Artificial Intelligence: A New Era in Drug Discovery

Devendra Kumar Mishra, Himani Awasthi*

Amity Institute of Pharmacy Lucknow, Amity University Uttar Pradesh, Sector 125, Noida, 201313, India.

ABSTRACT

Artificial intelligence (AI) is a simulation of the process of human intelligence through computers. The process involves obtaining information, developing rules for using information, making possible or accurate conclusions, and self-correcting. The development of new drug residues begins when basic scientists learn about biological targets (receptor, enzyme, protein, and gene). These targets involve the biological processes that occur in patients with a disease. Drug discovery can be through target identification, target verification, lead identification, and effectiveness of lead. AI can offer revolutionary insights into medicine, through data from genetics, proteomics and other life sciences that advance the process of discovery and development. Artificial Intelligence (AI) has recently been developed as a fiery element in the medical care industry. AI has exciting potential for prosperity in the field of biopharmaceutical. The biopharmaceutical industry makes efforts to approach AI to improve drug discovery, reduce research and development costs, reduce the time and cost of early drug discovery, and support predicting potential risks/side effects in late trials that can be very useful in avoiding traumatic events in clinical trials and ultimately clinical trials. Usually, drug development takes five years to go to trial, but the AI drug takes just 12 months. The rapid growth in life sciences and machine learning algorithms has led to enormous statistical access to the growth of AI-based startups focused on drug innovation in recent years.

Key words: Artificial intelligence, drug discovery, biopharmaceutical, clinical phase.

INTRODUCTION

The term "artificial intelligence" was given by John McCarthy at the Dartmouth Convention in 1956 to describe "the science and engineering of intelligent machines" [1]. Artificial intelligence (AI) is the process of human intelligence in computers [2]. AI is a field dealing with the design and application of algorithms for analyzing, learning and interpreting data [3]. The process of AI involves obtaining information, developing rules for using information, approximate or accurate conclusions, and self-correction [2]. Common applications of AI methods include appropriate information selection, data modelling, classification and reorganization, optimization and predictability [3]. The discovery and development of completely new drugs are intended for those who have an activity that is different from the already approved ones and the clinical indicator that is not addressed by the approved drug [4]. Better treatments that make repetitive improvements in current medicines are important as they can provide benefits with existing medicines such as potency, safety, tolerance, or convenience, but they usually do not include deceptions of biologically targeted targets that are different from those directly affected by existing drugs [5]. Nowadays artificial intelligence is widely used in the healthcare system for the following purposes:

- Research
- Digital health monitoring and diagnostics
- Patient data & risk analysis
- Surgery
- Mental health
- Hospital Management
- Virtual assistant
- Drug discovery
- Wearable

PRINCIPLE OF AI

Artificial techniques and tools that offer traditional mathematical, statistical and veterinary techniques that are ineffective or
inefficient can provide solutions to life-science problems \cite{1}. The roles of AI machine learning, information management, and multi-agent systems can make a huge contribution to experimental execution \cite{4}. The fields of agencies, human-computer interaction, natural language processing, vision and syntactic can provide technical support for integrating and building human and robot capabilities \cite{5}. Machine learning provides the framework for available discovery documents and the prioritization of bioactive compounds for the desired pharmacological effects and their efficacy as drug-like leads \cite{3}. Now there are emerging areas of biological target identification and protein design application. In many machines learning approaches in molecular informatics, chemocentric methods have found widespread application \cite{6}.

**FUNCTION OF AI**

AI has exciting potential for prosperity in the field of biopharmaceutical. Current AI programs of top biopharmaceutical companies include:

- **Mobile platform for improving health outcomes**: The ability to recommend patients and improve patient outcomes through real-time data collection.

- **Personalized medicine**: Ability to evaluate large patient data to identify treatment options using a cloud-based system.

- **Acquisitions galore**: A new startup companies combine artificial intelligence and healthcare to feed the startup needs of large biotech firms.

- **Drug discovery**: Pharma companies in association with software companies are trying the most advanced technology for the cutting cost and extensive process of drug discovery \cite{7}.

**DRUG DEVELOPMENT PROCESS**

Drug discovery is a long and complex process that can be broadly divided into four main categories:

- Target selection and validation
- Compound screening and lead optimization
- Preclinical studies
- Clinical trials \cite{6}

First, it is necessary to identify the target of a particular disease. This requires cellular and genetic target evaluation, genetic and proteomic analysis, and bioinformatics assays. After that, the next step means identification, where computers are identified from libraries of molecules through using some of the methods such as chemical synthesis, high-throughput, and virtual screening. In silico studies are used in the iterative cycle to improve the functional properties of newly synthesized drug candidates in combination with structure-function and cellular functional assays. Subsequently, in vivo studies like pharmacokinetic investigations and toxicity tests are performed in animal models \cite{8}. Finally, a drug candidate, who has successfully completed all preclinical tests, will be given to patients in a clinical trial. This step is characterized by three stages that require the drug to pass through them in succession. Phase I, includes the safety assessments of the drug on the small number of subjects; Phase II, includes the efficacy assessments drug with a small number of people who are affected by the targeted disease; and Phase III, efficacy studies with a larger number of subjects. Once the safety and efficacy of a drug candidate are confirmed at clinical stages, the area is reviewed by organizations such as the FDA for approval and operation shown in Fig. 1. \cite{8,9}.

<table>
<thead>
<tr>
<th>DRUG DISCOVERY PROCESS</th>
<th>DRUG DESIGN TOPICS</th>
<th>AI MODELS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Target identification and study</td>
<td>Prediction of protein folding and Proton pump inhibitors (PPIs)</td>
<td>CNN: predicting the residue contact and FD/DCA: detecting druggable PPI sites</td>
</tr>
<tr>
<td>Hit discovery</td>
<td>Drug repurposing</td>
<td>Network pathology</td>
</tr>
<tr>
<td></td>
<td>Virtual screening</td>
<td>SVM, AAE</td>
</tr>
<tr>
<td></td>
<td>Activity scoring</td>
<td>SVM, RF, 3D graph CNN</td>
</tr>
<tr>
<td>Hit lead</td>
<td>QSAR</td>
<td>Traditional machine learning, DNN</td>
</tr>
<tr>
<td></td>
<td>De novo design</td>
<td>Deep reinforcement learning, VAE, AAE</td>
</tr>
<tr>
<td>Lead optimization</td>
<td>Evaluation of ADME/T properties</td>
<td>CNN, multitask neural network</td>
</tr>
</tbody>
</table>
AI HELPS IN DRUG DISCOVERY

The successful finding of the new drugs in the development process is difficult and mainly the most difficult part. Since the initial stages of drug discovery, AI has been integrated to develop completely new leading computers that exhibit the required function in silico. The biological activity performed over the production of chemically correct compounds was learned by the “computer chemist” from well-known useful computers with help of combining computational de novo design with AI. The prediction of possible synthetic pathways for drug-like molecules was possible through AI in drug development, and as well as the pharmacological properties of the molecules, protein characteristics including the efficacy. The AI also helps in the drug combination, drug–target association and drug repurposing in the development process. Deep Learning (DL) has shown tremendous success in proposing potential drug candidates and accurately assessing their symptoms and potential toxic risks. To circumvent past problems in drug development – like analyzing large datasets, complex computer testing while minimizing standard error, requires huge R&D costs and more than the US $ 55 billion and over a decade - is now possible by using AI techniques. With the use of AI technology, new studies can be conducted to help identify new drug targets, rational drug design and drug repurposing shown in Table 2.

Ai in Understanding the Pathway or Finding Molecular Targets

In drug development, AI has revolutionized methodologies for the diagnoses and the treatment of the disease. This is likely due to the combination of genomics data, biochemical characteristics and target tractability.

Prediction of protein folding from the sequence: There are various diseases are related to protein inactivation. Thus, by studying protein composition, the drug formation techniques are supported that it can be used to identify small molecules that are active in reaching protein targets. Due to the powerful ability of feature extraction, intensive learning techniques have recently been used to assess secondary structure, backbone torsion angle and remaining protein contacts. At present, it is still a distant goal to accurately predict the 3D properties of proteins, and a comprehensive study approach has shown good promise for accelerating development in this field.

Prediction of protein–protein interactions (PPIs): It is defined as the protein–protein binding sites composed of multiple residues. It could be a new class of drug targets unique to traditional pharmacological targets such as G-protein coupled receptors (GPCRs), ion channels, kinases, and nuclear receptors. To achieve the concept of drug design depending on the structure of the protein–protein complex, it is important to study the PPI interface. Deep learning algorithms can extract the most relevant sequence features to predict PPI encounters, indicating obvious improvements compared to other machine learning methods.

Physical and chemical properties: The early identification of molecules with physical or chemical

---

**Figure 1: Drug development process**

- **Initial Synthesis**
  - Preclinical testing, research and development (1-3 Years)
  - 30-days safety review
    - Short-term animal testing
  - Phase I
    - Clinical research and development (2-10 Years)
    - Long-term animal testing
  - Phase II
  - Phase III
  - NDA submission
  - Report of Advisory committee
  - NBA approval
  - Post-marketing surveillance
  - NDA review (2 months - 7 years)
properties in the drug discovery pipeline greatly reduces the risk of failure. Thus, AI helps identify binding sites and modulates the stable composition that provides stability \cite{25}.

**Ai in Finding the Hit-To-Lead Optimization**

The approach of AI in the discovery of small drug-like molecules is related to the use of chemical space \cite{8}. The chemical space provides a platform for the identification of new and high-quality organic molecules because it is possible to further synthesize potential organic molecules. Whereas, the ML techniques and model software also contribute to the identification of direct-specific molecules and molecular interactions in their targeting while enhancing safety and efficacy capabilities \cite{26}.

**Generative models for de novo design:** De novo drug design means designing new chemical entities to modify the target of interest \cite{25}. The traditional de novo method similar to the component-based method can produce new molecules from scratch. However, many of them are difficult to synthesize due to the complexity and impossibility of the molecular structure \cite{28}. Due to the strong generative and learning ability, deep learning methods have been used to automatically create new structures with certain desired characteristics \cite{22}.

**Virtual testing:** Virtual testing refers to be in early drug development process intended as the use of algorithm and software to obtain bioactive molecules (hits) from internal compound collections or commercial chemical libraries, providing a highly efficient method for detecting novel hits and filtering compounds with unfavourable scaffolding \cite{29}. It includes docking-based, pharmacophore-based similarity search and machine learning techniques \cite{22}.

**Activity scoring:** A major component of molecular docking is the scoring function, which is designed to evaluate the binding affinities of interest-like molecules toward the target of interest \cite{30}. Due to the strong ability of nonlinear map estimation, machine-based scores show better performance by extracting various features successfully, such as geometric features, chemical characteristics, and magnetic field characteristics \cite{31}.

**QSAR:** In the hit-to-lead optimization process, QSAR analysis can be used to find powerful leading compounds from a series of hits analogues by assessing the bioactivity of the analogues. QSAR mainly uses mathematical methods to study quantitative mapping between the structural or physicochemical properties of compounds and their associated biological activities \cite{32}. QSAR analysis mainly consists of data collection, selection and execution of molecular definitions, the development of statistical models, the evaluation and interpretation of models, and the use of models \cite{33}.

**Drug repurposing:** It is also called drug repositioning, and defined as the process to estimate or find a new approach of the approved drugs \cite{22}. The process of drug repurposing is more attractive and practical with the help of AI. The idea of using existing treatments for a new disease is advantageous because the new appropriate drug bypasses the Phase I trial which includes the toxicity studies and it goes directly to Phase II clinical trials with a different indication \cite{54}. Drug repositioning is possible to begin because most drugs may have multiple targets and targets may have multiple effects, leading to higher variability in the drug-disease relationship with drug disorders. AI applies to drug repositioning as it provides short association information to the target population \cite{35}.

**Predicting the Mode-of-Action of Compounds Using AI**

The major approach of the AI platform is to predict the on-and-off effects of the target and the in vivo safety profile of the compounds before they are developed extends to those involved in the drug development process - especially those working in medicinal chemistry. This platform is intended to reduce drug development time, R&D costs and attractiveness rates \cite{36}.

**AI in Population Selection for Clinical Trials**

An appropriate AI tool to aid in clinical trials should identify the disease in patients, identify genetic targets and evaluate the impact of the designed molecule as well as on and off-target effects \cite{37}. The development of AI methods for detecting and predicting disease-related biomarkers in humans allows the recruitment of a specific population of patients in Phase II and III clinical trials. AI predictive modelling is successful in clinical trials in selected patient populations \cite{38}.

**AI in Polypharmacology**

Currently, there is a deep understanding of the pathological processes in diseases at the molecular level thus the “one-disease-multi-target model” dominates the “one-disease-one-target model”. This one disease multi-targeting is called polypharmacology and hence this AI works well toward polypharmacology to better understand the desired target of diseases resulting in best results \cite{39}.

**Recent Advancement for AI in Drug Discovery**

“Bloomberg Technology” reported that Microsoft has developed a technology that is used to support doctors in finding the right cancer treatment. Microsoft is also working on a project called Hanover. The purpose of this machine is to remember the available information needed to treat cancer and thus help to predict the combination of drugs that will be effective in identifying each patient. One of these projects is based on the use of AI in the treatment of myeloid leukaemia \cite{7}.

For the first time in the history of artificial intelligence, it createda medicine intended to be used on humans for the treatment of the obsessive-compulsive disorder (OCD) and named the DSP-1181 by the “British start-up Exscientia and Japanese pharmaceutical firm Sumitomo Dainippon Pharma,” which is a long-acting, potent serotonin 5-HT1A receptor agonist. Now this new AI algorithm-based drug is ready to enter a phase I human clinical trial \cite{40}.

In February 2013, IBM announced the first commercial software for Watson software system which is to be used for decision-making for lung cancer management at the Memorial Sloan Kettering Cancer Center, New York City, in partnership with health insurance company WellPoint. In December 2016, IBM in partnership with Pfizer launched IBM Watson, a cloud-based drug discovery platform. It allows users to analyze personal data such as medical lab
CONCLUSION

The current research and development process includes drug identification, target verification, lead generation, lead optimization, preliminary research, and clinical study. To develop a new novel drug requires money and time both. Approximately it takes the amount of 2.558 billion USD and as well as the time interval of 10-15 years. However, after given the high investment, the success rate for a small molecule in the drug discovery and development process remains only 13%, with a high risk of failure ultimately. AI can enhance and speed up research and development efforts, reduce the time and cost of early detection of drugs, and justify the assessment of potential toxic risks/side effects in late trials that can be of great help in avoiding traumatic events in clinical trials. AI can offer revolutionary ideas for medication and therapies with data gained from genomics, proteins and other life sciences that can bring advances in the drug discovery and development process. Modern technological combines existing algorithms/artificial neural input strategies that provide exciting opportunities for major transformations for large biopharmaceutical industries in the coming years.

REFERENCES


