

AI-POWERED DRUG REPURPOSING: A NOVEL APPROACH TO ACCELERATE DRUG DEVELOPMENT

Mahvish Akhlaqe¹, Kuldeep Singh^{1,*}, Arun Kumar¹, Shom Prakash Kushwaha¹, Suvaiv¹, and Piyush Kumar²

Received: September 01, 2025; Revised: December 15, 2025; Accepted: December 29, 2025

Abstract

The conventional drug development pipeline is expensive, time-consuming, and prone to failure. Artificial intelligence (AI) aids in drug repurposing by allowing for the prompt discovery of previously unknown therapeutic uses for approved drugs. This manuscript examines AI approaches used in drug repurposing, including machine learning, deep learning, and multi-omics data integration. We demonstrate how these methods enable tailored treatment, speed up virtual screening, and reveal hidden drug-target correlations. AI overcomes the drawbacks of conventional methods by evaluating extensive biological and clinical datasets; this is demonstrated by its crucial role in the quick identification of COVID-19 therapies. Finally, integrating AI results into clinical practice presents a unique set of challenges. It needs interdisciplinary cooperation and a deep comprehension of regulatory frameworks to close the gap between cutting-edge computational methods and practical healthcare applications. To successfully include AI into drug development processes, it is crucial to make sure that findings derived from AI are both scientifically solid and applicable in clinical settings. In conclusion, although AI has revolutionary potential to improve drug development and develop patient-specific treatments, its successful application in healthcare depends on resolving issues like data quality, algorithm interpretability, and the complexities of clinical translation.

Keywords: Algorithm Interpretability; Artificial Intelligence; Computational Methods; Multi-Omics; Virtual Screening

¹ Department of Pharmacy, Integral University, Lucknow, UP, India-226026. E-mail: mahvisha@student.iul.ac.in, ORCID: 0009-0001-8988-3238; kuldeep@iul.ac.in*, ORCID: 0000-0001-5418-945X; arun@iul.ac.in, ORCID: 0000-0002-6323-7377; shompk@iul.ac.in, ORCID: 0000-0002-9844-400X; suvaiv@student.iul.ac.in, ORCID: 0000-0003-0240-8029

² Department of Chemistry, Indian Institute of Technology, Jammu, Jammu and Kashmir, India. E-mail: piyush.kumar@iitjammu.ac.in, ORCID: 0000-0002-4475-1557

* Corresponding author

DOI: <https://doi.org/10.55766/sujst11313>

Introduction

The pharmaceutical business is leading the way as one of the major beneficiaries of the current expansion of artificial intelligence (AI) applications into other sectors. In recent years, the pharmaceutical industry has witnessed a notable rise in data digitization. However, this transition presents challenges in acquiring, analyzing, and utilizing that information to address challenging clinical issues. This is where AI proves beneficial, as it can efficiently handle large volumes of data through enhanced automation (Ashburn and Thor, 2004). AI comprises a range of advanced tools and networks that mimic human intelligence, without fully replacing the need for human involvement. It utilizes systems and software capable of interpreting and learning from data inputs, enabling it to make independent decisions to achieve specific objectives. The McKinsey Global Institute predicts that the swift development of AI-driven automation would drastically alter society's work culture (Crevier, 1993; Chong and Sullivan, 2007).

The drug development process could be substantially transformed by drug repurposing. Repurposing can significantly reduce the time and expense normally associated with creating new therapeutics by finding new therapeutic uses for already-approved treatments. Developing a new medication from the ground up takes more than ten years on average, costs more than \$1 billion, and has a high failure rate in clinical trials. Conversely, repurposing a drug that has already been approved for a different use allows for a far faster admission into clinical trials (Crevier, 1993; Jordan and Mitchell, 2015; Mucke, 2015).

Because the safety profiles of repurposed medications are already established, this simplified strategy not only decreases costs by about 30–40% but also lowers the risks involved in creating new medications. Additionally, drug repurposing can help address unmet medical needs, especially for complex or rare diseases where it can be expensive and time-consuming to find new medicines. The immense promise of this approach is demonstrated by the success of drugs like remdesivir, which was used to treat COVID-19, and thalidomide, which was first created as a sedative and then repurposed for multiple myeloma. There is enormous potential to discover novel therapeutics by repurposing existing drug libraries and utilizing AI and machine learning technology. This would give patients fresh hope and greatly improve healthcare results (Oprea and Mestres, 2012; Pushpakom *et al.*, 2019).

Drug Repurposing

Drug repurposing, also known as drug repositioning, is a promising drug-development strategy that identifies new therapeutic uses for existing, approved, investigational, or previously discontinued drugs. This approach can shorten the time required for drug discovery and development, improve productivity, and avoid some of the most costly stages of the conventional drug-development process. Because the safety and pharmacokinetic profiles of repurposed drug candidates are often already established, drug repurposing may reduce the overall cost and risk of bringing a treatment to market. Although the terms “drug repurposing” and “drug repositioning” are often used interchangeably in the literature, this review uses the term “drug repurposing” throughout for consistency (Aliper *et al.*, 2016; Alzubaidi *et al.*, 2021). One well-known example is sildenafil, which was initially investigated for hypertension and was later approved for the treatment of erectile dysfunction.

Another notable example of drug repurposing is dimethyl fumarate, which had been used to treat psoriasis in Europe for more than 20 years before being approved in 2013 for the treatment of multiple sclerosis (Mrowietz and Asadullah, 2005; Kappos *et al.*, 2008). Thalidomide is another well-known example; its antiangiogenic and immunomodulatory properties contributed to its repurposing for the treatment of multiple myeloma (Singhal *et al.*, 1999).

Drug repurposing can help address several challenges in drug development by evaluating approved, investigational, or discontinued drugs for new therapeutic indications. It also offers opportunities to develop therapies for rare diseases and reduce development costs by lowering the risk of failure during the drug-discovery process (Korkmaz *et al.*, 2014).

The Rise of Artificial Intelligence

Early achievements in logic-based problem-solving in 1956 marked the beginning of the area of artificial intelligence (AI). But there was a downturn because of the processing power constraints and the dependence on rule-based systems (Lima *et al.*, 2016). AI is already being used in domains such as drug repurposing to analyze large biological datasets because of advances in computing, especially GPUs, and the shift toward data-driven machine learning. In the 1980s, artificial

intelligence (AI) saw a comeback because of “expert systems,” which turned expert knowledge into rules. But these methods had drawbacks, and funding decreased once more. By enabling deeper neural networks and resolving training issues, deep learning (DL) transformed artificial intelligence in 2006. Big data and more processing power are essential for DL, a branch of machine learning (ML), which enables computers to “learn” from data. These days, DL algorithms are highly effective in domains like computer vision and natural language processing, influencing everything from economics to science. By facilitating analysis, prediction, and evidence-based decision-making, machine learning (ML), a fundamental branch of AI, is revolutionizing data-intensive sciences and providing a more efficient alternative to manual programming. These days, AI methods are widely used in many scientific fields, spurring innovation and improving society (Wassermann *et al.*, 2011). There are numerous applications of AI in public health. The use of artificial intelligence models to find and expedite the repurposing of remdesivir, an antiviral medication initially created for Ebola, as a treatment for COVID-19 is a well-known example of AI-powered drug repurposing (Hameed *et al.*, 2018). In order to forecast remdesivir’s effectiveness against SARS-CoV-2 by mimicking its binding to viral RNA polymerase and reducing off-target effects, artificial intelligence (AI) algorithms examined enormous datasets during the 2020 pandemic, including viral-host interactomes,

drug-target networks, and clinical trial data. In the wake of the COVID-19 epidemic, China created a digital health quick response (QR) code system that leverages mobile big data to determine a person’s health status after first identifying potential contact with confirmed COVID-19 patients (Madugula *et al.*, 2021). Red indicates a verified COVID-19 patient, orange indicates a risky situation requiring quarantine, and green indicates a healthy person who may go about their everyday lives. The general public finds it easy to use, and many other nations have adopted it. The COVID-19 pandemic has been prevented and controlled globally thanks in large part to the health QR code (Alkan and Erten, 2017; Lee *et al.*, 2020).

AI can Address the Limitations of Traditional Drug Repurposing

A multi-step procedure for finding and creating novel therapeutic medicines is represented by the conventional drug discovery pipeline (Figure 1) (Herráiz-Gil *et al.*, 2025). Finding a particular chemical or pathway within a pathogen that is essential to its survival or virulence is the main goal of the first stage (Wei *et al.*, 2019). Computational modeling, genetic engineering, and biochemical testing may all be used in this (Khurana *et al.*, 2023). Finding possible therapeutic compounds that have the ability to interact with the verified target is the next step (Lv *et al.*, 2021). This has historically depended on extracts from natural products of sizable chemical libraries. Furthermore, medicinal

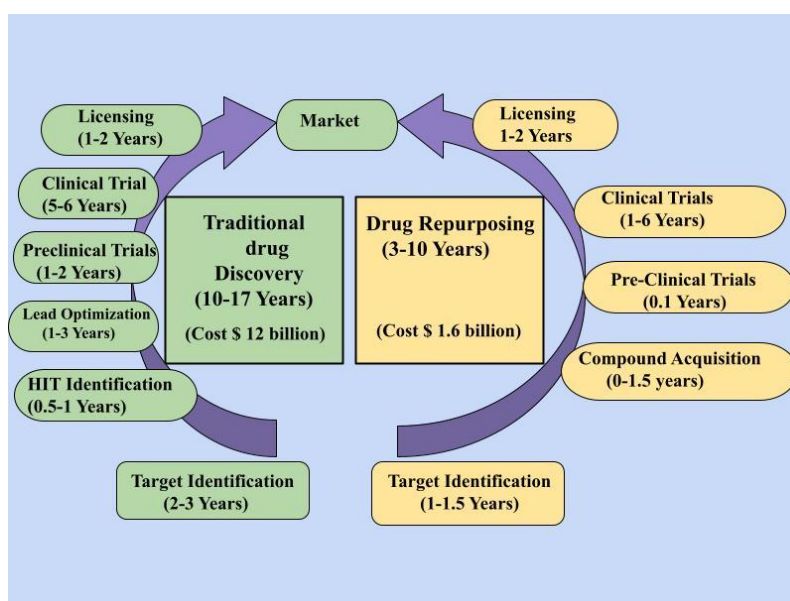


Figure 1. Target authentication, lead discovery, clinical and animal studies, approval by the FDA, and follow-up after market are all included in traditional pharmaceutical research

chemistry is essential for maximizing the potency, selectivity, and pharmacokinetic (absorption, distribution, metabolism, and excretion) characteristics of elements with lead (Zhou *et al.*, 2020). To evaluate their effectiveness and safety, promising lead compounds go through a rigorous in vitro and in vivo research process. The impact of the chemical on isolated cells or pathogen-infected tissues is assessed using in vitro experiments. Important information on the drug candidate's efficacy, toxicity, and metabolism is provided by in vivo investigations, which are usually carried out in animal models (Wang, 2020). The medication candidate moves on to human clinical trials if preclinical research is positive. In order to verify efficacy and safety in comparison to current treatments, with Phase I examining how well it is tolerated in healthy volunteers, Phase II evaluating results in a handful of patients with the target condition, and Phase III incorporating larger patient groups, these trials frequently take place in steps (Liu *et al.*, 2022). The medication candidate goes through a thorough assessment process by regulatory bodies after completing successful clinical studies. Marketing approval is based on proving the drug's quality, safety, and effectiveness. After a medicine is licensed, post-marketing surveillance is essential for tracking its long-term safety profile and spotting any unanticipated side effects (Serafim *et al.*, 2021).

AI Methods for Drug Repurposing

Vast quantities of cellular and chemical data can be analyzed by AI systems to find undetected relationships between current illicit substances, disease targets, and novel treatments. The following noteworthy automated projects are changing the fabrication and reuse of medications for infectious sicknesses: (Zhou *et al.*, 2020).

AI-driven Virtual Screening

Monitoring practically uses computer approaches to figure-out enormous collections of compounds in order to discover that may interact

with a given biological target AI-driven virtual screening relies heavily on Methods of deep computing and automated learning which help identify repurposed prospects faster (Hu *et al.*, 2021).

ML Assisted Virtual Screening

Data mining (Machine learning (ML)), a type of artificial intelligence, expedites drug repurposing by forecasting interactions through examination of drug-target data (Kappos *et al.*, 2008). This makes it possible to prioritize promising prospects and conduct effective virtual screening. AI's data analysis quickly identifies links between new targets and current medications. This method was essential for finding possible COVID-19 therapies quickly (Baker *et al.*, 2022). ML drastically cuts down on the span of time and money required for data-driven data advancement of drugs (Röösli *et al.*, 2021).

Kinds of Machine Learning

Supervised Learning

In supervised learning, algorithms are trained on labelled datasets that combine known outputs (like therapeutic efficacy or illness association) with input features (like gene expression patterns, molecular descriptors, or drug-target interactions) (Geldsetzer, 2020). The algorithm is an effective tool for drug repurposing since it learns a mapping function to anticipate results for new data. In this case, supervised learning uses patterns in biological, chemical, and clinical data to find current medications that can be repurposed to treat novel illnesses. Compared to conventional approaches, these models greatly cut down on the time and expense of drug development by predicting drug-disease pairings, ranking candidates, or estimating binding affinities by utilizing databases such as DrugBank or ChEMBL (Zhang, 2022). For instance, supervised learning can shorten the time needed to produce repurposed medications from 10 to 15 years to 1 to 3 years. Common supervised-learning algorithms used in drug-repurposing research include Random Forest and Support Vector Machines (SVM), as described below and summarized in Table 1.

Table 1. Application of Supervised Learning Algorithms in Drug Repurposing

Algorithm	Application in Drug Repurposing	Example Drug/Disease	References
Random forest	Predicting drug candidates for Alzheimer's by analyzing chemical structures and gene expression profiles	Donepezil/Alzheimer's	Rodriguez <i>et al.</i> (2021)
Support Vector Machines (SVM)	Identifying antiviral drugs for COVID-19 by modeling drug-target interactions	Remdesivir/COVID-19	Rodriguez <i>et al.</i> (2021)

The Algorithm

- Algorithms are trained using labelled data, or known drug-target interactions.
- To differentiate between medications that work and those that don't, they learn patterns.
- New drug-target interactions are predicted by the trained model.
- Promising drug candidates are given priority for additional testing (Ioakeim-Skoufa *et al.*, 2023)

Unsupervised Learning

Unsupervised learning is a machine learning paradigm where algorithms analyze unlabeled data to identify patterns, clusters, or relationships without predefined outputs.

Principles of Unsupervised Learning in Drug Repurposing

Unsupervised learning algorithms group or organize data based on inherent similarities or structures, making them ideal for:

- Clustering: Grouping drugs, targets, or diseases with similar properties (e.g., chemical scaffolds like coumarins or shared gene expression patterns) (Ioakeim-Skoufa *et al.*, 2023).
- Dimensionality Reduction: Simplifying high-dimensional datasets (e.g., molecular descriptors or omics data) to identify key features driving drug activity.
- Network Analysis: Mapping drug-target or drug-disease relationships in heterogeneous networks to uncover novel connections.
- It examines unlabeled data, such as medical information or chemical structures.
- Within this data, algorithms uncover hidden links and patterns.
- This aids in identifying novel relationships between medications and possible targets.
- It's useful for bringing to light details that conventional approaches overlook (Ioakeim-Skoufa *et al.*, 2023; Wang *et al.*, 2023).

Table 2. Selected AI-powered automated virtual screening platforms for drug repurposing

Tool/Platform	Justification	References
ZairaChem https://github.com/ersilia-os/zaira-chem	AI and ML-powered online platform for automated quantitative structure-activity relationship (QSAR) forecasting	Turon <i>et al.</i> (2023)
Insilico Medicine https://insilico.com/	AI-driven extensive platform that provides lead optimization tools and has a major emphasis on AI and deep learning.	Ivanenkov <i>et al.</i> (2023)
OpenEye Scientific https://www.eyesopen.com/	Software suite with various tools for computational drug discovery, including a module for AI-powered VS. Plenty of commercial and open-source choices available, including an adaptable framework with capabilities beyond that of Visual Studio.	Hawkins <i>et al.</i> (2007)
BenevolentAI https://www.benevolent.com/	Platform for AI-driven discovery of drugs that locates targets and repurposes medications using machine learning and network analysis. Focuses on applying moral concerns to huge data sets and network analysis to networks.	Richardson <i>et al.</i> (2020)
DrugRep https://cao.labshare.cn/drugrep/	Online system that uses AI and machine learning to identify targets, repurpose drugs, and practice personalized medicine. Includes target prediction, repurposing, and personalized medicine features.	Gan <i>et al.</i> (2023)
AtomNet PoseRanker (Atomwise) https://www.atomwise.com/	A graph convolutional neural network developed to identify and rerank crystal-like ligand poses for structure-based virtual screening and drug discovery.	Stafford <i>et al.</i> (2022)
BioSolveIT https://www.biosolveit.de/	AI-powered software platform for drug creation, molecular docking, and virtual screening. provides extensive features beyond Visual Studio and an intuitive user interface.	Rarey <i>et al.</i> (1996)
Cloudscreen (Cloudpharm) https://cloudpharm.eu/	This cloud-based medication discovery and repurposing tool provides AI-powered virtual evaluation and lead optimization services. Easy-to-use display, online access, and lead optimization tools.	Matsoukas <i>et al.</i> (2023)

DL Assisted Virtual Screening

Deep learning (DL), a powerful area of machine learning, uses multi-layered neural networks to analyze hard data, especially big databases of chemical structure and interaction between proteins (Herráiz-Gil *et al.*, 2025). DL is very good at analyzing images (such protein structures) and predicting drug-target binding with convolutional neural networks (CNNs). Even though DL's accuracy rises with data collection, issues like interpretability and computational cost must be resolved. Recent research shows that AI-assisted virtual screening is effective in drug repurposing, effectively discovering potential treatments for conditions including Parkinson's and pancreatic cancer. Large-scale virtual screening is also improved by AI-driven pipelines, which increase the effectiveness and precision of locating possible repurposed medications (Qiu and Cheng, 2024). A selection of such automated platforms is summarized in Table 2.

Integration of Multi-Omics Data

To comprehend human health, multi-omics data (DNA, transcript genome, proteins, and metabolism) must be studied, and gene sequencing has made this data accessible to everyone. Combining this data with clinical data offers a comprehensive, system-level understanding of disease and biological processes. Sequential or simultaneous multi-omics integration clarifies molecular interactions and the information flow from genotype to phenotype (Huang *et al.*, 2018). This method improves predictive accuracy and illness prognostics, which eventually results in better preventative and treatment plans. According to recent research, combining multi-omics datasets improves our understanding of biology. Combining ChIP-Seq and RNA-Seq data, for instance, demonstrated how histone changes unique to cancer affect gene expression in head and neck cancer (Jamal *et al.*, 2015). The influence of chromosomal amplifications on mRNA and protein levels was highlighted by the identification of important driver genes in colorectal cancer by the integration of proteomics with genomics and transcriptomics (Selvaraj *et al.*, 2021). Additionally, sphingosine was discovered as a possible biomarker and disturbed signaling pathways were identified as prospective therapeutic targets through the integration of metabolomics and transcriptomics in prostate cancer research (Singh, 2024). Information from many biological steps, like the DNA, proteome, metabolome, epigenome, & others like the lipidome and phosphoproteome, is included in

multi-omics data. When these datasets are examined collectively, the continuity of biological information is shown, so underlying mechanisms are clarified. Multi-omics patient datasets are accessible through publicly accessible resources (World Health Organization, 2021).

Machine Learning

When we're looking for new ways to use old drugs, machine learning is really important. To start, researchers will usually use some classic methods, things like Logistic Regression, Random Forest, Support Vector Machines, K-Nearest Neighbors, and RotatE. However, over the last few decades, deep learning has looked as an essential tool for signaling and identifying drugs that can be repurposed. Deep learning uses techniques like Recurrent and Convolutional Neural Networks, along with Transformers, for complex data analysis (Hasin *et al.*, 2017). These advanced techniques are particularly good at extracting meaningful features from molecular data and mapping these structures into spaces where potential therapeutic uses can be identified. Additionally, flow-based models are gaining interest because they can effectively modify the distribution of features, enhancing the analysis process (Subramanian *et al.*, 2020).

One of among the most essential instruments of medicine repurposing presently is machine learning, especially in light of the COVID-19 pandemic, as scientists globally have turned to these methods to find effective treatments (Kristensen and Györfy, 2021). This highlights the need for more interpretable models and causal learning techniques to work alongside traditional drug discovery methods. Another important consideration is how we can better describe molecules and their changing shapes, which is vital for extracting meaningful features (Li *et al.*, 2023). By advancing machine learning techniques, we have the possibility to Boost up the drug discovery Workflow and make significant improvements human health like never before. Algorithms supporting learning by machine have become essential for world of drug discovery, helping researchers uncover new treatment options. Naive Bayesian (NB), Random Forest (RF), and Support Vector Machine (SVM) are important machine learning techniques for early-stage medication repurposing studies (Pushpakom *et al.*, 2019). While NB forecasts drug-disease relationships, RF efficiently manages complicated data to anticipate drug-target interactions. SVM predicts pharmacological action and categorizes substances. These fundamental approaches offer a solid starting point for contrasting more

sophisticated deep learning strategies. These tools are particularly useful in the early stages of research, as they allow scientists to analyze complex data and find hidden patterns that might lead to breakthroughs (Perdomo-Quinteiro and Belmonte-Hernández, 2024).

Deep Learning

These days, deep learning algorithms are leading the way in a number of scientific and technological fields. A major component of this movement is the revival of artificial neural networks (NNs), which date back to the 1950s and demonstrate how AI has been effectively incorporated with conventional techniques (Zhou *et al.*, 2018; Tanoli *et al.*, 2021). These algorithms advance fields like picture and speech recognition by enabling computers to comprehend complex data in a more nuanced manner. Deep learning methods are changing the game in the field of drug discovery. They aid in the discovery of attractive lead compounds, new targets, and therapeutic activity predictions (Jobin *et al.*, 2019). Convolutional neural networks (CNNs), recurrent neural networks (RNNs), and deep neural networks (DNNs) are the three primary neural network types utilized in this field. DNNs stand out in particular because they analyze data in a simple manner, going from input to output while using examples to learn and generate precise results. Deep learning can address difficult problems like the creation of new chemical compounds or the prediction of their properties by utilizing these cutting-edge techniques. One of the most advanced methods in drug development is the Quantitative Structure-Activity Relationship (QSAR) model, which allows researchers to investigate the connections between chemical structures and biological activity (Chen *et al.*, 2020). They can even investigate how these molecules' shapes affect their interactions with their targets using 3D-QSAR. Although these methods have advanced drug discovery, it's crucial to keep in mind that they aren't flawless and have drawbacks of their own. However, there is no denying their influence in the field (Chen *et al.*, 2020)

Natural Language Processing (NLP)

The field of creating machines that can control human language, or data that mimics human language, in its written, spoken, and organized forms is known as natural language processing, or NLP. Instead of creating theoretical frameworks, natural language processing (NLP) is an engineering

discipline that aims to create technology that can do practical tasks. It developed from computational linguistics, which employs computer science to study the principles of language (Love *et al.*, 2025). Natural language generation (NLG), which focuses on text production by a machine, and natural language understanding (NLU), which focuses on semantic analysis or figuring out the intended meaning of text, are two overlapping subfields of natural language processing (NLP). With applications involving tasks like translation, summarization, text synthesis, and sentiment analysis, natural language processing (NLP) is one of the rapidly expanding research areas in artificial intelligence (Love *et al.*, 2025). An increasing variety of internal (such as identifying insurance fraud, assessing consumer sentiment, and improving aircraft maintenance) and user-facing (such as Google Translate) applications are powered by natural language processing (NLP). Future NLP professionals can start by learning the fundamentals of AI, which include basic math operations, Python coding, and the use of algorithms such as logistic regression, decision trees, and Naive Bayes. You can strengthen your foundation by taking online classes (Love *et al.*, 2025). As you go into more specialized subjects, they can also be beneficial. Working knowledge of neural networks, frameworks like PyTorch and TensorFlow, and other data preprocessing methods are necessary for NLP specialization. One particularly significant architecture is the transformer architecture, which has transformed the field since its introduction in 2017 (Love *et al.*, 2025). NLP is a fascinating and fulfilling field with the potential to significantly improve the world in a number of ways. Furthermore, deception can be produced using advanced language models. Large model training results in significant greenhouse gas emissions, which is a bigger worry (Secăreanu *et al.*, 2024).

Text Mining and Information Extraction:

- NLP techniques like named entity recognition (NER) identify key entities in literature, such as drug names, diseases, genes, proteins, and molecular pathways. For example, NER can extract mentions of a drug like "metformin" and associate it with diseases like "diabetes" or "cancer."
- Relation extraction identifies relationships between entities, such as drug-disease associations (e.g., "Drug X treats Disease Y") or drug-target interactions (e.g., "Drug Z inhibits Protein A").
- Tools like PubTator or BioBERT (a biomedical-specific NLP model) are used to extract structured data from unstructured texts in PubMed abstracts or full-text articles (Secăreanu *et al.*, 2024).

Knowledge Graph Construction:

- NLP helps build knowledge graphs by integrating extracted entities and relationships into a structured network. For instance, a knowledge graph might connect drugs, targets, and diseases based on literature evidence, revealing potential repurposing opportunities (e.g., a drug used for one disease may target a pathway relevant to another).
- These graphs enable researchers to identify novel drug-disease or drug-target pairs by traversing relationships not explicitly stated in a single paper (Secăreanu *et al.*, 2024).

Literature-Based Discovery (LBD):

- NLP supports LBD by identifying implicit connections across literature. For example, if Drug A is linked to Protein B in one study, and Protein B is linked to Disease C in another, NLP can infer a potential Drug A-Disease C connection for further investigation.
- Techniques like co-occurrence analysis or semantic similarity help uncover these hidden relationships (Wu *et al.*, 2021).

Sentiment and Contextual Analysis:

- NLP analyzes the context of drug mentions to determine their efficacy, side effects, or mechanisms of action. For instance, sentiment analysis can assess whether a drug is described positively (e.g., effective) or negatively (e.g., toxic) in specific contexts.
- This helps prioritize drugs for repurposing based on favorable literature evidence (Secăreanu *et al.*, 2024).

Processing Heterogeneous Data Sources:

- NLP integrates information from diverse sources like clinical trial databases (e.g., ClinicalTrials.gov), patents, and electronic health records alongside scientific literature. This holistic approach identifies repurposing candidates by cross-referencing findings.
- For example, NLP can extract adverse event data from clinical reports to assess a drug's safety profile for a new indication.

Semantic Search and Querying:

- NLP-powered semantic search tools allow researchers to query literature with complex questions, like "Which drugs targeting inflammation are also associated with neurodegenerative diseases?" Models like BERT or SciBERT understand context and retrieve relevant articles efficiently.

Automation and Scalability:

- Manual literature review is time-consuming and infeasible for millions of articles. NLP automates the process, enabling large-scale analysis of PubMed, PMC, or other repositories to identify repurposing candidates quickly.
- Machine learning models fine-tuned on biomedical texts (e.g., BioBERT, PubMedBERT) enhance accuracy in extracting domain-specific insights.

Examples of Applications

- **Drug-Disease Repurposing:** NLP identified thalidomide's potential for multiple myeloma by analyzing literature on its immunomodulatory effects, originally studied for other conditions.
- **COVID-19 Drug Repurposing:** During the pandemic, NLP tools scanned literature to identify drugs like remdesivir or dexamethasone for repurposing based on their antiviral or anti-inflammatory properties.
- **Adverse Effect Analysis:** NLP extracts side-effect profiles from literature to assess whether a drug's safety profile supports its use in a new indication.

Tools and Models

- **Pre-trained Models:** BioBERT, PubMedBERT, or ClinicalBERT for biomedical text processing.
- **Frameworks:** SpaCy, NLTK, or Scikit-learn for custom NLP pipelines.
- **Databases:** PubMed, PMC, or DrugBank for literature and drug data.
- **Platforms:** Systems like DRKG (Drug Repurposing Knowledge Graph) integrate NLP-extracted data for repurposing (Secăreanu *et al.*, 2024).

Applications and Case Studies (Real-World Impact)**Oncology**

In recent years, numerous companies have substantially ramped up their research and development efforts focused on anti-cancer drugs. This increase has resulted in a higher number of extensive, long-term clinical trials, leading to new opportunities for treating more cancer patients. A recent report from the American Cancer Society highlighted a promising statistic: the three-year survival rate for lung cancer has improved from 21% to nearly 31% between 2014 and 2021

(Wu *et al.*, 2021). This advancement is primarily attributed to the success of targeted therapies and immunotherapies, which are currently being investigated for various solid tumors. By investing more in these strategies, we can unlock the full potential of precision medicine, ultimately enhancing the long-term survival rates for those fighting cancer. Journey of developing new anti-cancer drugs is quite complex. Finding appropriate targets, investigating possible candidates, honing those leads, and finally choosing medication candidates for preclinical and clinical development are the steps in the process. Anti-cancer medication design is challenging because of issues like cancer spread, tumor variations, chemotherapy resistance, and hard-to-target proteins. Conventional approaches frequently fall short, resulting in treatments that are useless. Because of this urgency, new drug design techniques are needed to reduce prices and trial durations, which will eventually improve health and life expectancy. Drug development requires the use of Computer-Aided Drug Design (CADD), which dates back to the early 1980s. It speeds up and saves time and money by assisting scientists in the testing of medication interactions and the making of

novel compounds through simulations. Advanced computers and AI are transforming the medication design process. They make it possible to analyze biological data and medication interactions more quickly and accurately, predicting the efficacy of compounds before expensive studies are conducted (Benjamins *et al.*, 2020). Deep learning is great for designing new drugs from the ground up, and reinforcement learning lets scientists gradually build complex chemical structures, like the O-C-O example, which is really useful (Figure 2) (Benjamins *et al.*, 2020).

Neurodegenerative Diseases

Millions of older Americans suffer with Alzheimer's disease (AD), a neurological condition whose prevalence is predicted to increase dramatically by 2050 (Bakkar *et al.*, 2018). Effective medicines are desperately needed, but growing prices and a high clinical trial attrition rate continue to be major obstacles. Donnan demonstrated promise in decreasing cognitive loss in early AD, while lecanemab (Leqembi) recently became the first FDA-approved disease-modifying medication. However, its use can be restricted due to the monthly infusion requirement and related

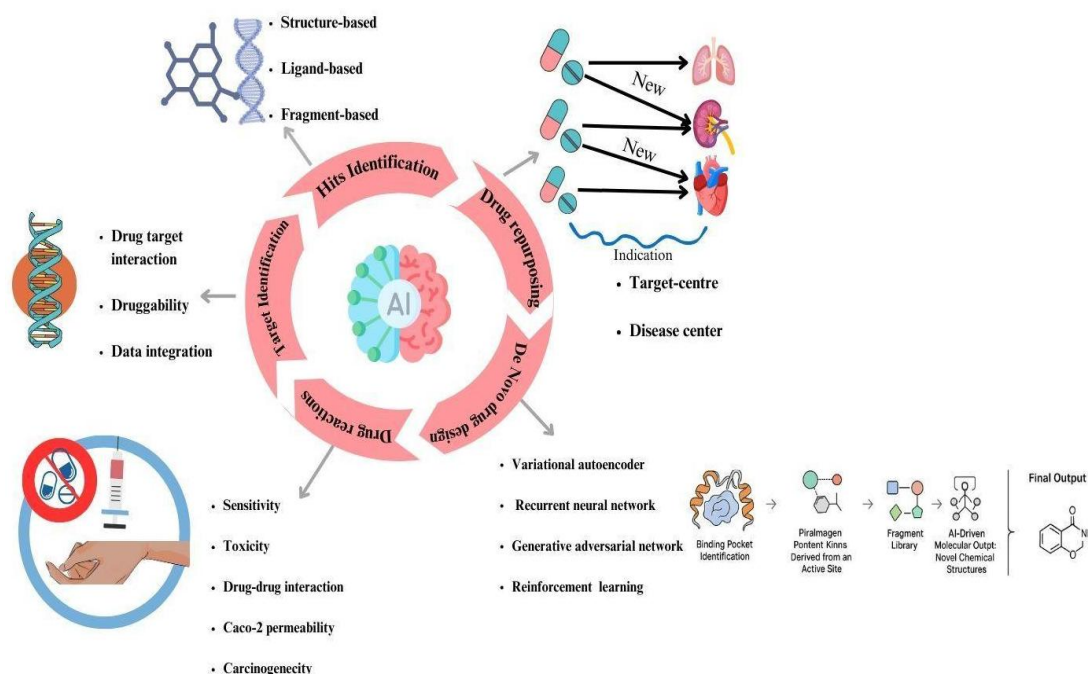


Figure 2. Schematic representation of the AI-driven drug repurposing workflow. The process begins with heterogeneous data acquisition from biomedical literature, omics databases, and clinical records, followed by natural-language processing (NLP)-based text mining, knowledge-graph construction, and machine-learning prediction to reveal novel drug-disease associations

safety issues. Comprehending the pathophysiology of AD is essential for creating novel treatments. The amyloid theory, which might not address late-stage AD, has received the most attention. Therefore, research on amyloid-independent targets and reasonably priced oral medicines is crucial (Bakkar *et al.*, 2018).

Infectious Diseases

The goal of the artificial intelligence (AI) pioneers of the 1950s was to create robots that were able to sense, reason, and think like people. General AI is a proof-of-concept for this aim. One significant advances in AI have been made possible by the creation of sophisticated algorithms, an unparalleled amount of data, and the quick expansion of processing power and memory capacity (Chen *et al.*, 2016). Applications of AI are found in many different domains, including digital pathology, data analysis, speech recognition, computer vision, and natural language understanding. Similarly, by uncovering hidden patterns and evidence in biomedical data, AI has been transforming drug discovery. AI employed by start-ups or pharmaceutical corporations for medication development and discovery.

AI in Drug Repurposing for Infectious Diseases

Drug repurposing involves identifying new uses for existing approved drugs, bypassing early-stage development to expedite responses to outbreaks. NLP and ML extract insights from vast biomedical corpora (e.g., PubMed, clinical trials), while graph neural networks model drug-target-disease interactions.

Specific Example: IBM's Watson Health; The Watson Health unit at IBM developed Watson for Drug Discovery, a cloud-based platform that applied NLP, machine-learning and data-integration to biomedical literature, patent filings and proprietary datasets, enabling researchers to surface novel drug-target and drug-disease hypotheses. For example, in collaboration with Pfizer (from 2016) the tool was used to assist immuno-oncology target discovery (Singhal *et al.*, 1999).

However, by 2019 IBM announced it would cease new sales of the WDD product and shift its focus towards clinical development tools rather than early-stage discovery, citing limited commercial uptake.

Even though WDD did not become a large-scale commercial repurposing engine in the way originally envisioned, its ambitions and prototypes helped shape the field of AI-driven repurposing—the

notion of mining vast, heterogeneous biomedical data to find hidden connections remains influential (Mrowietz and Asadullah, 2005).

AI in Vaccine Development

Vaccine development for infectious diseases traditionally spans 10–15 years, but AI compresses this by predicting immunogenic epitopes, simulating immune responses, and optimizing trial designs. Reverse vaccinology (RV) uses ML to screen pathogen genomes for antigens, while generative AI designs mRNA sequences (Boolell *et al.*, 1996).

- **Epitope Prediction and Antigen Design:**

Tools like Vaxign-ML employ random forests and convolutional neural networks (CNNs) to identify protective antigens from bacterial/viral proteomes. For SARS-CoV-2, AI predicted T-cell epitopes across HLA alleles, enabling universal vaccine blueprints that cover diverse populations. In HIV research, ML models forecast vaccine responses by mapping genotype-phenotype links, improving efficacy predictions by 20–30%.

- **mRNA Vaccine Acceleration:**

During COVID-19, AstraZeneca integrated AI for biomarker detection and thermostability optimization in its Covishield vaccine, analyzing genomic data 30% faster than manual methods. Recent 2024 studies used generative adversarial networks (GANs) for dengue and influenza vaccines, designing peptide-based candidates with high thermostability.

- **Clinical Trial Optimization:**

AI predicts trial outcomes and recruits diverse cohorts. For RSV vaccines, ML stratified risk groups, reducing failure rates in Phase III trials by identifying non-responders early.

Challenges include data scarcity for rare pathogens, addressed by transfer learning from related viruses (e.g., SARS-CoV to SARS-CoV-2) (Boolell *et al.*, 1996).

AI in Outbreak Prediction and Surveillance

Outbreak prediction integrates multi-modal data (e.g., social media, mobility, genomics) via ML to forecast spread, enabling proactive interventions. Early-warning systems use NLP for real-time signal detection.

- **Epidemiological Modeling:**

Hybrid AI-SIR (Susceptible-Infectious-Recovered) models predict trajectories with 85–95% accuracy. BlueDot, a Canadian AI platform, flagged COVID-19 nine days before WHO by analyzing news, flights, and animal disease reports using NLP and ensemble ML. For influenza, HealthMap scans global web

sources to detect anomalies, integrating with genomic data for variant tracking.

- **Genomic Surveillance:** AI analyzes sequencing data for mutations and AMR signatures. During moped outbreaks (2022–2024), graph neural networks traced transmission routes across 100+ countries, informing contact tracing. Metabiota's AI quantified Ebola impacts, combining environmental and social signals for rural predictions.

- **Resource Allocation:** In 2023 Taiwan studies, decision-tree ML from national health databases predicted influenza-like illness severity, aiding ventilator distribution during surges.

Limitations involve bias in underreported regions, mitigated by federated learning across global datasets.

AI in Diagnosis and Prognosis

Diagnosis leverages imaging AI and symptom triage for rapid, point-of-care detection, while prognosis models stratify risks for personalized care.

- **Imaging and Pathogen Identification:** CNNs diagnose tuberculosis from chest X-rays with 95% sensitivity, outperforming radiologists in low-resource settings. For COVID-19, hybrid CNN-RNN models analyzed CT scans to predict severity, identifying ground-glass opacities linked to poor outcomes. In HIV, ML from routine bloodwork detects early neuropathy.

- **Symptom-Based Triage:** XGBoost classifiers distinguish influenza from RSV using EHRs, reducing misdiagnosis by 40%. During COVID-19, random forests processed lab variables for sepsis prediction in infected patients.

- **Genomic and AMR Detection:** Support vector machines (SVMs) compare sequences to databases, identifying AMR in real-time for tailored antibiotics. For emerging threats like multidrug-resistant bacteria, AI flags resistance genes from metagenomic samples (Boolell *et al.*, 1996).

Individualized Medications

Individualized medicine, sometimes referred to as precision medicine, customizes medical care by considering the unique biological, physiological in nature, ecological, and behaviour features that define every patient. Rather than employing a one-size-fits-all strategy, this approach seeks to customize illness prevention and treatment. With the help of technological advancements in wireless health monitoring, imaging, and DNA proteomics, customized medicine has accelerated since the end of human DNA sequencing in 2003 (Kappos *et al.*, 2008). The field has enormous

potential to improve healthcare results and quality of life, despite current obstacles taking the requirement for human cell culture models and customized treatments. By evaluating intricate datasets to forecast medication reactions unique to each patient's genetic composition, pharmacogenomics' integration of artificial intelligence (AI) seeks to enhance treatment outcomes. This individualized strategy can improve the efficacy of medications, lessen side effects, and open the door for the creation of novel drugs (Hays *et al.*, 2001; Lin *et al.*, 2025). All things considered, the combination of AI with pharmacogenomics marks a substantial breakthrough in customized medicine that improves patient care while lowering medical expenses. Pharmacogenomics is undergoing a revolution thanks to AI and machine learning, which are handling the growing amount of high-throughput data. Clinical applications depend on these technologies' ability to recognize patterns and forecast the pharmacological characteristics of medication targets. They use big data analytics to speed up drug discovery by fusing pharmacokinetics and genetics (Singh and Dhiman, 2025).

Personalized medicine is aided by applications such as virtual chemical screening, in silico medication repositioning, and the identification of risk genes linked to certain diseases. From early drug discovery to clinical application, advanced AI approaches are crucial across the pharmacology continuum, surpassing conventional statistical methods. They make it possible to thoroughly investigate the relationships between illnesses, genes, and medications, which advances our knowledge of personalized treatment and enhances patient outcomes (Supriyono *et al.*, 2024; Zhou *et al.*, 2025).

Challenges and Limitations (Critical Evaluation)

AI has grown into a vital tool in the battle against infectious diseases, particularly in the area of drug development, where it supports synthetic drug evaluation and molecular layout. Quantity and quality data present difficulties, though, and can provide skewed findings. Data curation, teamwork for data sending, and creating AI systems that work well with short experimental data are all necessary to address these problems. Interpretability may be hampered by the intricacy of certain AI models, making it challenging for researchers to comprehend the reasoning behind results. This makes determining risks and evaluating pharmacological targets difficult. The goal of

explainable AI (XAI) methods like SHAP and LIME is to increase openness in this area. The significance of interpretability in healthcare applications is highlighted by a suggested taxonomy of XAI techniques specifically designed for drug design (Velásquez *et al.*, 2024).

AI can greatly improve current procedures, but it cannot replace conventional drug discovery techniques. Effective communication amid AI systems with human expertise are necessary for well-coordinated integration, with researchers directing AI development and using the results to find and improve viable medication consumers. Kim (2020) offers thorough analysis of ‘AI’ in drug research, highlighting uses or lingering difficulties. The availability of appropriate data is a significant problem because AI techniques need vast amounts of reliable, consistent data (Maji *et al.*, 2023). Data limitations may have an impact on how accurate the results are. Since biased training data may provide unjust or inaccurate projections, addressing these ethical issues is essential for the correct use of AI in the development of novel therapies. Techniques like data augmentation can be filed to improve instructions of AI by augmenting existing datasets in order to overcome these challenges. By using explainable AI techniques, bias issues can be addressed and consumers can have a better understanding of prediction mechanisms (Maji *et al.*, 2023). It’s also critical to keep in mind that AI cannot take the role of conventional experimental techniques or human researchers’ knowledge. Rather, combining AI with these techniques can speed up the creation of new drugs and improve the drug discovery. More successful results in the field may result from combining human understanding with AI’s predictive capabilities. Strong data integration and sophisticated machine learning models are key advantages of using AI in drug repurposing. Accurate prediction-making requires combining many datasets. There are still issues, though, like the reliance on accurate and comprehensive data. Poor prediction accuracy can emerge from models that do not effectively depict biological processes due to incomplete data, which can be caused by gaps in clinical histories or restricted experimental records. Additionally, the generalizability of models across diverse groups may be jeopardized by biased or underrepresented data. Datasets with imbalanced samples present serious problems as well since algorithms may favor the majority class, raising the possibility of losing out on possible drug repurposing prospects (Maji *et al.*, 2023).

Careful data curation, the creation of algorithms that control imbalances, and making sure datasets are complete are crucial to overcoming these obstacles. By overcoming these obstacles, AI may be more successful in finding novel therapeutic uses. Finding novel applications for current medications requires combining information from pharmacological, clinical, genetic, and biological sources (Maji *et al.*, 2023; Velásquez *et al.*, 2024). The discovery process can be sped up with techniques like data imputation, strong algorithm development, and suitable evaluation measures. By encouraging a group approach to difficult problems, collaborative platforms that let researchers share data and approaches might strengthen this endeavour even more.

Conclusion

The integration of Artificial Intelligence (AI) into the pharmaceutical sector has revolutionized the landscape of drug discovery and development, particularly in the domain of drug repurposing. AI’s ability to efficiently process vast and complex biological datasets, identify hidden patterns, and predict drug-target interactions has accelerated the identification of new therapeutic applications for existing drugs. This approach not only reduces the time, cost, and risks associated with traditional drug discovery but also provides new avenues for addressing unmet medical needs, especially in rare and complex diseases (Maji *et al.*, 2023).

Furthermore, advanced AI methodologies such as machine learning, deep learning, and natural language processing have enhanced the accuracy of virtual screening, multi-omics data integration, and literature-based discovery, making drug development more data-driven and evidence-based.

The success stories of repurposed drugs like remdesivir and thalidomide underscore the transformative potential of combining AI with pharmacological innovation.

In essence, AI-powered drug repurposing represents a paradigm shift toward a more intelligent, efficient, and sustainable pharmaceutical future. By bridging the gap between computational science and clinical application, AI continues to pave the way for faster therapeutic breakthroughs, improved patient outcomes, and a more responsive global healthcare system (Maji *et al.*, 2023).

Funding

The authors received no specific financial support for the research, authorship, and/or publication of this article.

Author Contributions

Mahvish Akhlaq: Conceptualization, methodology, investigation, software, writing – original draft, and writing – review and editing.

Kuldeep Singh: Conceptualization, visualization, supervision, investigation, data curation, and validation.

Piyush Kumar: Visualization and supervision.

Arun Kumar: Supervision.

Shom Prakash Kushwaha: Supervision.

Suvaiv: Conceptualization and writing – review and editing.

Ethics Statement

Ethical approval was not required for this study because it did not involve human participants, animals, identifiable personal data, or materials requiring institutional ethical review.

Declaration of Generative AI Use

The authors declare that no generative AI tools were used in the preparation of this manuscript.

Conflict of Interest

The authors declare that they have no conflicts of interest related to this work.

Data Availability Statement

No new data were generated or analyzed in this study.

Acknowledgements

The authors sincerely acknowledge Integral University, Lucknow, for providing research facilities and assigning manuscript communication number IU/R&D/2025-MCN0004167 to this article.

References

- Aliper, A., Plis, S., Artemov, A., Ulloa, A., Mamoshina, P., & Zhavoronkov, A. (2016). Deep learning applications for predicting pharmacological properties of drugs and drug repurposing using transcriptomic data. *Molecular Pharmaceutics*, 13(7), 2524–2530. <https://doi.org/10.1021/acs.molpharmaceut.6b00248>
- Alkan, F., & Erten, C. (2017). RedNemo: Topology-based PPI network reconstruction via repeated diffusion with neighborhood modifications. *Bioinformatics*, 33(4), 537–544. <https://doi.org/10.1093/bioinformatics/btw655>
- Alzubaidi, L., Zhang, J., Humaidi, A. J., Al-Dujaili, A., Duan, Y., Al-Shamma, O., Santamaria, J., Fadhel, M. A., Al-Amidie, M., & Farhan, L. (2021). Review of deep learning: Concepts, CNN architectures, challenges, applications, future directions. *Journal of Big Data*, 8(1), Article 53. <https://doi.org/10.1186/s40537-021-00444-8>
- Ashburn, T. T., & Thor, K. B. (2004). Drug repositioning: identifying and developing new uses for existing drugs. *Nature Reviews Drug Discovery*, 3(7), 673–683. <https://doi.org/10.1038/nrd1468>
- Baker, R. E., Mahmud, A. S., Miller, I. F., Rajeev, M., Rasambainarivo, F., Rice, B. L., Takahashi, S., Tatem, A. J., Wagner, C. E., Wang, L.-F., Wesolowski, A., & Metcalf, C. J. E. (2022). Infectious disease in an era of global change. *Nature Reviews Microbiology*, 20(4), 193–205. <https://doi.org/10.1038/s41579-021-00639-z>
- Bakkar, N., Kovalik, T., Lorenzini, I., Spangler, S., Lacoste, A., Sponaugle, K., Ferrante, P., Argentinis, E., Sattler, R., & Bowser, R. (2018). Artificial intelligence in neurodegenerative disease research: Use of IBM Watson to identify additional RNA-binding proteins altered in amyotrophic lateral sclerosis. *Acta Neuropathologica*, 135(2), 227–247. <https://doi.org/10.1007/s00401-017-1785-8>
- Benjamens, S., Dhunnoo, P., & Meskó, B. (2020). The state of artificial intelligence-based FDA-approved medical devices and algorithms: An online database. *npj Digital Medicine*, 3, 118. <https://doi.org/10.1038/s41746-020-00324-0>
- Boolell, M., Gepi-Attee, S., Gingell, J. C., & Allen, M. J. (1996). Sildenafil, a novel effective oral therapy for male erectile dysfunction. *British Journal of Urology*, 78(2), 257–261. <https://doi.org/10.1046/j.1464-410X.1996.10220.x>
- Chen, X., Xie, H., Zou, D., & Hwang, G. J. (2020). Application and theory gaps during the rise of artificial intelligence in education. *Computers and Education: Artificial Intelligence*, 1, 100002. <https://doi.org/10.1016/j.caeai.2020.100002>
- Chen, Y., Argentinis, J. E., & Weber, G. (2016). IBM Watson: How cognitive computing can be applied to big data challenges in life sciences research. *Clinical Therapeutics*, 38(4), 688–701. <https://doi.org/10.1016/j.clinthera.2015.12.001>
- Chong, C. R., & Sullivan, D. J. (2007). New uses for old drugs. *Nature*, 448(7154), 645–646. <https://doi.org/10.1038/448645a>
- Crevier, D. (1993). *AI: The Tumultuous History Of The Search For Artificial Intelligence*. Basic Books.
- Gan, J. H., Liu, J. X., Liu, Y., Chen, S.-W., Dai, W.-T., Xiao, Z.-X., & Cao, Y. (2023). DrugRep: An automatic virtual screening server for drug repurposing. *Acta Pharmacologica Sinica*, 44(4), 884–892. <https://doi.org/10.1038/s41401-022-00996-2>

- Geldsetzer, P. (2020). Knowledge and perceptions of COVID-19 among the general public in the United States and the United Kingdom: A cross-sectional online survey. *Annals of Internal Medicine*, 173(2), 157–160. <https://doi.org/10.7326/M20-0912>
- Hameed, P. N., Verspoor, K., Kusljic, S., & Halgamuge, S. (2018). A two-tiered unsupervised clustering approach for drug repositioning through heterogeneous data integration. *BMC Bioinformatics*, 19, Article 129. <https://doi.org/10.1186/s12859-018-2123-4>
- Hao, Z. (2019). Deep learning review and discussion of its future development. *MATEC Web of Conferences*, 277, 02035. <https://doi.org/10.1051/mateconf/201927702035>
- Hasin, Y., Seldin, M., & Lusis, A. (2017). Multi-omics approaches to disease. *Genome Medicine*, 9, Article 25. <https://doi.org/10.1186/s13059-017-1215-1>
- Hawkins, P. C. D., Skillman, A. G., & Nicholls, A. (2007). Comparison of shape-matching and docking as virtual screening tools. *Journal of Medicinal Chemistry*, 50(1), 74–82. <https://doi.org/10.1021/jm0603365>
- Hays, J. T., Hurt, R. D., Rigotti, N. A., Niaura, R., Gonzales, D., Durcan, M. J., Sachs, D. P., Wolter, T. D., Buist, A. S., Johnston, J. A., & White, J. D. (2001). Sustained-release bupropion for pharmacologic relapse prevention after smoking cessation: A randomized, controlled trial. *Annals of Internal Medicine*, 135(6), 423–33. <https://doi.org/10.7326/0003-4819-135-6-200109180-00011>
- Herráiz-Gil, S., Nygren-Jiménez, E., Acosta-Alonso, D. N., León, C., & Guerrero-Aspizua, S. (2025). Artificial intelligence-based methods for drug repurposing and development in cancer. *Applied Sciences*, 15(5), 2798. <https://doi.org/10.3390/app15052798>
- Hu, B., Guo, H., Zhou, P., & Shi, Z. L. (2021). Characteristics of SARS-CoV-2 and COVID-19. *Nature Reviews Microbiology*, 19(3), 141–154. <https://doi.org/10.1038/s41579-020-00459-7>
- Huang, C., Clayton, E. A., Matyunina, L. V., McDonald, L. D., Benigno, B. B., Vannberg, F., & McDonald, J. F. (2018). Machine learning predicts individual cancer patient responses to therapeutic drugs with high accuracy. *Scientific Reports*, 8, 16444. <https://doi.org/10.1038/s41598-018-34753-5>
- Ioakeim-Skoufa, I., Tobajas-Ramos, N., Menditto, E., Aza-Pascual-Salcedo, M., Gimeno-Miguel, A., Orlando, V., González-Rubio, F., Fanlo-Villacampa, A., Lasala-Aza, C., Ostasz, E., & Vicente-Romero, J. (2023). Drug repurposing in oncology: A systematic review of randomized controlled clinical trials. *Cancers*, 15(11), 2972. <https://doi.org/10.3390/cancers15112972>
- Ivanenkov, Y. A., Polykovskiy, D., Bezrukov, D., Zagribelnyy, B., Aladinskiy, V., Kamya, P., Aliper, A., Ren, F., & Zhavoronkov, A. (2023). Chemistry42: An AI-driven platform for molecular design and optimization. *Journal of Chemical Information and Modeling*, 63(3), 769–779. <https://doi.org/10.1021/acs.jcim.2c01191>
- Jamal, S., Goyal, S., Shanker, A., & Grover, A. (2015). Checking the STEP-associated trafficking and internalization of glutamate receptors for reduced cognitive deficits: A machine learning approach-based cheminformatics study and its application for drug repurposing. *PLOS ONE*, 10(6), e0129370. <https://doi.org/10.1371/journal.pone.0129370>
- Jobin, A., Ienca, M., & Vayena, E. (2019). The global landscape of AI ethics guidelines. *Nature Machine Intelligence*, 1(9), 389–399. <https://doi.org/10.1038/s42256-019-0088-2>
- Jordan, M. I., & Mitchell, T. M. (2015). Machine learning: Trends, perspectives, and prospects. *Science*, 349(6245), 255–260. <https://doi.org/10.1126/science.aaa8415>
- Kappos, L., Gold, R., Miller, D. H., MacManus, D. G., Havrdova, E., Limmroth, V., Polman, C. H., Schmierer, K., Yousry, T. A., Yang, M., Eraksoy, M., Meluzinova, E., Rektor, I., Dawson, K. T., Sandrock, A. W., & O'Neill, G. N. (2008). Efficacy and safety of oral fumarate in patients with relapsing-remitting multiple sclerosis: A multicentre, randomised, double-blind, placebo-controlled phase IIb study. *Lancet*, 372(9648), 1463–1472. [https://doi.org/10.1016/S0140-6736\(08\)61619-0](https://doi.org/10.1016/S0140-6736(08)61619-0)
- Khurana, D., Koli, A., Khatter, K., & Singh, S. (2023). Natural language processing: State of the art, current trends and challenges. *Multimedia Tools and Applications*, 82, 3713–3744. <https://doi.org/10.1007/s11042-022-13428-4>
- Kim, H., Kim, E., Lee, I., Bac, 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.
- Korkmaz, S., Zararsiz, G., & Goksuluk, D. (2014). Drug/nondrug classification using support vector machines with various feature selection strategies. *Computer Methods and Programs in Biomedicine*, 117(2), 51–60. <https://doi.org/10.1016/j.cmpb.2014.08.009>
- Lee, J., Yoon, W., Kim, S., Kim, D., Kim, S., So, C. H., & Kang, J. (2020). BioBERT: A pre-trained biomedical language representation model for biomedical text mining. *Bioinformatics*, 36(4), 1234–1240. <https://doi.org/10.1093/bioinformatics/btz682>
- Li, C., Wood, J. C., Vu, A. H., Hamilton, J. P., Lopez, C. E. R., Payne, R. M. E., Guerrero, D. A. S., Gase, K., Yamamoto, K., Vaillancourt, B., Caputi, L., O'Connor, S. E., & Buell, C. R. (2023). Single-cell multi-omics in the medicinal plant *Catharanthus roseus*. *Nature Chemical Biology*, 19, 1031–1041. <https://doi.org/10.1038/s41589-023-01327-0>
- Lima, A. N., Philot, E. A., Trossini, G. H., Scott, L. P., Maltarollo, V. G., & Honorio, K. M. (2016). Use of machine learning approaches for novel drug discovery. *Expert Opinion on Drug Discovery*, 11(3), 225–239. <https://doi.org/10.1517/17460441.2016.1146250>
- Lin, M., Weis, J., Fattah, H. M. A., & Fan, J. (2025). Rare disease study identification (RDSI): A natural language processing assisted search and visualization tool. *Journal of Clinical and Translational Science*, 9(s1), 110. <https://doi.org/10.1017/cts.2024.984>
- Liu, Z., Chen, X., Carter, W., Moruf, A., Komatsu, T. E., Pahwa, S., Chan-Tack, K., Snyder, K., Petrick, N., Cha, K., Lal-Nag, M., Hatim, Q., Thakkar, S., Lin, Y., Huang, R., Wang, D., Patterson, T. A., & Tong, W. (2022). AI-powered drug repurposing for developing COVID-19 treatments. *Reference Module in Biomedical Sciences*. 2022:B978-0-12-824010-6.00005-8. <https://doi.org/10.1016/B978-0-12-824010-6.00005-8>
- Love, A. S., Niu, C., & Labay-Marquez, J. (2025). Artificial intelligence in public health education: Navigating ethical challenges and empowering the next generation of professionals. *Health Promotion Practice*. <https://doi.org/10.1177/15248399251320989>
- Lv, H., Shi, L., Berkenpas, J. W., Dao, F.-Y., Zulfiqar, H., Ding, H., Zhang, Y., Yang, L., & Cao, R. (2021). Application of artificial intelligence and machine learning for COVID-19 drug discovery and vaccine design. *Briefings in Bioinformatics*, 22(6), bbab320. <https://doi.org/10.1093/bib/bbab320>
- Madugula, S. S., John, L., Nagamani, S., Gaur, A., Candasamy, S., Poroikov, V. V., & Sastry, G. N. (2021). Molecular descriptor analysis of approved drugs using unsupervised learning for drug repurposing. *Computers in Biology and Medicine*, 138, Article 104856. <https://doi.org/10.1016/j.compbiomed.2021.104856>

- Maji, S., Badavath, V. N., & Ganguly, S. (2023). Drug repurposing and computational drug discovery for viral infections and COVID-19. In: *Drug Repurposing and Computational Drug Discovery* (pp. 59–76). Apple Academic Press. <https://doi.org/10.1201/9781003347705-3>
- Matsoukas, M.-T., Panagiotopoulos, V., Ouzounis, S., Bafiti, V., Giatro, S. M., Kanterakis, A., Zoumpoulakis, P., & Katsila, T. (2023). Cloudscreen: A “one-stop-shop” platform for drug repurposing. REXPO23 Conference. <https://doi.org/10.58647/REXPO.23000029.v1>
- Menyhárt, O., & Györfy, B. (2021). *Multi-omics approaches in cancer research with applications in tumor subtyping, prognosis, and diagnosis*. Computational and Structural Biotechnology Journal, 19, 949–960.
- Mrowietz, U., & Asadullah, K. (2005). *Dimethylfumarate for psoriasis: More than a dietary curiosity*. Trends in Molecular Medicine, 11(1), 43–48.
- Mucke, H. A. M. (2015). A new journal for the drug repurposing community. *Drug Repurposing, Rescue, and Repositioning*, 1(1), 3–4. <https://doi.org/10.1089/drr.2014.0002>
- Oprea, T. I., & Mestres, J. (2012). *Drug repurposing: Far beyond new targets for old drugs*. The AAPS Journal, 14(4), 759–763.
- Perdomo-Quinteiro, P., & Belmonte-Hernández, A. (2024). Knowledge graphs for drug repurposing: A review of databases and methods. *Briefings in Bioinformatics*, 25(6), bbae461. <https://doi.org/10.1093/bib/bbae461>
- Pushpakom, S., Iorio, F., Ecker, G. F., Campillos, M., Kuhn, M., & Mestres, J. (2019). Drug repurposing: Progress, challenges and recommendations. *Nature Reviews Drug Discovery*, 18(1), 41–58. <https://doi.org/10.1038/nrd.2018.168>
- Qiu, Y., & Cheng, F. (2024). Artificial intelligence for drug discovery and development in Alzheimer's disease. *Current Opinion in Structural Biology*, 85, 102776. <https://doi.org/10.1016/j.sbi.2024.102776>
- Rarey, M., Kramer, B., Lengauer, T., & Klebe, G. (1996). A fast flexible docking method using an incremental construction algorithm. *Journal of Molecular Biology*, 261(3), 470–489. <https://doi.org/10.1006/jmbi.1996.0477>
- Richardson, P., Griffin, I., Tucker, C., Smith, D., Oechsle, O., Phelan, A., Rawling, M., Savory, E., & Stebbing, J. (2020). Baricitinib as potential treatment for 2019-nCoV acute respiratory disease. *The Lancet*, 395(10223), e30–e31. [https://doi.org/10.1016/S0140-6736\(20\)30304-4](https://doi.org/10.1016/S0140-6736(20)30304-4)
- Rodriguez, S., Hug, C., Todorov, P., Moret, N., Boswell, S. A., Evans, K., Zhou, G., Johnson, N. T., Hyman, B. T., Sorger, P. K., Albers, M. W., & Sokolov, A. (2021). Machine learning identifies candidates for drug repurposing in Alzheimer's disease. *Nature Communications*, 12, Article 1033. <https://doi.org/10.1038/s41467-021-21330-0>
- Röösli, E., Rice, B., & Hernandez-Boussard, T. (2021). Bias at warp speed: How AI may contribute to the disparities gap in the time of COVID-19. *Journal of the American Medical Informatics Association*, 28(1), 190–192. <https://doi.org/10.1093/jamia/ocaa210>
- Russell, S. J., & Norvig, P. (1995). *Artificial intelligence: A modern approach*. Prentice Hall.
- Secăreanu, G., Buhuș, Ș., & Cioantă-Păcuraru, I.-M. (2024). Artificial intelligence in educational management: Evolution, contributors and future directions. In: *Proceedings of the International Management Conference*. <https://doi.org/10.24818/IMC/2024/05.10>
- Selvaraj, G., Kaliyamurthi, S., Pelsherbe, G. H., & Wei, D. (2021). Application of artificial intelligence in drug repurposing: A mini-review. *Current Chinese Science*, 1(3), 333–345. <https://doi.org/10.2174/2210298101666210204162006>
- Serafim, M. S. M., Gertrudes, J. C., Costa, D. M. A., Oliveira, P. R., Maltarollo, V. G., & Honorio, K. M. (2021). Knowing and combating the enemy: A brief review on SARS-CoV-2 and computational approaches applied to the discovery of drug candidates. *Bioscience Reports*, 41(3), BSR20202616. <https://doi.org/10.1042/BSR20202616>
- Singh, A. (2024). Artificial intelligence for drug repurposing against infectious diseases. *Artificial Intelligence Chemistry*, 2, 100071. <https://doi.org/10.1016/j.aichem.2024.100071>
- Singh, J., & Dhiman, G. (2025). A survey on artificial intelligence-powered drug discovery and development in real-life environments including neonatal therapeutics. *Journal of Neonatal Surgery*, 14(5S), 809–819. <https://doi.org/10.52783/jns.v14.2156>
- Singhal, S., Mehta, J., Desikan, R., Ayers, D., Roberson, P., Eddlemon, P., Munshi, N., Anaissie, E., Wilson, C., Dhodapkar, M., Zeldis, J., Siegel, D., Crowley, J., & Barlogie, B. (1999). Antitumor activity of thalidomide in refractory multiple myeloma. *New England Journal of Medicine*, 341(21), 1565–1571. <https://doi.org/10.1056/NEJM199911183412102>
- Stafford, K. A., Anderson, B. M., Sorenson, J., & van den Bedem, H. (2022). AtomNet PoseRanker: Enriching ligand pose quality for dynamic proteins in virtual high-throughput screens. *Journal of Chemical Information and Modeling*, 62(5), 1178–1189. <https://doi.org/10.1021/acs.jcim.1c01250>
- Subramanian, I., Verma, S., Kumar, S., Jere, A., & Anamika, K. (2020). Multi-omics data integration, interpretation, and its application. *Bioinformatics and Biology Insights*, 14, 1177932219899051. <https://doi.org/10.1177/1177932219899051>
- Supriyono, S., Wibawa, A. P., Suyono, S., & Kurniawan, F. (2024). Advancements in natural language processing: Implications, challenges, and future directions. *Telecommunications and Radio Engineering*, 16, 100173. <https://doi.org/10.1016/j.teler.2024.100173>
- Tanoli, Z., Seemab, U., Scherer, A., Wennerberg, K., Tang, J., & Vähä-Koskela, M. (2021). Exploration of databases and methods supporting drug repurposing: A comprehensive survey. *Briefings in Bioinformatics*, 22(2), 1656–1678. <https://doi.org/10.1093/bib/bbaa003>
- Turon, G., Hlozek, J., Woodland, J. G., Kumar, A., Chibale, K., & Duran-Frigola, M. (2023). First fully-automated AI/ML virtual screening cascade implemented at a drug discovery centre in Africa. *Nature Communications*, 14, 5736. <https://doi.org/10.1038/s41467-023-41512-2>
- Velásquez, P. A., Hernandez, J. C., Galeano, E., Hincapié-García, J., Rugeles, M. T., & Zapata-Builes, W. (2024). Effectiveness of drug repurposing and natural products against SARS-CoV-2: A comprehensive review. *Clinical Pharmacology: Advances and Applications*, 16, 1–25. <https://doi.org/10.2147/CPAA.S429064>
- Wang, J. (2020). Fast identification of possible drug treatment of coronavirus disease-19 (COVID-19) through computational drug repurposing study. *Journal of Chemical Information and Modeling*, 60(6), 3277–3286. <https://doi.org/10.1021/acs.jcim.0c00179>
- Wang, L., Song, Y., Wang, H., Zhang, X., Wang, M., He, J., Li, S., Zhang, L., Li, K., & Cao, L. (2023). Advances of artificial intelligence in anti-cancer drug design: A review of the past decade. *Pharmaceuticals*, 16(2), 253. <https://doi.org/10.3390/ph16020253>
- Wassermann, A. M., Geppert, H., & Bajorath, J. (2011). Application of support vector machine-based ranking strategies to search for target-selective compounds. In: Bajorath, J. (eds) *Chemoinformatics and Computational Chemical Biology. Methods in Molecular Biology*, 672, 517–530. https://doi.org/10.1007/978-1-60761-839-3_21

- Wei, C. H., Allot, A., Leaman, R., & Lu, Z. (2019). PubTator central: Automated concept annotation for biomedical full text articles. *Nucleic Acids Research*, 47(W1), W587-W593. <https://doi.org/10.1093/nar/gkz389>
- World Health Organization. (2021). Repurposing of medicines in oncology: The underrated champion of sustainable innovation [Policy brief].
- Wu, E., Wu, K., Daneshjou, R., Ouyang, D., Ho, D. E., & Zou, J. (2021). How medical AI devices are evaluated: Limitations and recommendations from an analysis of FDA approvals. *Nature Medicine*, 27(4), 582–584. <https://doi.org/10.1038/s41591-021-01312-x>
- Zhang, X. (2022). Decoding China's COVID-19 health code apps: The legal challenges. *Healthcare*, 10(8), 1479. <https://doi.org/10.3390/healthcare10081479>
- Zhou, X.-Q., Huang, S., Shi, X.-M., Liu, S., Zhang, W., Shi, L., Lv, M.-H., & Tang, X.-W. (2025). Global trends in artificial intelligence applications in liver disease over seventeen years. *World Journal of Hepatology*, 17(3), 10172. <https://doi.org/10.4254/wjh.v17.i3.101721>
- Zhou, Y., Hou, Y., Shen, J., Huang, Y., Martin, W., & Cheng, F. (2020). Network-based drug repurposing for novel coronavirus 2019-nCoV/SARS-CoV-2. *Cell Discovery*, 6, Article 14. <https://doi.org/10.1038/s41421-020-0153-3>
- Zhou, Y., Tian, X., Browning, B. L., & Browning, S. R. (2018). POPdemog: Visualizing population demographic history from simulation scripts. *Bioinformatics*, 34(16), 2854–2855. <https://doi.org/10.1093/bioinformatics/bty184>
- Zhou, Y., Wang, F., Tang, J., Nussinov, R., & Cheng, F. (2020). Artificial intelligence in COVID-19 drug repurposing. *The Lancet Digital Health*, 2(12), e613–e614. [https://doi.org/10.1016/S2589-7500\(20\)30192-8](https://doi.org/10.1016/S2589-7500(20)30192-8)