

AI-Driven Drug Discovery Approaches for Pharmaceutical Innovation

Honglin Li

Professor, East China University of Science and Technology, China

Abstract

Artificial intelligence has emerged as an important technological approach for transforming drug discovery and pharmaceutical innovation. Conventional drug discovery processes can require substantial time, financial resources, experimental screening, and repeated optimization. AI-driven approaches can support these activities by applying machine learning, deep learning, natural language processing, molecular modeling, virtual screening, and predictive analytics to large and complex biomedical datasets. This study examines AI-driven drug discovery approaches and their contribution to pharmaceutical innovation through a conceptual and literature-based research approach. The study focuses on target identification, molecular property prediction, virtual screening, de novo molecular design, drug repurposing, toxicity prediction, pharmacokinetic assessment, and candidate optimization.

AI systems can identify relationships within biological and chemical datasets that may support the prioritization of drug candidates and reduce the number of compounds requiring extensive experimental investigation. Machine-learning models can assist in predicting molecular properties, while deep-learning approaches can support structure-based analysis and molecular generation. Natural language processing can facilitate the extraction of relevant information from scientific literature and biomedical databases. However, AI-driven drug discovery also faces challenges related to data quality, limited datasets, model interpretability, biological complexity, computational requirements, validation, regulatory expectations, intellectual property, and integration with laboratory experimentation. AI should therefore be considered a complementary technology rather than an independent replacement for pharmaceutical research and experimental validation.

The study concludes that AI can strengthen pharmaceutical innovation by accelerating information analysis, improving candidate prioritization, supporting molecular design, and enabling more data-driven research workflows. Future research should emphasize explainable AI, multimodal biomedical data integration, generative models, physics-informed approaches, secure data sharing, laboratory automation, and stronger integration between computational prediction and experimental validation.

Keywords: Artificial Intelligence, Drug Discovery, Pharmaceutical Innovation, Machine Learning, Deep Learning, Virtual Screening, Molecular Design, Drug Repurposing, Predictive Analytics, Computational Drug Discovery.

1. Introduction

Drug discovery is a complex and multidisciplinary process involving biological target identification, compound screening, molecular optimization, pharmacological evaluation, toxicity assessment, and experimental validation. Traditional approaches frequently require extensive laboratory experimentation and sequential decision-making across multiple development stages. The increasing availability of chemical, genomic, proteomic, clinical, and biomedical datasets has created opportunities for computational approaches that can analyze information more rapidly. Artificial intelligence has consequently emerged as a significant technology for supporting pharmaceutical research and innovation.

AI-driven drug discovery involves the application of machine learning, deep learning, natural language processing, molecular modeling, and other computational techniques to different stages of pharmaceutical research. These methods can analyze relationships between molecular structures, biological targets, pharmacological properties, and experimental outcomes. Machine-learning models can assist with prediction tasks, while deep-learning approaches can identify complex patterns within high-dimensional datasets. Such capabilities can support researchers in prioritizing promising candidates before extensive experimental testing.

Virtual screening represents another important application of AI in drug discovery. Instead of experimentally testing every available compound, computational models can rank molecules according to predicted interactions, properties, or biological activity. AI can also support de novo molecular design by generating or identifying candidate structures with desirable characteristics. Drug repurposing is another area in which computational models can analyze existing medicines and biological information to identify potential alternative therapeutic applications.

Despite its potential, AI-driven drug discovery involves important limitations. Pharmaceutical datasets can contain missing information, experimental variation, biases, inconsistent measurements, and limited representation of biological conditions. Models trained on one dataset may not always generalize effectively to different chemical or biological environments. Interpretability is also important because researchers need to understand why a computational model prioritizes a particular compound or target. Experimental validation remains essential for determining whether computational predictions translate into reliable biological outcomes.

AI therefore should be integrated into pharmaceutical research as a complementary capability that supports rather than replaces experimental science. Effective implementation requires high-quality data, appropriate algorithms, domain expertise, computational infrastructure, laboratory validation, and responsible governance. This study examines AI-driven drug discovery approaches for pharmaceutical innovation, focusing on major technological applications, potential contributions, implementation challenges, and future research directions.

2. Objectives of the Study

2.1 Primary Objective

The primary objective of this study is to examine AI-driven approaches used in drug discovery and analyze how artificial intelligence, machine learning, deep learning, virtual screening, molecular

modeling, natural language processing, and predictive analytics can support target identification, molecular design, candidate prioritization, drug repurposing, and pharmaceutical innovation.

The study also aims to identify the technological and scientific requirements necessary for reliable AI integration within drug discovery. Particular attention is given to data quality, model validation, interpretability, computational resources, biological complexity, experimental verification, regulatory considerations, and collaboration between computational scientists, pharmaceutical researchers, and laboratory professionals.

2.2 Examination of AI Technologies in Drug Discovery

The first objective is to examine the major AI technologies supporting drug discovery, including machine learning, deep learning, generative models, natural language processing, molecular prediction, and computational screening. The study considers how these technologies can process chemical and biomedical information.

The objective also examines the application of AI across target identification, molecular property prediction, virtual screening, molecular generation, toxicity prediction, pharmacokinetic assessment, and candidate optimization. Emphasis is placed on the ability of AI to prioritize research activities while recognizing that experimental validation remains necessary to establish biological and pharmaceutical reliability.

2.3 Assessment of Pharmaceutical Innovation Applications

The second objective is to assess how AI-driven approaches can contribute to pharmaceutical innovation by improving candidate selection, accelerating computational analysis, supporting drug repurposing, and assisting molecular optimization. The study considers AI as a tool for increasing the efficiency of research workflows.

The objective further evaluates challenges affecting AI implementation, including data limitations, model interpretability, biological complexity, generalization, computational requirements, regulatory expectations, and experimental validation. Understanding these challenges is necessary for developing AI-supported drug discovery workflows that are scientifically reliable, reproducible, transparent, and appropriately integrated with laboratory research.

3. Review of Literature

3.1 Artificial Intelligence in Drug Discovery

Artificial intelligence has increasingly been investigated as a computational approach for improving different stages of drug discovery. Vamathevan et al. (2019) reviewed machine learning applications across drug discovery and development and identified opportunities in target validation, compound screening, biomarker identification, and clinical development. AI can process large datasets and identify relationships that may be difficult to detect through conventional analysis.

Ekins et al. (2019) examined the role of machine learning and artificial intelligence in drug discovery and emphasized the importance of appropriate data, validation, and domain expertise. Computational models can assist researchers in prioritizing candidates and predicting properties, but

their usefulness depends on the quality and relevance of training information. Integration with experimental research remains necessary.

3.2 Virtual Screening and Molecular Prediction

Virtual screening is an important computational strategy for prioritizing compounds according to predicted biological activity or molecular interactions. Gawehn, Hiss, and Schneider (2016) discussed the role of deep learning in drug discovery and highlighted its potential for molecular property prediction and compound prioritization. Deep-learning methods can process complex molecular representations and learn relationships from chemical datasets.

Machine-learning models can also support prediction of properties related to absorption, distribution, metabolism, excretion, and toxicity. These predictions may help researchers identify potential problems earlier in the discovery process. However, predictive accuracy depends on the quality, diversity, and representativeness of available datasets. Computational predictions should therefore be evaluated against appropriate experimental evidence.

3.3 Generative AI and Pharmaceutical Innovation

Generative approaches have created new opportunities for designing molecules with desired characteristics. Gómez-Bombarelli et al. (2018) demonstrated the application of deep generative models to molecular design, illustrating how machine learning can be used to explore chemical spaces computationally. Generative approaches can potentially increase the number of candidate structures considered during early drug discovery.

Computational molecular design can be combined with optimization objectives related to activity, selectivity, toxicity, and physicochemical properties. However, generated molecules require careful assessment because computational novelty does not automatically indicate biological effectiveness. Experimental synthesis, testing, and validation remain necessary. AI-driven generation should therefore operate within multidisciplinary pharmaceutical workflows that combine computational modeling with chemical and biological expertise.

4. Research Methodology

4.1 Research Design

This study adopts a conceptual and literature-based research design to examine AI-driven approaches for drug discovery and pharmaceutical innovation. The research focuses on machine learning, deep learning, virtual screening, molecular modeling, generative approaches, natural language processing, drug repurposing, and predictive analytics.

The study is descriptive and analytical rather than experimental. Existing academic literature is examined to identify major AI applications, technological capabilities, pharmaceutical benefits, implementation requirements, and challenges. The approach provides a structured conceptual assessment without presenting fabricated laboratory measurements, clinical trial outcomes, experimental compound results, or unverified pharmaceutical performance claims.

4.2 Literature Collection and Organization

The literature was organized around major AI-driven drug discovery themes, including target identification, virtual screening, molecular property prediction, molecular generation, drug repurposing, toxicity prediction, pharmacokinetic analysis, and candidate optimization. Academic studies were reviewed to identify recurring findings and technological applications.

The literature was categorized according to AI method, drug discovery stage, data type, analytical function, and potential pharmaceutical contribution. Machine-learning and deep-learning approaches were considered in relation to prediction and classification, while generative models were examined for molecular design. Natural language processing was considered for scientific information extraction and knowledge discovery.

4.3 Analytical Framework

The analytical framework evaluates AI-driven drug discovery across six dimensions: biomedical data acquisition, computational representation, predictive modeling, candidate prioritization, experimental validation, and pharmaceutical decision-making. Biomedical and chemical data provide the foundation for computational analysis.

Molecular representations and biological features are processed through machine-learning or deep-learning models. Predictive outputs can support candidate prioritization and molecular optimization. Experimental validation is considered a critical stage for evaluating computational predictions. Pharmaceutical decision-making integrates computational and experimental evidence to determine whether candidates warrant further investigation. This framework emphasizes AI as a supportive component of multidisciplinary drug discovery.

4.4 Data Analysis Approach

A thematic analytical approach was applied to the selected literature to identify recurring patterns in AI-driven drug discovery. The analysis focused on technologies, discovery stages, pharmaceutical applications, potential contributions, and implementation limitations.

The identified themes were organized into computational, chemical, biological, experimental, regulatory, and organizational categories. Comparative interpretation was used to examine how AI approaches can support different stages of drug discovery. Particular attention was given to predictive accuracy, data quality, model interpretability, generalization, experimental validation, computational efficiency, and integration with established pharmaceutical research workflows.

5. Analysis

5.1 AI Technologies Supporting Drug Discovery

AI-driven drug discovery involves multiple computational technologies that address different research requirements. Machine learning can support classification and prediction, while deep learning can identify complex patterns in chemical and biological datasets. Natural language processing can extract information from scientific literature, and virtual screening can prioritize candidate compounds.

Generative models can support molecular design, while predictive analytics can assist toxicity and pharmacokinetic assessment. Table 1 summarizes major AI technologies and their principal drug discovery functions.

Table 1: AI Technologies and Their Drug Discovery Applications

AI Technology	Primary Function	Drug Discovery Application
Machine Learning	Prediction and classification	Candidate prioritization
Deep Learning	Complex pattern recognition	Molecular property prediction
Generative Models	Molecular generation	De novo drug design
Natural Language Processing	Information extraction	Scientific knowledge discovery
Virtual Screening Models	Compound ranking	Candidate screening
Molecular Modeling	Interaction prediction	Target analysis
Predictive Analytics	Risk and property estimation	Candidate evaluation
Graph Neural Networks	Molecular representation	Structure prediction
Reinforcement Learning	Iterative optimization	Molecular optimization
Knowledge Graphs	Relationship modeling	Drug-target discovery

Interpretation: Table 1 demonstrates that AI technologies can support complementary stages of drug discovery. Machine learning and deep learning can assist prediction and classification, while generative models can explore potential molecular structures. Natural language processing can help researchers organize information from scientific literature, and virtual screening can prioritize compounds for further investigation.

Molecular modeling and graph-based approaches can represent relationships between chemical structures and biological targets. Reinforcement learning can support iterative optimization, while knowledge graphs can connect drug, target, disease, and biological information. These technologies can collectively strengthen computational pharmaceutical research when appropriately validated.

5.2 Drug Discovery Stages and AI Contributions

AI can be integrated across several stages of the drug discovery process. Target identification can use computational analysis of biological information, while virtual screening can prioritize compounds. Molecular design approaches can explore chemical structures, and predictive models can estimate properties associated with efficacy, toxicity, and pharmacokinetics. Drug repurposing can analyze relationships among existing compounds, targets, and diseases. Table 2 presents major discovery stages and potential AI contributions.

Table 2: AI Applications Across Drug Discovery Stages

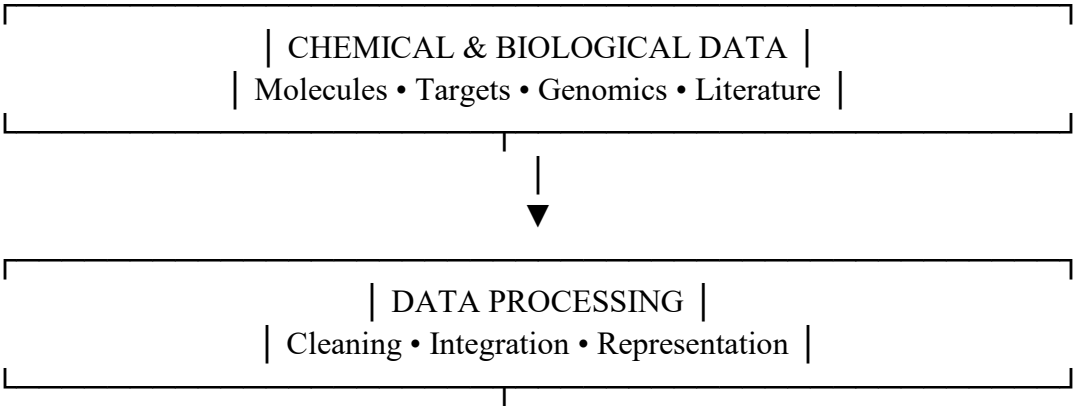
Drug Discovery Stage	AI Application	Potential Contribution
Target Identification	Biological data analysis	Target prioritization
Target Validation	Predictive modeling	Target assessment
Virtual Screening	Machine-learning ranking	Compound prioritization
Molecular Design	Generative AI	Candidate generation
Lead Optimization	Predictive modeling	Property improvement
Toxicity Prediction	AI classification	Early risk assessment
Pharmacokinetic Prediction	Machine learning	ADME assessment
Drug Repurposing	Knowledge analysis	New therapeutic hypotheses
Candidate Selection	Predictive analytics	Research prioritization
Experimental Planning	Data-driven modeling	Efficient testing strategies

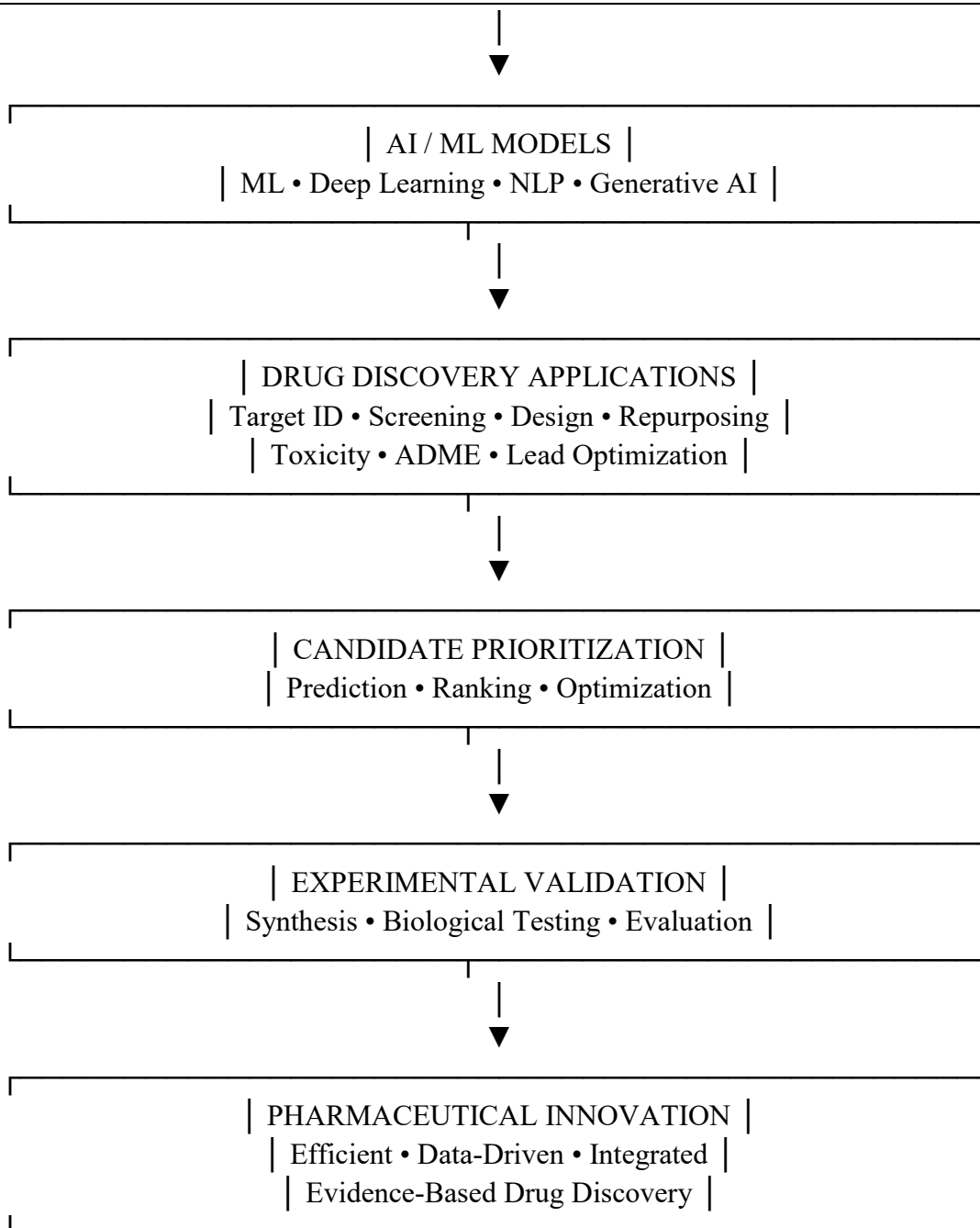
Interpretation: Table 2 illustrates that AI can contribute throughout multiple stages of drug discovery rather than being limited to compound screening. Biological data analysis can assist target identification, while predictive models can support target validation. Virtual screening can prioritize compounds, and generative AI can explore molecular structures.

Lead optimization can use predictive models to evaluate desirable properties, while toxicity and pharmacokinetic models can identify potential concerns earlier. Knowledge-based approaches can support drug repurposing by connecting existing compounds with potential therapeutic targets. These applications can help researchers prioritize computational and experimental activities more efficiently.

5.3 Conceptual Framework for AI-Driven Drug Discovery

Figure 1: AI-Driven Pharmaceutical Innovation Framework





Interpretation: The framework shows how AI-driven drug discovery begins with chemical, biological, genomic, and scientific information. Data processing prepares information for computational modeling, after which machine learning, deep learning, natural language processing, and generative AI can support different discovery activities. Computational outputs help prioritize and optimize candidate molecules before experimental testing. Laboratory synthesis and biological validation remain essential for determining whether computational predictions have practical value.

The final integration of computational and experimental evidence can support more efficient, data-driven, and evidence-based pharmaceutical innovation while maintaining scientific and regulatory requirements.

6. Discussion

AI-driven drug discovery can strengthen pharmaceutical research by enabling computational analysis of large and complex chemical and biological datasets. Machine-learning and deep-learning methods can support prediction, classification, virtual screening, and molecular property assessment. Generative approaches can expand the range of molecular structures considered during early discovery, while natural language processing can assist scientific knowledge extraction. These capabilities can help researchers prioritize promising candidates and reduce unnecessary computational or experimental effort. However, AI outputs should be considered research-support information rather than definitive evidence of pharmaceutical effectiveness.

The integration of AI across multiple drug discovery stages creates opportunities for more connected research workflows. Target identification, compound screening, molecular design, toxicity prediction, pharmacokinetic assessment, and lead optimization can all benefit from computational modeling. AI can also support drug repurposing by identifying relationships between existing compounds, biological targets, and diseases. Nevertheless, biological systems are highly complex, and computational predictions may not fully represent real biological responses. Laboratory experiments, pharmacological studies, and appropriate validation remain essential for confirming AI-generated hypotheses.

The future of AI-driven pharmaceutical innovation depends on responsible integration between computational technologies and experimental science. High-quality datasets, transparent models, reproducible workflows, domain expertise, and appropriate validation are necessary for trustworthy AI applications. Researchers must also address issues involving interpretability, data ownership, regulatory expectations, computational costs, and intellectual property. A collaborative model involving pharmaceutical scientists, computational researchers, biologists, chemists, clinicians, and regulatory experts can help ensure that AI technologies are used effectively while maintaining scientific rigor and appropriate standards of pharmaceutical development.

7. Challenges

7.1 Data Quality and Availability

AI-driven drug discovery depends heavily on the quality, quantity, diversity, and relevance of available chemical and biological datasets. Pharmaceutical datasets may contain missing values, inconsistent measurements, experimental variability, limited sample sizes, or biased representations of chemical structures and biological conditions. These limitations can influence model performance and reduce the reliability of computational predictions.

Improving data quality requires standardized data collection, appropriate preprocessing, careful annotation, and transparent documentation of experimental conditions. Researchers should also evaluate whether datasets adequately represent the intended discovery problem. External validation using independent datasets can help determine whether models generalize beyond training information.

Reliable data infrastructure is therefore essential for developing robust and reproducible AI-supported drug discovery systems.

7.2 Model Interpretability and Validation

Many advanced AI models can provide highly complex predictions that are difficult to interpret. In pharmaceutical research, understanding why a model prioritizes a particular molecule or predicts a specific biological property can be important for scientific evaluation. Limited interpretability may reduce confidence in computational findings and make it difficult to identify model weaknesses.

Validation should therefore include appropriate performance measures, independent datasets, comparison with established methods, and experimental verification. Explainable AI techniques can provide additional information about model predictions and influential features. However, explanations should not be treated as proof of biological validity. Combining computational transparency with laboratory evidence is necessary to establish confidence in AI-driven drug discovery.

7.3 Biological Complexity and Generalization

Drug discovery involves complex biological systems in which molecular interactions may vary according to biological context, disease state, cellular environment, and experimental conditions. An AI model trained using a particular dataset may perform well within that dataset but show weaker performance when applied to new compounds, targets, or biological environments.

Generalization therefore represents an important challenge. Models should be tested using appropriate external datasets and relevant experimental settings. Researchers should carefully define the intended application domain before interpreting model outputs. Multimodal datasets combining chemical, genomic, proteomic, structural, and biological information may improve modeling capabilities, but they also introduce additional complexity that requires careful methodological management.

7.4 Experimental, Regulatory, and Computational Requirements

AI can accelerate computational analysis, but pharmaceutical development still requires synthesis, laboratory testing, pharmacological evaluation, safety assessment, and other forms of experimental evidence. Computational predictions cannot independently establish therapeutic effectiveness or safety. Integrating AI into established research pipelines can also require specialized computational infrastructure, skilled personnel, software, and data-management systems.

Regulatory expectations create additional considerations when AI contributes to pharmaceutical research and development. Models should be appropriately documented, validated, and reproducible. Researchers must also consider intellectual property and data ownership when using proprietary chemical or biomedical datasets. Effective implementation therefore requires coordinated computational, experimental, organizational, and regulatory planning.

8. Future Research

Future research should investigate multimodal AI systems capable of integrating chemical structures, biological targets, genomic information, proteomic data, scientific literature, and experimental results.

Such models may provide more comprehensive representations of drug discovery problems. Research should also examine explainable and uncertainty-aware AI methods that help researchers understand prediction confidence and model limitations.

Further research should explore generative AI, graph neural networks, reinforcement learning, physics-informed machine learning, and AI-assisted laboratory automation. Integration between computational prediction and automated experimental platforms could create iterative discovery workflows in which computational models propose candidates and laboratory systems generate evidence for subsequent model improvement.

Future empirical studies should evaluate AI-driven drug discovery systems using independent datasets and experimental validation. Research should assess predictive accuracy, generalization, reproducibility, computational efficiency, interpretability, and practical pharmaceutical value. Comparative studies between AI-assisted and conventional discovery workflows may provide stronger evidence regarding where AI delivers meaningful improvements in research efficiency and candidate prioritization.

9. Conclusion

AI-driven drug discovery represents an important technological development in pharmaceutical innovation. Machine learning, deep learning, virtual screening, molecular modeling, generative approaches, and natural language processing can support multiple stages of drug discovery. These technologies can process large datasets, identify patterns, prioritize compounds, support molecular design, and assist researchers in evaluating potential candidates before extensive experimental investigation.

The analysis demonstrates that AI can contribute to target identification, virtual screening, molecular generation, lead optimization, toxicity prediction, pharmacokinetic assessment, drug repurposing, and candidate selection. Different AI methods provide complementary capabilities, allowing pharmaceutical researchers to develop more integrated computational workflows. These applications can support research prioritization and potentially improve the efficiency of early-stage drug discovery.

Nevertheless, AI-driven pharmaceutical research faces significant challenges involving data quality, model interpretability, biological complexity, generalization, computational requirements, experimental validation, regulatory considerations, and intellectual property. AI predictions should not be considered substitutes for laboratory evidence. Appropriate experimental testing remains necessary to determine biological activity, safety, pharmacological behavior, and therapeutic relevance.

Overall, AI can become a valuable component of pharmaceutical innovation when computational approaches are integrated with experimental science and domain expertise. Future development should emphasize reliable datasets, explainable models, multimodal learning, generative approaches, experimental automation, transparent validation, and responsible governance. A balanced integration of artificial intelligence, pharmaceutical science, laboratory experimentation, and regulatory expertise can support more efficient, data-driven, and scientifically rigorous drug discovery.

References

1. Mak, K. K., & Pichika, M. R. (2019). Artificial intelligence in drug development: Present status and future prospects. *Drug Discovery Today*, 24(3), 773–780. <https://doi.org/10.1016/j.drudis.2018.11.014>
2. Zhang, L., Tan, J., Han, D., & Zhu, H. (2017). From machine learning to deep learning: Progress in machine intelligence for rational drug discovery. *Drug Discovery Today*, 22(11), 1680–1685. <https://doi.org/10.1016/j.drudis.2017.08.010>
3. Kim, H., Kim, E., Lee, I., Bae, B., Park, M., & Nam, H. (2020). Artificial intelligence in drug discovery: A comprehensive review of data-driven and machine learning approaches. *Biotechnology and Bioprocess Engineering*, 25(6), 895–930. <https://doi.org/10.1007/s12257-020-0049-y>
4. Sellwood, M. A., Ahmed, M., Segler, M. H. S., & Brown, N. (2018). Artificial intelligence in drug discovery. *Future Medicinal Chemistry*, 10(17), 2025–2028. <https://doi.org/10.4155/fmc-2018-0212>
5. Lavecchia, A. (2019). Deep learning in drug discovery: Opportunities, challenges and future prospects. *Drug Discovery Today*, 24(10), 2017–2032. <https://doi.org/10.1016/j.drudis.2019.07.006>
6. Hessler, G., & Baringhaus, K.-H. (2018). Artificial intelligence in drug design. *Molecules*, 23(10), 2520. <https://doi.org/10.3390/molecules23102520>
7. Jing, Y., Bian, Y., Hu, Z., Wang, L., & Xie, X.-Q. (2018). Deep learning for drug design: An artificial intelligence paradigm for drug discovery in the big data era. *The AAPS Journal*, 20, 58. <https://doi.org/10.1208/s12248-018-0210-0>
8. Popova, M., Isayev, O., & Tropsha, A. (2018). Deep reinforcement learning for de novo drug design. *Science Advances*, 4(7), eaap7885. <https://doi.org/10.1126/sciadv.aap7885>
9. Harel, S., & Radinsky, K. (2018). Prototype-based compound discovery using deep generative models. *Molecular Pharmaceutics*, 15(10), 4406–4416. <https://doi.org/10.1021/acs.molpharmaceut.8b00474>
10. Jørgensen, P. B., Schmidt, M. N., & Winther, O. (2018). Deep generative models for molecular science. *Molecular Informatics*, 37(1–2). <https://doi.org/10.1002/minf.201700133>
11. Elton, D. C., Boukouvalas, Z., Fuge, M. D., & Chung, P. W. (2019). Deep learning for molecular design—A review of the state of the art. *Molecular Systems Design & Engineering*, 4, 828–849. <https://doi.org/10.1039/C9ME00039A>
12. Schwaller, P., Laino, T., Gaudin, T., Bolgar, P., Hunter, C. A., Bekas, C., & Lee, A. A. (2019). Molecular Transformer: A model for uncertainty-calibrated chemical reaction prediction. *ACS Central Science*, 5(9), 1572–1583. <https://doi.org/10.1021/acscentsci.9b00576>
13. Lee, A. A., Yang, Q., Sresht, V., Bolgar, P., Hou, X., Klug-McLeod, J. L., & Butler, C. R. (2019). Molecular Transformer unifies reaction prediction and retrosynthesis across pharma chemical space. *Chemical Communications*, 55, 12152–12155. <https://doi.org/10.1039/C9CC05122H>
14. Brown, N., Fiscato, M., Segler, M. H. S., & Vaucher, A. C. (2019). GuacaMol: Benchmarking models for de novo molecular design. *Journal of Chemical Information and Modeling*, 59(3), 1096–1108. <https://doi.org/10.1021/acs.jcim.8b00839>
15. Lim, J., Ryu, S., Kim, J. W., & Kim, W. Y. (2018). Molecular generative model based on conditional variational autoencoder for de novo molecular design. *Journal of Cheminformatics*, 10, 31. <https://doi.org/10.1186/s13321-018-0286-7>

16. Gómez-Bombarelli, R., Wei, J. N., Duvenaud, D., Hernández-Lobato, J. M., Sánchez-Lengeling, B., Sheberla, D., Aguilera-Iparraguirre, J., Hirzel, T. D., Adams, R. P., & Aspuru-Guzik, A. (2018). Automatic chemical design using a data-driven continuous representation of molecules. *ACS Central Science*, 4(2), 268–276. <https://doi.org/10.1021/acscentsci.7b00572>
17. Korotcov, A., Tkachenko, V., Russo, D. P., & Ekins, S. (2017). Comparison of deep learning with multiple machine learning methods and metrics using diverse drug discovery data sets. *Molecular Pharmaceutics*, 14(12), 4462–4475. <https://doi.org/10.1021/acs.molpharmaceut.7b00578>
18. Cai, J., Li, C., Liu, Z., Du, J., Ye, J., Gu, Q., & Xu, J. (2017). Predicting DPP-IV inhibitors with machine learning approaches. *Journal of Computer-Aided Molecular Design*, 31(4), 393–402. <https://doi.org/10.1007/s10822-017-0009-6>
19. Pandey, M., Fernandez, M., Gentile, F., Isayev, O., Tropsha, A., Stern, A. C., & Cherkasov, A. (2022). The transformational role of GPU computing and deep learning in drug discovery. *Nature Machine Intelligence*, 4, 211–221. <https://doi.org/10.1038/s42256-022-00463-x>
20. You, Y., Lai, X., Pan, Y., Zheng, H., Vera, J., Liu, S., Deng, S., et al. (2022). Artificial intelligence in cancer target identification and drug discovery. *Signal Transduction and Targeted Therapy*, 7, 156. <https://doi.org/10.1038/s41392-022-00994-0>