin, Z., et al. (2023). Evolutionary-scale prediction of atomic-level protein structure with a language model. *Science*, *379*(6634), 1123–1130. https://doi.org/10. 1126/science. Ade 2574.

[5] Mead, R. J., Shan, N., Reiser, H. J., Marshall, F., & Shaw, P. J. (2023). Amyotrophic lateral sclerosis: a neurodegenerative disorder poised for successful therapeutic translation. *Nature Reviews Drug Discovery*, *22*(3), 185–212. https://doi.org/10. 1038/ s41573-022-00612-2.

[6] Jones, D. T., & Thornton, J. M. (2022). The impact of AlphaFold2 one year on. *Nature Methods*, *19*(1), 15–20. https://doi.org/10.1038/s41592-021-01365-3.

[7] Schrödinger. (2021). Morphic Therapeutic leverages digital chemistry strategy to design a novel small molecule inhibitor of α4β7 integrin.

[8] Chen, J., Bolhuis, D. L., Laggner, C., et al. (2023). AtomNet-Aided OTUD7B Inhibitor Discovery and Validation. *Cancers*, *15*(2), 517. https://doi.org/10.3390/cancers15020517.

[9] Niazi, S. K. (2025). Artificial Intelligence in Small-Molecule Drug Discovery: A Critical Review of Methods, Applications, and Real-World Outcomes. *Pharmaceuticals*, *18*(9), 1271. https://doi.org/10.3390/ph18091271.

[10] Bender, A., & Cortés-Ciriano, I. (2021). Artificial intelligence in drug discovery: what is realistic, what are illusions? Part 1. *Drug Discovery Today*, *26*(2), 511–524. https://doi.org/10. 1016/j. drudis.2020.12.009.

[11] Cardinale, A., Castrogiovanni, A., Gaudin, T., et al. (2023). Fuelling the Digital Chemistry Revolution with Language Models. *Chimia*, *77*(7–8), 484–488. https://doi.org/ 10.2533/ chimia. 2023.484.

[12] Coley, C. W., Thomas, D. A., Lummiss, J. A. M., et al. (2019). A robotic platform for flow synthesis of organic compounds informed by AI planning. *Science*, *365*(6453), eaax1566. https:// doi.org/10.1126/science. aax1566.

[13] Bera, K., Schalper, K. A., Rimm, D. L., Velcheti, V., & Madabhushi, A. (2019). Artificial intelligence in digital pathology - new tools for diagnosis and precision oncology. *Nature Reviews Clinical Oncology*, *16*(11), 703–715. https:// doi. org/10.1038/s41571-019-0252-y.

[14] Purpura, C. A., Garry, E. M., Honig, N., Case, A., & Rassen, J. A. (2022). The Role of Real-World Evidence in FDA-Approved New Drug and Biologics License Applications. *Clinical Pharmacology & Therapeutics*, *111*(1), 135–144. https:// doi.org/10.1002/cpt.2474.

[15] Obermeyer, Z., Powers, B., Vogeli, C., & Mullainathan, S. (2019). Dissecting racial bias in an algorithm used to manage the health of populations. *Science*, *366*(6464), 447–453. https://doi.org/10.1126/science.aax2342.