International Journal for Multidisciplinary Research (IJFMR)
E-ISSN: 2582-2160 ● Website: www.ijfmr.com ● Email: editor@ijfmr.com

Next-Generation Artificial Intelligence for Drug Repurposing: Integrating Machine Learning, Deep Learning, Multi-Omics, Knowledge Graphs, and Clinical Translation: A Review

Sushma N[1], Rachana SN[2], Aruna MV[3], Sunil kumar Busi[4], Gopinath E[5]

[1,2,3]Student, Shantha College of Pharmacy, Peresandra, Chickballapur-562104
[4]Professor, Department of Pharmaceutical Chemistry, Shantha College of Pharmacy, Peresandra, Chickballapur-562104
[5]Principal, Shantha College of Pharmacy, Peresandra, Chickballapur-562104

Abstract

Drug repurposing has emerged as a promising strategy to accelerate drug discovery by identifying new therapeutic indications for existing approved or investigational drugs, thereby reducing development time, cost, and clinical risk compared with traditional de novo drug development. The rapid expansion of biomedical big data, together with advances in artificial intelligence (AI) and machine learning (ML), has transformed computational drug repurposing into a data-driven and highly efficient discipline. Conventional machine learning algorithms, including Support Vector Machines, Random Forests, and gradient boosting methods, have demonstrated significant utility in predicting drug-target and drug-disease associations. More recently, deep learning architectures such as Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), Autoencoders, Transformer-based models, and Graph Neural Networks (GNNs) have enabled the integration of heterogeneous datasets, including chemical structures, transcriptomics, proteomics, metabolomics, pharmacogenomics, protein-protein interaction networks, and electronic health records, substantially improving prediction accuracy.

This review provides a comprehensive overview of AI-driven drug repurposing, covering computational strategies, publicly available biomedical databases, feature representation methods, machine learning and deep learning algorithms, and their applications in cancer, infectious diseases, neurological disorders, cardiovascular diseases, and rare diseases. Furthermore, recent advances in knowledge graphs, explainable artificial intelligence (XAI), federated learning, foundation models, and large language models (LLMs) are discussed as emerging technologies capable of improving prediction reliability, interpretability, and clinical applicability. Current challenges, including data heterogeneity, limited external validation, algorithmic bias, model interpretability, and regulatory barriers, are critically evaluated. Finally, future perspectives focusing on multimodal multi-omics integration, digital twins, real-world evidence, and precision medicine are presented.

Keywords: Drug repurposing; Drug repositioning; Artificial intelligence; Machine learning; Deep learning; Graph neural networks; Multi-omics; Knowledge graphs; Explainable AI; Precision medicine.

1. Introduction

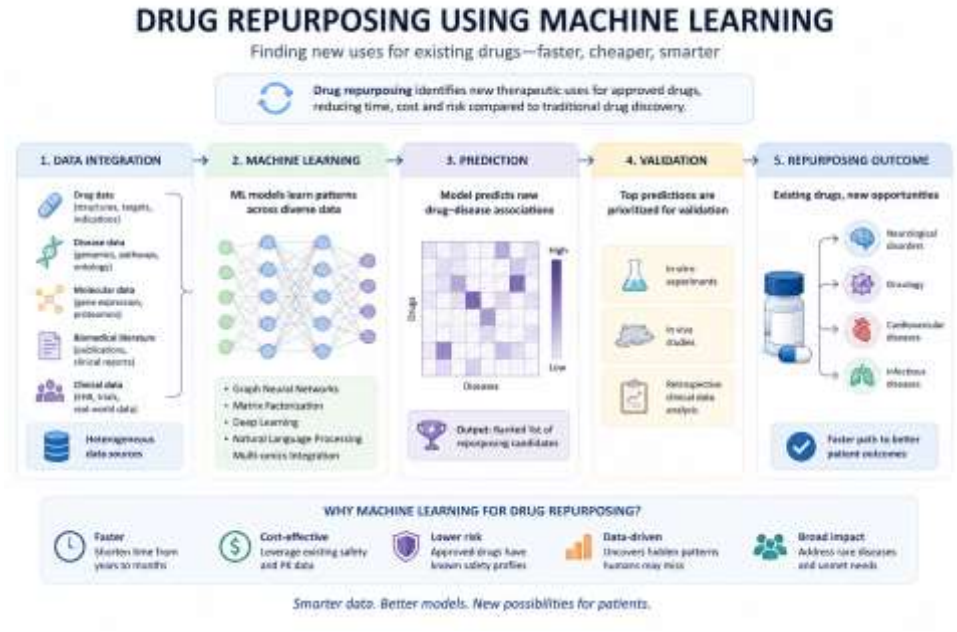
Drug discovery is a lengthy, costly, and high-risk process, typically requiring 10-15 years and billions of dollars to bring a new therapeutic agent from initial discovery to regulatory approval. Furthermore, only a small proportion of candidate molecules successfully complete clinical development because of inadequate efficacy, unacceptable toxicity, or unfavourable pharmacokinetic characteristics. These limitations have driven increasing interest in drug repurposing (drug repositioning), a strategy that identifies new therapeutic indications for existing approved or investigational drugs. Since repurposed drugs often have well-established pharmacokinetic profiles, safety data, and manufacturing processes, they can substantially reduce development time, cost, and clinical risk.

Recent advances in artificial intelligence (AI) and machine learning (ML) have transformed computational drug repurposing by enabling rapid analysis of large-scale biomedical datasets. Unlike conventional computational methods that rely on a single source of information, modern AI integrates diverse data modalities including chemical structures, drug-target interactions, transcriptomic signatures, proteomics, metabolomics, protein-protein interaction networks, electronic health records, and biomedical literature. These approaches improve the prediction of previously unrecognized drug-disease associations and accelerate therapeutic discovery.

Early machine learning methods primarily used supervised algorithms such as Support Vector Machines, Random Forests, and logistic regression to classify therapeutic indications based on chemical similarity or molecular descriptors. The emergence of deep learning introduced neural-network architectures capable of learning complex nonlinear relationships directly from high-dimensional biological data. Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), Autoencoders, and particularly Graph Neural Networks (GNNs) have demonstrated superior performance by modelling intricate interactions among drugs, targets, genes, pathways, and diseases.

The COVID-19 pandemic further highlighted the importance of AI-assisted drug repurposing. Within months of the emergence of SARS-CoV-2, computational platforms screened thousands of approved drugs using transcriptomic analysis, molecular docking, knowledge graphs, and deep learning, greatly accelerating candidate prioritization. Similar approaches are increasingly applied to oncology, neurodegenerative diseases, cardiovascular disorders, autoimmune diseases, rare diseases, and antimicrobial resistance.

Despite these advances, several challenges continue to hinder clinical translation. Many predictive models rely on incomplete or biased datasets, lack external validation, and often function as black boxes with limited interpretability. The integration of heterogeneous multi-omics datasets remains difficult, and standardized benchmark datasets for evaluating AI models are lacking. Emerging technologies including foundation models, large language models (LLMs), knowledge graphs, federated learning, digital twins, and explainable AI (XAI) offer promising opportunities to address these limitations and improve precision drug repurposing.



2. Fundamentals of Drug Repurposing

Drug repurposing (also known as drug repositioning) is the process of identifying new therapeutic indications for approved, discontinued, investigational, or shelved drugs beyond their original medical use. Unlike conventional drug discovery, which requires the identification of novel chemical entities followed by extensive preclinical and clinical evaluation, drug repurposing exploits existing pharmacological, toxicological, and clinical information to shorten the development timeline and reduce financial risk.

Historically, many successful repurposed drugs were discovered serendipitously. Modern drug repurposing, however, has become increasingly systematic through the integration of computational biology, bioinformatics, pharmacogenomics, and artificial intelligence. Large biomedical databases combined with machine learning algorithms enable the identification of hidden relationships among drugs, molecular targets, genes, pathways, and diseases, facilitating the prediction of novel therapeutic applications.

2.2 Types of Drug Repurposing

Drug repurposing strategies can be broadly classified into three categories.

A. On-target Repurposing

- Same molecular target
- Different disease indication
- Example: kinase inhibitors used across multiple cancer types.

B. Off-target Repurposing

- Drug interacts with a previously unknown molecular target
- New therapeutic indication arises from this interaction.
- Often identified through AI-based target prediction.

C. Disease Network-Based Repurposing

- Based on similarities among diseases, pathways, gene expression profiles, or protein interaction networks.

- Frequently uses systems biology and graph-based computational models.

2.3 Drug Repurposing Workflow

A typical AI-assisted drug repurposing pipeline consists of the following steps:

- Collection of biomedical data.
- Data pre-processing.
- Feature representation.
- AI/ML model development.
- Prediction of drug-disease associations.
- Biological validation.
- Clinical validation.

Table 1: Comparison Between Traditional Drug Discovery and Drug Repurposing

Parameter	Traditional Drug Discovery	Drug Repurposing
Development time	10-15 years	2-6 years
Development cost	Very high	Significantly lower
Safety profile	Unknown	Mostly established
Failure rate	High	Lower
Clinical phases	Phase I-III	Often begins at Phase II/III
Regulatory risk	High	Moderate
Success probability	Low	Higher

2.4 Advantages of Drug Repurposing

Drug repurposing provides several scientific and economic advantages.

- Reduced development cost.
- Shorter regulatory approval timeline.
- Existing pharmacokinetic and toxicological data.
- Lower probability of late-stage clinical failure.
- Rapid response during public health emergencies (e.g., COVID-19).
- Improved return on pharmaceutical investment.
- Accelerated translation into clinical practice.
- Increased opportunities for treating rare and neglected diseases.

2.5 Challenges and Limitations

Despite its advantages, drug repurposing faces several challenges.

Scientific Challenges

- Incomplete understanding of disease mechanisms.
- Limited availability of high-quality multi-omics datasets.
- Biological variability among patient populations.
- Drug resistance and disease heterogeneity.

Computational Challenges

- Data imbalance.
- Missing biological information.
- Heterogeneous data integration.

- Poor model interpretability.
- Overfitting in deep learning models.

Clinical Challenges

- Lack of prospective clinical validation.
- Limited randomized clinical trials.
- Regulatory uncertainty.
- Intellectual property and patent issues.

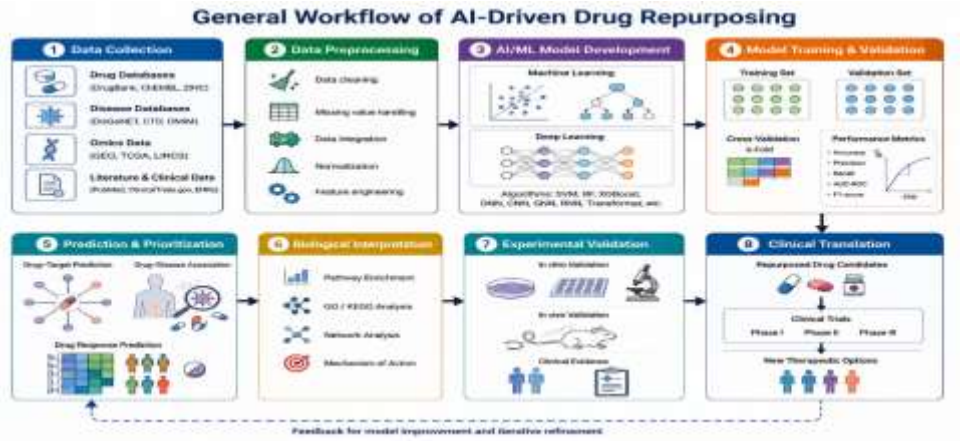
Table 1.2: Conventional vs AI-Driven Drug Repurposing

Feature	Conventional Repurposing	AI-Driven Repurposing
Discovery approach	Experimental	Computational + Experimental
Data sources	Literature	Multi-omics + clinical + chemical
Speed	Slow	Rapid
Cost	High	Lower
Scalability	Limited	High
Prediction accuracy	Moderate	High
Personalization	Limited	Precision medicine enabled

2.6 Emerging Trends in Drug Repurposing

Recent advances are transforming drug repurposing into a precision medicine platform. These include:

- Multi-omics integration.
- Knowledge graph-based drug discovery.
- Graph Neural Networks (GNNs).
- Transformer-based foundation models.
- Large Language Models (LLMs).
- Explainable Artificial Intelligence (XAI).
- Federated learning for privacy-preserving healthcare.
- Digital twins for patient-specific therapy.
- Real-world evidence using electronic health records.



3. Biomedical Databases and Multi-Omics Resources for AI-Driven Drug Repurposing

3.1 Introduction

The remarkable success of artificial intelligence (AI) and machine learning (ML) in drug repurposing relies heavily on the availability of comprehensive, high-quality biomedical datasets. Unlike conventional computational approaches that use a single data source, AI-driven drug repurposing integrates heterogeneous information from chemical, genomic, transcriptomic, proteomic, metabolomic, pharmacological, clinical, and literature databases. These resources enable computational models to identify hidden relationships among drugs, targets, genes, proteins, biological pathways, and diseases, thereby improving prediction accuracy and accelerating therapeutic discovery.

Recent advances in multi-omics technologies have further expanded the scope of drug repurposing. Integration of genomics, transcriptomics, proteomics, metabolomics, epigenomics, and single-cell sequencing data provides a systems-level understanding of disease mechanisms and drug responses.

3.2 Multi-Omics Data Sources

Multi-omics integration combines complementary biological information to better understand disease mechanisms.

Table 2.1: Multi-Omics Data Used for Drug Repurposing

Omics Layer	Biological Information	AI Applications
Genomics	DNA mutations	Target identification
Transcriptomics	Gene expression	Drug signature matching
Proteomics	Protein abundance	Biomarker prediction
Metabolomics	Metabolite profiles	Disease metabolism
Epigenomics	DNA methylation	Regulatory mechanism
Single-cell omics	Cell heterogeneity	Personalized medicines

3.3 Chemical Databases

Chemical databases provide molecular descriptors, physicochemical properties, molecular fingerprints, and bioactivity information required for machine learning algorithms.

These databases support:

- Molecular similarity analysis
- Drug-target interaction prediction
- Molecular docking
- QSAR modelling
- Virtual screening
- Deep learning-based molecular representation

3.4 Pharmacogenomic Databases

Pharmacogenomic resources link genetic variation with drug response and toxicity.

Important information includes:

- Drug-response biomarkers
- Drug sensitivity
- Drug resistance
- Precision medicine targets

- Pharmacokinetics
- Pharmacodynamics

These datasets are increasingly used for personalized drug repurposing.

3.5 Transcriptomic Databases

Transcriptomic repositories are among the most valuable resources for AI-based drug repurposing.

Drug-induced gene expression signatures are compared with disease-associated signatures to identify drugs capable of reversing disease phenotypes.

The LINCS L1000 and Connectivity Map databases have become standard resources for transcriptomics-driven drug repositioning.

3.6 Protein Interaction Databases

Protein-protein interaction (PPI) databases enable network-based drug discovery.

PPI networks help identify:

- Disease modules
- Hub proteins
- Essential signaling pathways
- Novel therapeutic targets

3.7 Clinical Data Resources

Clinical information complements molecular datasets.

Major sources include:

- Electronic Health Records (EHRs)
- Clinical trial repositories
- Adverse event databases
- Insurance claims databases
- Real-world evidence registries

Clinical datasets improve the translational relevance of AI predictions.

3.8 Challenges in Biomedical Data Integration

Despite the abundance of public databases, several limitations remain.

- Missing values
- Batch effects
- Data imbalance
- Experimental variability
- Different data formats
- Limited clinical annotations
- Poor interoperability
- Small sample sizes for rare diseases

These challenges necessitate advanced data preprocessing, feature engineering, and robust AI algorithms.

Future Perspectives

The next generation of AI-driven drug repurposing will increasingly rely on:

- Integrated multi-omics datasets
- Knowledge graphs
- Graph Neural Networks (GNNs)
- Foundation models

- Large Language Models (LLMs)
- Explainable AI
- Federated learning
- Digital twins
- Real-world evidence

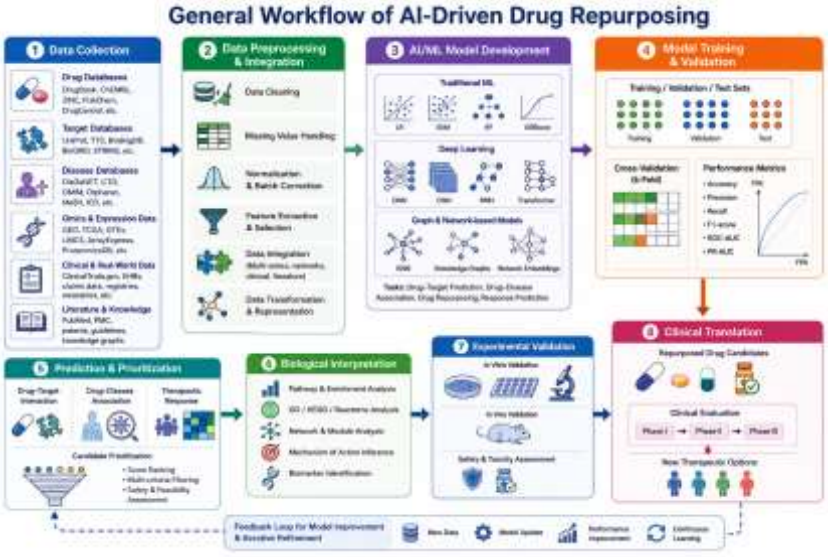
These advances are expected to substantially improve prediction accuracy and facilitate precision medicine.

4. Machine Learning and Deep Learning Algorithms for Drug Repurposing

4.1 Introduction

Artificial intelligence (AI) has revolutionized computational drug repurposing by enabling the analysis of large-scale biomedical datasets that are difficult to interpret using conventional statistical methods. Machine learning (ML) and deep learning (DL) algorithms can identify complex nonlinear relationships among drugs, genes, proteins, diseases, and biological pathways, allowing rapid prediction of novel drug-disease associations.

AI-based drug repurposing generally consists of four major steps: data acquisition, feature engineering, model development, and biological validation. Depending on the availability of labelled data and prediction objectives, supervised, unsupervised, semi-supervised, and reinforcement learning approaches have been applied. Recently, deep learning architectures including convolutional neural networks (CNNs), recurrent neural networks (RNNs), graph neural networks (GNNs), and transformer-based models have become state-of-the-art methods because of their ability to learn hierarchical representations directly from heterogeneous biomedical data.



4.2 Classical Machine Learning Algorithms

Traditional machine learning methods remain widely used because of their simplicity, robustness, and relatively low computational requirements.

Support Vector Machine (SVM)

SVM is one of the earliest and most widely applied algorithms in computational drug repurposing. It identifies an optimal hyperplane that separates different therapeutic classes while maximizing the

classification margin. SVM performs well on high-dimensional datasets such as transcriptomic and pharmacogenomic profiles.

Applications

- Drug-target interaction prediction
- Drug classification
- Molecular docking
- Virtual screening

Advantages

- High predictive accuracy
- Effective with small datasets
- Robust to high-dimensional data

Limitations

- Difficult interpretation
- Computationally intensive for very large datasets

4.3 Deep Learning Algorithms

Deep learning has dramatically improved computational drug repurposing by automatically extracting informative features from complex biomedical datasets.

4.3.1 Deep Neural Networks (DNN)

DNNs consist of multiple hidden layers capable of modeling nonlinear biological relationships. Applications include:

- Drug indication prediction
- Drug response prediction
- Transcriptomic analysis
- Pharmacological classification

4.3.2 Convolutional Neural Networks (CNN)

CNNs are primarily used to analyze molecular images, protein structures, and chemical fingerprints. Applications:

- Molecular property prediction
- Drug toxicity
- Chemical structure analysis

4.3.3 Recurrent Neural Networks (RNN) and Long Short-Term Memory (LSTM)

RNNs process sequential biological information such as genomic and protein sequences. Applications:

- DNA sequence analysis
- Protein sequence prediction
- Time-series clinical data

4.3.4 Autoencoders

Autoencoders compress high-dimensional omics datasets into informative latent representations. Applications:

- Feature extraction
- Data denoising
- Drug similarity analysis

4.3.5 Variational Autoencoders (VAE)

VAEs learn probability distributions of molecular structures and generate novel drug candidates.

4.4 Graph Neural Networks (GNN)

Graph Neural Networks represent the latest generation of AI algorithms in drug repurposing. Instead of treating drugs independently, GNNs model relationships among drugs, diseases, proteins, genes, and biological pathways. These networks efficiently propagate information through biological interaction graphs. Applications:

- Drug-target interaction
- Drug-disease association
- Protein interaction prediction
- Network pharmacology

4.5 Transformer Models

Transformer architectures employ self-attention mechanisms that capture long-range relationships among molecular entities. Applications include:

- Molecular representation learning
- Protein language models
- Drug discovery
- Biomedical text mining

4.6 Large Language Models (LLMs)

Recent foundation models are increasingly used for:

- Literature mining
- Hypothesis generation
- Knowledge extraction
- Drug indication prediction
- Clinical decision support

4.7 Explainable Artificial Intelligence (XAI)

Most deep learning models function as black boxes, limiting clinician confidence. Explainable AI improves transparency by identifying the biological features responsible for predictions. Common XAI methods include:

- SHAP
- LIME
- Attention visualization
- Feature attribution
- Saliency mapping

Explainability is becoming increasingly important for regulatory approval and clinical adoption.

Table 3.2: Comparison of ML and DL Approaches

Parameter	Machine Learning	Deep Learning
Feature engineering	Manual	Automatic
Data requirement	Moderate	Very large
Computational cost	Low	High
Interpretability	Better	Lower

Accuracy	Moderate	High
Scalability	Moderate	Excellent
Multi-omics integration	Limited	Excellent

Challenges

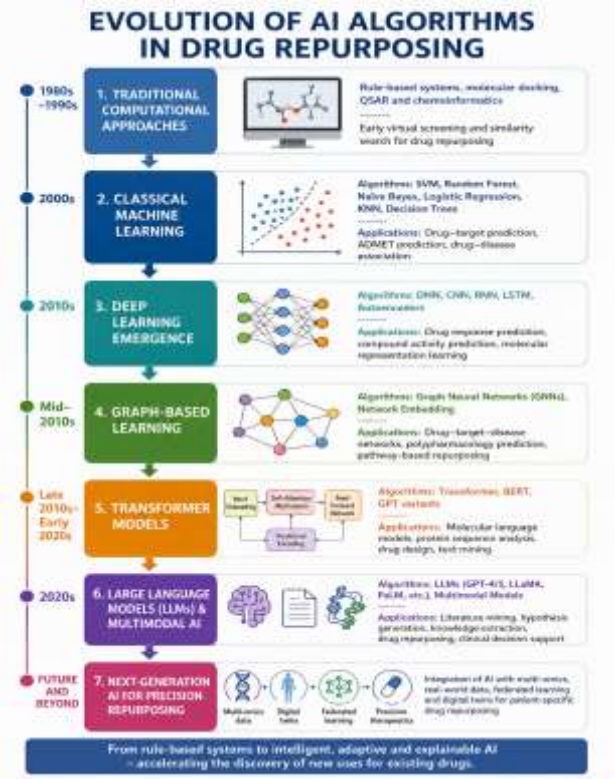
Although AI has transformed drug repurposing, several limitations remain:

- Limited labelled biomedical datasets
- Class imbalance
- Data heterogeneity
- Poor external validation
- Black-box predictions
- High computational cost
- Lack of standardized benchmarks
- Regulatory uncertainty

Future Directions

Future AI systems are expected to combine:

- Multi-modal foundation models
- Knowledge graphs
- Graph Neural Networks
- Explainable AI
- Federated learning
- Digital twins
- Real-world evidence
- Precision medicine



5. Disease-Specific Applications of AI-Driven Drug Repurposing

5.1 Introduction

Artificial intelligence (AI) and machine learning (ML) have significantly expanded the scope of drug repurposing across a wide range of diseases. By integrating chemical structures, molecular targets, transcriptomic profiles, protein-protein interaction networks, electronic health records (EHRs), and clinical data, AI-based approaches have accelerated the identification of promising therapeutic candidates. Disease-specific models can capture unique molecular mechanisms, enabling more accurate prediction of drug-disease associations and supporting precision medicine. This section summarizes the major applications of AI-driven drug repurposing in oncology, infectious diseases, neurological disorders, cardiovascular diseases, metabolic diseases, and rare diseases.

5.2 Cancer

Cancer is one of the most extensively studied areas for AI-assisted drug repurposing because of its complex molecular heterogeneity and the availability of large multi-omics datasets such as TCGA, GEO, LINCS, and CCLE. Deep learning, graph neural networks (GNNs), and network pharmacology have been widely used to identify novel anticancer indications for approved drugs.

Several studies have demonstrated the effectiveness of transcriptomic signature matching, drug-target interaction prediction, and knowledge graph analysis in discovering candidate therapies. AI models integrate genomic mutations, gene expression patterns, signalling pathways, and protein interaction networks to identify drugs capable of reversing tumor-associated molecular signatures. These approaches have successfully proposed new therapeutic candidates for breast cancer, lung cancer, colorectal cancer, leukemia, melanoma, and glioblastoma.

5.3 Infectious Diseases

The COVID-19 pandemic demonstrated the value of AI-driven drug repurposing for rapidly identifying therapeutic candidates. Large-scale screening of approved drugs using transcriptomic signatures, molecular docking, and deep learning enabled rapid prioritization of antiviral agents. Similar approaches have also been applied to tuberculosis, influenza, HIV/AIDS, malaria, and neglected tropical diseases.

Machine learning models combine viral genome information, host response signatures, drug-target interactions, and molecular docking scores to identify compounds with potential antiviral activity. Although several AI-predicted candidates entered clinical trials during the COVID-19 pandemic, the experience also highlighted the need for rigorous experimental validation before clinical implementation.

5.4 Neurological Disorders

Neurological diseases such as Alzheimer's disease, Parkinson's disease, Huntington's disease, and amyotrophic lateral sclerosis (ALS) present significant challenges because of their multifactorial pathogenesis. AI integrates genomics, transcriptomics, neuroimaging, and protein interaction data to identify drugs capable of modulating neurodegenerative pathways.

Deep learning models have been used to predict compounds targeting amyloid-beta deposition, tau phosphorylation, oxidative stress, mitochondrial dysfunction, and neuroinflammation. Knowledge graph approaches have further facilitated the identification of shared molecular mechanisms among neurodegenerative disorders.

5.5 Cardiovascular Diseases

AI has icreasingly been applied to cardiovascular medicine by integrating clinical data, genomics, proteomics, metabolomics, and imaging datasets. Machine learning models identify novel therapeutic

opportunities for hypertension, heart failure, myocardial infarction, and atherosclerosis.

Drug repurposing in cardiovascular diseases focuses on reducing inflammation, improving endothelial function, and modulating lipid metabolism. Electronic health records and real-world evidence have become particularly valuable resources for identifying clinically relevant drug-disease associations.

5.6 Metabolic Diseases

Type 2 diabetes mellitus, obesity, and non-alcoholic fatty liver disease (NAFLD) are major targets for AI-assisted drug repurposing. Deep learning algorithms analyze metabolic pathways, insulin signaling, gut microbiome interactions, and transcriptomic profiles to identify candidate therapies.

Several approved drugs originally developed for cardiovascular or inflammatory diseases have shown promise for metabolic disorders through AI-based prediction frameworks.

5.7 Rare Diseases

Rare diseases often suffer from limited patient populations and insufficient commercial incentives for conventional drug development. Drug repurposing offers an efficient strategy for identifying treatments for orphan diseases.

AI integrates genomic sequencing, literature mining, knowledge graphs, and electronic health records to predict therapeutic candidates despite limited clinical data. Network-based methods are particularly useful because they leverage shared biological pathways among related diseases.

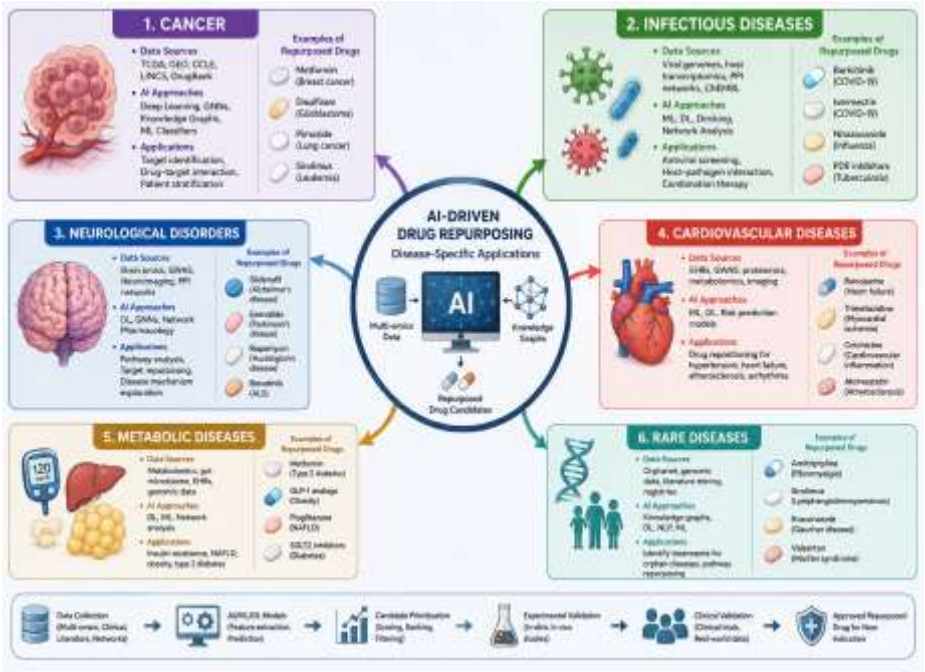
Table 5.1: Summary of Disease-Specific Applications

Disease Category	Common AI Methods	Major Data Sources	Main Challenge
Cancer	Deep Learning, GNN	TCGA, GEO, LINCS	Tumor heterogeneity
Infectious diseases	ML, Docking	Viral genomes, transcriptomics	Viral mutation
Neurological disorders	Knowledge Graph, DL	Multi-omics	Complex pathogenesis
Cardiovascular diseases	ML	EHRs, metabolomics	Clinical variability
Metabolic diseases	Deep Learning	Multi-omics	Lifestyle influences
Rare diseases	Network AI	Genomics	Limited datasets

Challenges in Disease-Specific AI Models

Despite significant progress, disease-specific AI models face several limitations:

- Limited availability of high-quality disease-specific datasets.
- Population heterogeneity and demographic bias.
- Difficulty integrating heterogeneous multi-omics data.
- Lack of external validation across independent cohorts.
- Limited interpretability of deep learning predictions.
- Regulatory and ethical challenges for clinical implementation.



Future Perspectives

Future disease-specific drug repurposing will increasingly rely on:

- Multi-modal AI integrating genomics, proteomics, metabolomics, and clinical data.
- Graph Neural Networks for disease network modeling.
- Foundation models and Large Language Models (LLMs) for biomedical knowledge synthesis.
- Explainable AI (XAI) to improve transparency.
- Federated learning for secure collaboration across healthcare institutions.
- Digital twins for patient-specific therapeutic prediction.
- Real-world evidence from electronic health records and wearable devices.

6. Current Challenges, Explainable AI, and Future Perspectives in AI-Driven Drug Repurposing

6.1 Introduction

Although artificial intelligence (AI) has significantly accelerated drug repurposing, its translation into routine clinical practice remains limited. The predictive performance of AI models depends heavily on the quality, diversity, and completeness of biomedical datasets. Challenges such as heterogeneous data sources, algorithmic bias, limited interpretability, lack of standardized benchmarks, and insufficient clinical validation continue to hinder widespread adoption. Addressing these issues is essential for developing reliable, transparent, and clinically applicable AI-driven drug repurposing frameworks.

6.2 Data-Related Challenges

High-quality biomedical data are fundamental for successful AI models. However, existing datasets often contain missing values, inconsistent annotations, batch effects, and class imbalance, which can adversely affect prediction accuracy. Major challenges include:

- Incomplete drug-target interaction data.
- Limited patient-specific multi-omics datasets.
- Inconsistent data formats across public databases.
- Population bias due to underrepresentation of diverse ethnic groups.

- Limited availability of rare disease datasets.
- Noise in transcriptomic and proteomic data.

Effective preprocessing, normalization, feature selection, and harmonization techniques are therefore crucial for improving model robustness.

6.3 Algorithmic Challenges

Despite impressive predictive performance, many AI algorithms face methodological limitations. These include:

- Overfitting due to small datasets.
- Poor generalization across independent datasets.
- High computational complexity.
- Hyperparameter optimization difficulties.
- Scalability limitations.
- Lack of standardized evaluation metrics.

Developing benchmark datasets and reproducible validation pipelines remains a high priority.

6.4 Explainable Artificial Intelligence (XAI)

One of the most important limitations of deep learning models is their lack of interpretability. Many high-performing algorithms operate as black boxes, making it difficult for clinicians and regulatory agencies to understand how predictions are generated. Explainable Artificial Intelligence (XAI) addresses this issue by providing transparent and interpretable insights into model decision-making. Common XAI techniques include:

- SHAP (Shapley Additive explanations)
- LIME (Local Interpretable Model-agnostic Explanations)
- Attention maps
- Feature attribution methods
- Saliency analysis

By highlighting the biological features that contribute most strongly to predictions, XAI improves confidence in AI-assisted drug repurposing and supports regulatory acceptance.

6.5 Federated Learning

Biomedical datasets are often distributed across hospitals and research institutions, making centralized data sharing difficult because of privacy regulations. Federated learning enables collaborative model training without transferring sensitive patient data.

Advantages

- Preserves patient privacy.
- Facilitates multi-center collaboration.
- Improves model generalizability.
- Supports regulatory compliance.

6.6 Foundation Models and Large Language Models

Foundation models trained on extensive biomedical datasets are increasingly being applied to drug discovery. Large Language Models (LLMs) can:

- Analyse biomedical literature.
- Summarize clinical evidence.
- Extract drug-gene-disease relationships.
- Generate scientific hypotheses.

- Support clinical decision-making.

When integrated with molecular and multi-omics data, foundation models have the potential to substantially accelerate therapeutic discovery.

6.7 Digital Twins

Digital twins are computational representations of individual patients that integrate genomic, clinical, imaging, and physiological information. Potential applications include:

- Personalized drug repurposing.
- Virtual clinical trials.
- Treatment optimization.
- Prediction of adverse drug reactions.

Although still in early development, digital twins represent a promising direction for precision medicine.

6.8 Regulatory and Ethical Considerations

The clinical implementation of AI-driven drug repurposing requires careful consideration of regulatory and ethical issues. Major concerns include:

- Patient privacy.
- Data security.
- Algorithmic bias.
- Model transparency.
- Reproducibility.
- Clinical accountability.
- Regulatory approval.

Collaboration among researchers, clinicians, regulatory agencies, and industry will be essential for establishing standardized evaluation frameworks.

6.9 Future Perspectives

Future AI-assisted drug repurposing will likely focus on integrating diverse biomedical data sources into unified predictive frameworks. Key research priorities include:

- Multi-modal multi-omics integration.
- Explainable and trustworthy AI.
- Foundation models for molecular biology.
- Large-scale knowledge graphs.
- Real-world evidence from electronic health records.
- Digital twins for personalized therapy.
- Federated learning across healthcare systems.
- Standardized benchmark datasets.
- Prospective clinical validation.
- AI-assisted precision medicine.

These advances have the potential to transform drug repurposing from a retrospective computational strategy into a proactive platform for personalized therapeutic discovery.

Table 6.1: Future Research Priorities

Research Area	Expected Impact
Multi-omics integration	Improved prediction accuracy
Explainable AI	Increased clinical trust

Knowledge graphs	Better biological interpretation
Foundation models	Automated knowledge discovery
Federated learning	Secure data sharing
Digital twins	Personalized treatment
Real-world evidence	Enhanced clinical translation
Standardized benchmarks	Reproducible AI evaluation

Conclusion

Drug repurposing has emerged as one of the most promising strategies for accelerating drug discovery by identifying new therapeutic indications for approved and investigational drugs. Compared with conventional de novo drug development, drug repurposing significantly reduces development time, cost, and clinical risk while taking advantage of existing pharmacokinetic, pharmacodynamics, and safety information. The rapid growths of biomedical big data and advances in artificial intelligence (AI) have transformed drug repurposing from a largely serendipitous process into a systematic, data-driven approach.

Machine learning (ML) algorithms, including Support Vector Machines (SVM), Random Forests (RF), Gradient Boosting, and XGBoost, have demonstrated considerable success in predicting drug-target interactions, drug-disease associations, and therapeutic responses. More recently, deep learning architectures such as Deep Neural Networks (DNNs), Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), Autoencoders, Transformer models, and Graph Neural Networks (GNNs) have enabled the analysis of highly complex biological datasets, substantially improving prediction performance.

The integration of diverse biomedical resources, including DrugBank, ChEMBL, BindingDB, LINCS, GEO, TCGA, KEGG, Reactome, and electronic health records (EHRs), has further strengthened AI-driven drug repurposing by enabling comprehensive analyses of chemical structures, transcriptomics, proteomics, metabolomics, protein-protein interactions, and clinical phenotypes. These integrated

datasets support systems-level investigations that enhance the identification of novel therapeutic opportunities across multiple disease areas.

AI-assisted drug repurposing has already demonstrated significant applications in oncology, infectious diseases, neurodegenerative disorders, cardiovascular diseases, metabolic disorders, and rare diseases. During the COVID-19 pandemic, computational repurposing approaches highlighted the potential of AI for rapidly prioritizing candidate therapeutics under urgent public health conditions. Similarly, increasing applications in precision oncology and personalized medicine illustrate the growing clinical value of computational drug repositioning.

Despite these achievements, several important challenges remain. Data heterogeneity, incomplete biological knowledge, limited external validation, class imbalance, insufficient interpretability, and regulatory uncertainties continue to limit clinical translation. Many current AI models function as black boxes, reducing clinician confidence and complicating regulatory approval. Furthermore, the lack of standardized benchmark datasets and prospective multicenter validation restricts objective comparison among different computational methods.

Emerging technologies are expected to overcome many of these limitations. Multi-omics integration, knowledge graphs, Graph Neural Networks, foundation models, Large Language Models (LLMs), Explainable Artificial Intelligence (XAI), federated learning, digital twins, and real-world evidence have the potential to substantially improve predictive accuracy, biological interpretability, and clinical applicability. These innovations are likely to facilitate personalized drug repurposing by integrating genomic, transcriptomic, proteomic, metabolomic, clinical, and lifestyle information into unified computational frameworks.

Overall, AI-driven drug repurposing represents a transformative paradigm in modern pharmaceutical research. Continued collaboration among computational scientists, biologists, clinicians, regulatory agencies, and the pharmaceutical industry will be essential for translating computational predictions into clinically effective therapies. Future research should prioritize transparent, explainable, and clinically validated AI models supported by standardized evaluation frameworks and large-scale multimodal biomedical datasets. Such efforts will accelerate the development of safer, more effective, and personalized therapeutics while advancing precision medicine and improving global healthcare outcomes.

References

1. Pan X, Lin X, Cao D, et al. Deep learning for drug repurposing: Methods, databases, and applications. WIREs Computational Molecular Science. 2022; 12:e1597.
2. Yang F, Zhang Q, Ji X, et al. Machine learning applications in drug repurposing. Interdisciplinary Sciences: Computational Life Sciences. 2022; 14:15-21.
3. Tanoli Z, Vaha-Koskela M, Aittokallio T. Artificial intelligence, machine learning, and drug repurposing in cancer. Expert Opinion on Drug Discovery. 2021; 16:977-989.
4. Urbina F, Puhl AC, Ekins S. Recent advances in drug repurposing using machine learning. Current Opinion in Chemical Biology. 2021; 65:74-84.
5. Li B, Dai C, Wang L, et al. A novel drug repurposing approach for non-small cell lung cancer using deep learning. PLoS ONE. 2020; 15:e0233112.
6. Koromina M, Pandi MT, Patrinos GP. Rethinking drug repositioning and development with artificial intelligence, machine learning, and omics. OMICS. 2019; 23:539-548.

7. Aliper A, Plis S, Artemov A, et al. Deep learning applications for predicting pharmacological properties of drugs and drug repurposing using transcriptomic data. Molecular Pharmaceutics. 2016; 13:2524-2530.
8. Napolitano F, Zhao Y, Moreira VM, et al. Drug repositioning: A machine-learning approach through data integration. Journal of Cheminformatics. 2013;5:30.
9. Kinnings SL, Liu N, Tonge PJ, et al. A machine learning-based method to improve docking scoring functions and its application to drug repurposing. Journal of Chemical Information and Modeling. 2011; 51:408-419.
10. Kim E, Choi A, Nam H. Drug repositioning of herbal compounds via a machine-learning approach. BMC Bioinformatics. 2019; 20(Suppl).
11. Bakhtiari N, Safaei AA, Ebadi MJ. Drug Repurposing Using Machine Learning and Deep Learning: A Systematic Literature Review. Current Computer-Aided Drug Design. 2026.
12. Zielinska A, Fornalik M, Szczepaniak M, et al. Artificial intelligence in drug research and development: a review of methods and applications in drug repurposing. Briefings in Bioinformatics. 2026.
13. Cheng F. Artificial Intelligence to Guide Repurposing of Drugs. Annual Review of Medicine. 2026.
14. Stokes JM, et al. AI-enabled antibiotic discovery. Cell. 2024.
15. Jumper J, et al. AlphaFold 3. Nature. 2024.
16. Fu T, et al. TxGNN: Foundation graph neural network for therapeutic discovery. Nature Machine Intelligence.
17. Zitnik M, et al. Graph neural networks for biomedicine. Nature Reviews Methods Primers.
18. Zitnik M, Leskovec J. Predicting multicellular function through graph representation learning.
19. Zitnik M. Machine learning for precision medicine.
20. Ekins S, Puhl AC. Artificial intelligence in drug discovery and repurposing.
21. Ramsundar B, et al. Deep Learning for the Life Sciences.
22. Gaudelet T, et al. Utilizing graph machine learning within drug discovery.
23. Vamathevan J, et al. Applications of machine learning in drug discovery and development. Nature Reviews Drug Discovery.
24. Schneider P, et al. Rethinking drug design in the artificial intelligence era. Nature Reviews Drug Discovery.
25. Zhavoronkov A. Artificial intelligence for drug discovery, biomarker development and generation of novel therapeutics.
26. Walters WP, Murcko MA. Assessing the impact of AI on medicinal chemistry.
27. Ekins S, et al. Exploiting machine learning for drug repositioning.
28. Chen H, Engkvist O, Wang Y, et al. The rise of deep learning in drug discovery.
29. Jimenez-Luna J, Grisoni F, Schneider G. Drug discovery with explainable artificial intelligence.
30. Gao K, et al. Interpretable artificial intelligence for drug discovery.
31. Inoue Y, Song T, Fu T. DrugAgent: Explainable Drug Repurposing Agent with Large Language Model-based Reasoning. 2024 (preprint).
32. Lu Y, Hu Y, Li C. DrugCLIP: Contrastive Drug-Disease Interaction for Drug Repurposing. 2024 (preprint).
33. Yang Z, et al. DrugMCTS: A drug repurposing framework combining multi-agent, RAG and Monte Carlo Tree Search. 2025 (preprint).

34. AlphaFold 3 applications in AI-driven drug discovery. Nature. 2024.
35. Knowledge graph applications in drug repurposing. Briefings in Bioinformatics. 2024.
36. Explainable AI for biomedical prediction. Nature Machine Intelligence. 2024.
37. Federated learning in precision medicine. Nature Medicine. 2024.
38. Large language models for biomedical discovery. Nature Biotechnology. 2024.
39. Multi-omics integration using deep learning for precision medicine. Nature Reviews Bioengineering. 2024.
40. Foundation models for computational biology and drug discovery. Nature Reviews Drug Discovery. 2025.