WWW.IJRPR.COM

International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Artificial Intelligence in Drug Discovery and Development

Shubham Bhusare*[1], Ms. Dipmala Ghorpade [2], Dr. Gajanan Sanap [3]

[1,2,3]**Department of pharmacy, LBYP College of Pharmacy, Pathri, India**

DOI : https://doi.org/10.55248/gengpi.6.0125.0307

ABSTRACT

Artificial Intelligence(AI) has lately started to gear- up its operation in colorful sectors of the society with the pharmaceutical assiduity as a frontrunner devisee. This review highlights the poignant use of AI in different areas of the pharmaceutical sectors viz., medicine discovery and development, medicine repurposing, perfecting pharmaceutical productivity, clinical trials, etc. to name a many, therefore reducing the mortal workload as well as achieving targets in a short period. The use of artificial intelligence(AI) has been adding in colorful sectors of society, particularly the pharmaceutical assiduity. In this review, we punctuate the use of AI in different sectors of the pharmaceutical assiduity, including medicine discovery and development, medicine repurposing, perfecting pharmaceutical productivity, and clinical trials, among others; similar use reduces the mortal workload as well as achieving targets in a short period of time.

Introduction

AI in drug discovery

AI can help with the drug development process's limitations, which include its time-consuming and costly nature and which can be caused by a lack of advanced technologies [1]. AI is able to distinguish between hit and lead chemicals, and offer a speedier confirmation of the therapeutic target as well as drug structure design optimization [2,3]. A project called "drug discovery" is driven by the circumstances in which an illness has no treatment or in which the treatment that is available has serious side effects or poor efficacy.

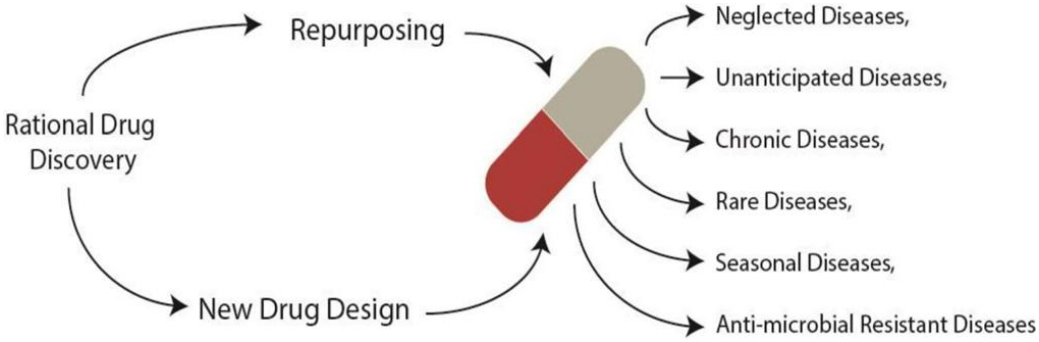

Fig1. The need for new drugs; they can be new drugs by rational drug development or repurposed drugs

The earliest step is to construct an underlying hypothesis, which involves target identification and target validation, that states that treatment effects for the disease are produced by activating or inhibiting a target (such as an enzyme, receptor, ion channel, etc.). Extensive assays will be conducted for the chosen target in order to find the hits and then the leads (i.e., drug candidates). This process will encompass lead optimization, hit-to-lead phase, and hit discovery. Following that, the medication candidates start preclinical research and clinical trials. Should the medication candidate prove effective, it might be introduced to the market as a medicinal product for treating the ailment. [4] Along with structure- and ligand-based approaches, a variety of in silico techniques for virtual screening compounds from virtual chemical spaces offer improved profile analysis, quicker elimination of nonlead compounds, and more cost-effective therapeutic molecule selection[2]. To choose a lead ingredient, drug design algorithms take into account the physical, chemical, and toxicological profiles, such as cou lomb matrices and molecular fingerprint recognition [5].

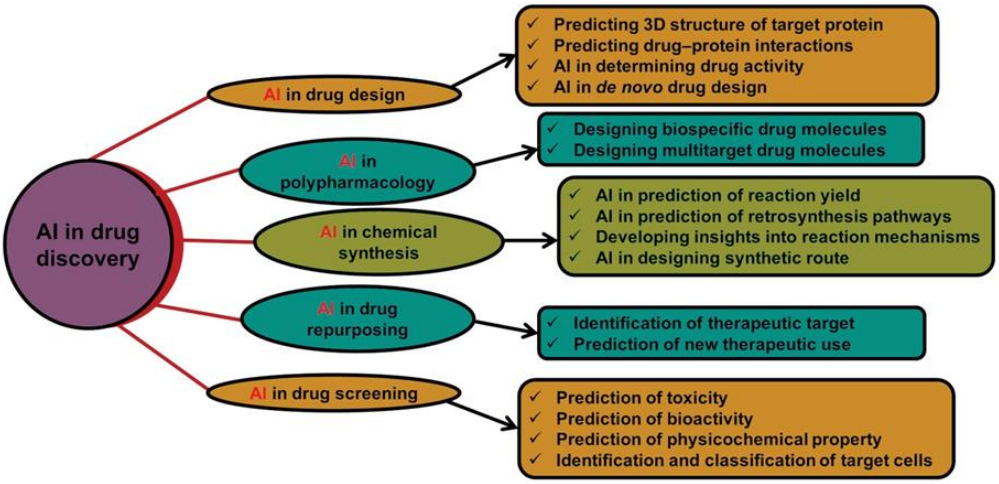

Fig2 shows how AI is being used in drug discovery.

To anticipate the intended chemical structure of a drug, a variety of factors can be used, including predictive models, molecular similarity, the molecule synthesis process, and the use of in silico techniques [3],[1]. When 95 000 decoys were tested against these receptors, the DeepVS method, which Pereira et al. demonstrated, performed exceptionally well for the docking of 40 receptors and 2950 ligands[6]. An alternative method evaluated the form similarity, pharmacological activity, and physicochemical characteristics of a cyclin-dependent kinase-2 inhibitor in order to optimize its potency profile using a multiobjective automated replacement algorithm [7].

THE IMPACT OF AI IN DRUG DISCOVERY: A NEW ERA UNFOLDS

Prediction of the physiochemical properties

The drug's physicochemical characteristics, such as its solubility, intrinsic permeability, degree of ionization, and partition coefficient (logP), indirectly influence its pharmacokinetics features and its target receptor family and so need to be taken into account when Formulating a Novel Drug[8] . There are various AI-based tools available. to forecast properties' physicochemical makeup. As an illustration, ML employs big data sets generated via earlier chemical optimization
to instruct the software [9]. Among the drug design algorithms are potential energy and molecular identifiers like SMILES strings coordinates, measurements, and the electron density surrounding the molecule of atoms in three dimensions, to create workable molecules using DNN and hence forecast its characteristics [10].

In order to ascertain the six physicochemical properties of environmental chemicals collected from the Environmental Protection Agency (EPA) and dubbed the Estimation Program Interface (EPI) Suite, Zang et al. developed a quantitative structure–property relationship (QSPR) methodology [9]. The lipophilicity and solubility of different substances have been predicted using neural networks based on the ADMET predictor and ALGOPS software [11]. DL techniques, like graph-based convolutional neural networks (CVNN) and undirected graph recursive neural networks, have been used to predict characteristics [10,12] by predicting a molecule's solubility.Using 745 compounds for training, Kumar et al. developed six predictive models [LDAs, SVMs, ANNs, knearest neighbor algorithms, probabilistic neural network algorithms, and partial least square (PLS)] that were then used to predict the

intestinal absorptivity of 497 compounds based on parameters such as molecular mass, total hydrogen count, molecular volume, logP, total polar surface area, the sum of E-states indices, solubility index (log S), and rotatable bonds [13]. In a similar vein, human intestinal absorption of various chemical substances was determined using RF and DNN-based in silico models [14]. As a result, AI plays a crucial role in drug development by helping to anticipate both the intended bioactivity and the desirable physicochemical features of the medicine.

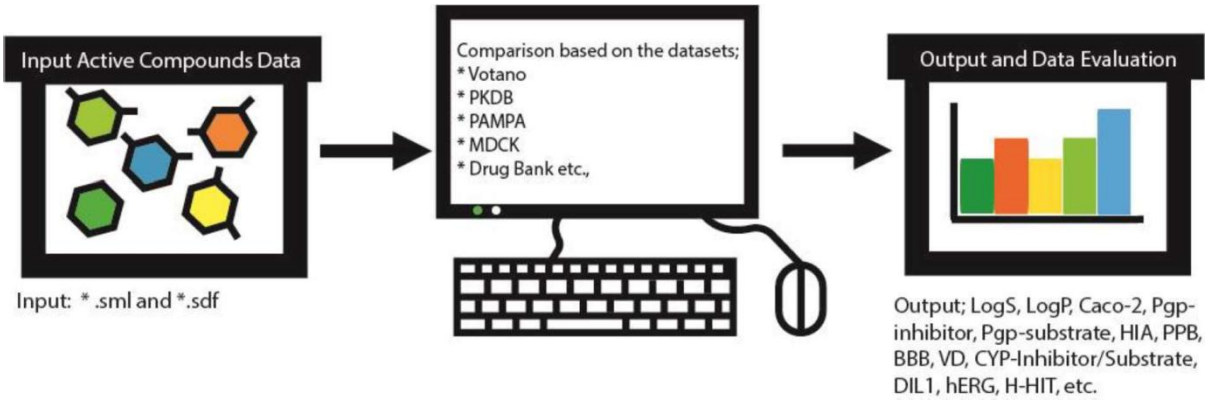

Fig 3: ML technique for predicting pharmacological and physicochemical features. After being converted to.sdf or.sml format, the structures developed molecular library is loaded into the computer. The machine is trained using a variety of data points from diverse sources, including PAMPA, Votano, DrugBank, and so on. Using encoded neural networks, the AI analyses the data and exports the physiological characteristics and ADMET as graphs and charts for comparative analysis .

Prediction of bioactivity

The effectiveness of medicinal compounds is determined by their affinity for the target protein or receptor.The therapeutic effect cannot be produced by drug molecules that do not interact with or have affinity for the targeted protein. It's also possible that produced medication molecules occasionally connect with unwanted proteins or receptors, which can be harmful. Therefore, the ability to anticipate drug-target interactions depends on drug target binding affinity (DTBA). AI-based techniques can calculate a drug's binding affinity by taking into account the characteristics or similarities between the drug and its target. Feature-based interactions identify the target and drug's chemical moieties in order to calculate the feature vectors. For predicting drug-target interactions, web applications like ChemMapper and the similarity ensemble approach (SEA) are available.DTBA has been ascertained using a variety of ML and DL techniques, including KronRLS, SimBoost, DeepDTA, and PADME. To calculate DTBA, machine learning (ML) techniques like Kronecker-regularized least squares (KronRLS) compare drug and protein molecule similarity. In a similar vein, SimBoost takes into account interactions based on similarity and features when predicting DTBA through regression trees. It is also possible to take into account drug features from SMILES, ligand maximum common substructure (LMCS), extended connectivity fingerprint, or a combination of these [14]. Since DL approaches use network-based techniques and do not rely on the availability of the 3D protein structure, they have demonstrated improved performance when compared to ML[15]. DL techniques such as DeepDTA, WideDTA, PADME, and DeepAffinity are employed to quantify DTBA. The amino acid sequence is entered for protein input data and for the 1D representation of the drug structure in DeepDTA, which accepts drug data in the form of SMILES.[16] The WideDTA CVNN DL method uses ligand SMILES (LS), amino acid sequences, LMCS, and protein domains and motifs as input data to determine the binding affinity [17].

Prediction of toxicity

To prevent harmful effects, it is essential to predict the toxicity of any drug molecule. To determine a compound's toxicity, animal studies are frequently conducted after preliminary research using cell-based in vitro assays, which drives up the cost of drug discovery. There are numerous web-based tools available to assist with cost reduction, including Toxtree, pkCSM, admetSAR, and LimTox [9]. Cutting-edge AI-based methods use input feature-based toxicity projections or search for commonalities between compounds. The US Food and Drug Administration (FDA), the Environmental Protection Agency (EPA), and the National Institutes of Health collaborated to organize the Tox21 Data Challenge, an effort to assess various computational methods for predicting the toxicity of 12,707 pharmaceuticals and environmental chemicals ; By identifying static and dynamic features within the chemical descriptors of the molecules, such as molecular weight (MW) and Van der Waals volume, an ML algorithm called DeepTox outperformed all other methods and could effectively predict the toxicity of a molecule based on predefined 2500 toxicophore features [18]. Table 1 is a list of the various AI tools used in drug discovery.

The safety target prediction of 656 marketed drugs was assessed using SEA against 73 unintended targets that could result in negative effects[15]. eToxPred, which was created using a machine learning approach, demonstrated accuracy of up to 72% when used to assess the toxicity and viability of small organic molecule synthesis [19]. Comparably, toxicity prediction also makes use of open-source tools like TargeTox and PrOCTOR[20]. TargeTox is a drug toxicity risk prediction method based on biological networks and targets. It applies the guilt-by-association principle, which states that entities in biological networks with similar functional properties share similarities [21].

Prediction of Carcinogenicity

To identify potential carcinogens in novel pharmaceuticals and evaluate the risk of environmental pollutants, carcinogenicity testing is necessary. Carcinogenic risk is commonly estimated by a two-year rat carcinogenicity study; however, this method is time- and resource-consuming. A DeepCarc model was developed by Li et al., (2021)[22] to forecast carcinogenicity. The DeepCarc model helps rank chemicals according to their likelihood of causing cancer and provides an early screening tool for assessing the potential carcinogenicity of new medications without the need for animal models. The liver cancer database of the National Centre for Toxicological Research provided 692 training set chemicals and 171 test chemicals that were used to develop the model. The effectiveness of Deepcarc was contrasted with that of other deep learning models, such as Chemistry Chainer Neural Fingerprint, Text Convolutional Neural Networks, and Convolutional Neural Network Fingerprint and "Multi-relational Graph Convolutional Networks with Edge Attention Based." With an accuracy of 0.74, specificity of 0.467, and sensitivity of 0.910, DeepCarc performed better than the other models.

Prediction of Ligand receptor binding affinity

The primary step in drug drug discovery is finding out the receptor related to disease and understanding its role in that particular disease mechanism. Based on the receptor's molecular nature, several chemical compounds with drug likeness are proposed , and the one with the highest receptor ligand binding score is assumed to be the hit. Various ML and DL-based artificial intelligence systems have been proposed to predict the binding score. SVR-Score and ID-Score are two models based on SVM. In order to investigate the receptor-ligand binding affinity, Ballester and Mitchell et al. developed a number of RF-based AIs, including RF-IChem, SFCscore RF, X-Score, and B2B Score. The large-scale protein-ligand docking website DockThor (https://www.dockthor.lncc.br/v2/) encodes RF-Score, which performed better among these. Once more, Ballester et al. verified that RF-Score-v3 outperformed X-Score when it came to 16 traditional scoring function sets. Decision trees (DTs) are used by RF as foundation learners. This makes the procedure more flexible and variable; the huge variance lowers the correlation between the trees. As a result, it raises the ensemble model's overall score prediction accuracy. In structure-based drug designing, the instruments used for molecular docking operate on varied sampling algorithms, docking, and simulation methodologies. To forecast the most precise binding score, these technologies additionally make use of a variety of scoring systems

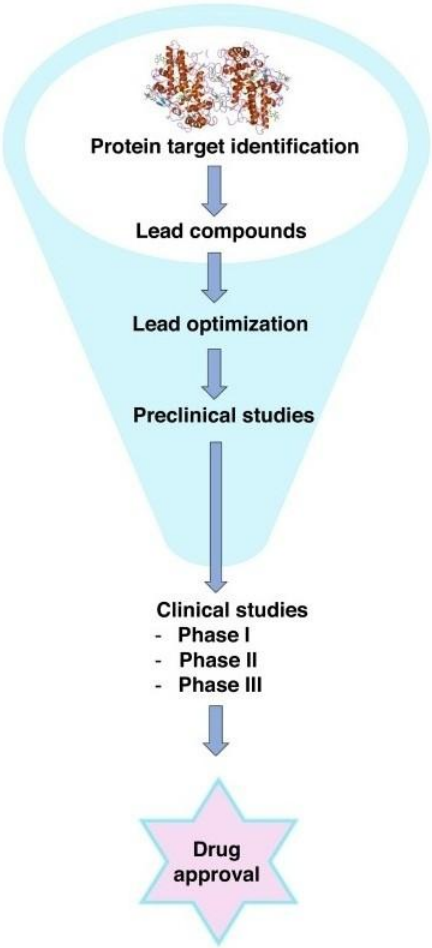

Fig.4 A flow chart of typical drug discovery

and characteristics. The techniques and operations rely on the ligand's three-dimensional structural characteristics, which are assessed by applying translational and rotational vectors [23]. The ANN, RF, or SVM-based machine learning methods previously discussed have a single drawback: they require fixed-length vectors to represent molecules. Deep Ducking is a unique deep learning framework for structure-based drug discovery that was introduced by Artem et al. in 2020. 1.36 billion molecules from the ZINC15 collection were docked against 12 target proteins using the Deep Ducking model and the FRED docking program; the outcomes were quite remarkable. It produced a decrease of 100-fold data without sacrificing any favorably docked entities, and an almost 6000-fold enrichment of high scoring molecules. The Deep Ducking method builds and trains the feed-forward DNN model based on QSAR properties and docking scores of subsets of a chemical library using the Keras Python package [23].

AI in de novo chemical synthesis

A novel drug's synthesis is invariably knowledge-driven. Drug development has become more logical thanks to computer-assisted synthesis planning (CASP), which eliminates the need for medicinal chemists to synthesize every molecule chosen after a thorough study of organic reaction pathways [23]. The following requirements must be met: (i) the molecular weight must be less than 500 Da; (ii) the number of H-bond donor atoms and acceptors must be fewer than ten; (iv) the octanol-water partition coefficient logP must be less than five [13,24]. Prior to initiating the synthesis, it is advisable to take into account additional factors such as the reaction yield with atom economy (AE), process mass intensity (PMI), and material costs [24]. Furthermore, it is important for the reaction reagents, catalysts, products, and byproducts to adhere to green synthesis guidelines and other safety criteria .

AI in preclinical and clinical trials

The synthesized drug has to pass pre-clinical studies in animals to enter the full passage of clinical trials . In phase I of the clinical trials, investigators use a small quantity of drugs on 20-80 healthy human volumteers (with no medical condition) for several months to study human pharmacology and evaluate ideal dosage. Phase II contains hundreds of infected volunteers (people with the ailment that the new treatment is supposed to treat) with the same dose for several years to examine the interaction and other therapeutic circumstances. Up to 3,000 randomly selected infected volunteers are monitored for several years during phase III. Phase III confirms the early phase's findings by a double-blind experiment in which neither the volunteers nor the observer knows which medicine they are taking. Here, the new medicine is approved in phase III; however, phase IV continues to monitor its safety and other therapeutic uses[24]. The failure rate of suggested medications in clinical trials is very high because of two factors: (i) ineffective volunteer selection; and (ii) ineffective observation monitoring[25]. To address these casualties in a clinical trial, ML and DL techniques have been proposed to set up the study, control necessary parameters, and continuously assess trial success rates. In Phase II/III trials, several AI technologies are employed to predict illness biomarkers relevant to humans in order to enlist a particular patient population [26,27]. The machine is designed in such a manner that it note down every change in the patient medical condition electronically. IBM Watson maintains and analyses patient electronic medical records, both structured and unstructured, using a DL-based clinical trial matching system to identify and construct appropriate patient profiles.[28] PrOCTOR forecasts the likelihood of toxicity. AiCure is a smartphone app that tracks phase II clinical trial data for individuals with schizophrenia; compared to conventional "modified directly observed therapy," monitoring data improved by 25%. [2]

AI in drug repurposing

A new, clever, and reasonable strategy in the rational drug discovery process is the repurposing of licensed medications and underdevelopment medications (failed projects); this is done to support cryptic therapeutic necessities of unanticipated, unusual, and disregarded disorders. Drug repurposing is effective because (i) several diseases share genetic components and biological pathways, and (ii) medicines have multiple targets. Repurposing involves a lot of data from diverse perspectives[29,30]. Three key algorithms—supervised learning, unsupervised learning, and semisupervised learning—are employed in drug repurposing investigations. Various models are utilised for drug repurposing, including the supervised model DTINet, the unsupervised model MANTRA, the semi-supervised model LapRLS and advanced NetLapRLS, LPMIHN, BLM with neighbor-based interaction-profile inferring (BLM-NII), the network consistency-based prediction (NetCBP) method, and network-based deepDR[31,32]. Even though these models have promised improved performance, no projections have been made as of yet[33].

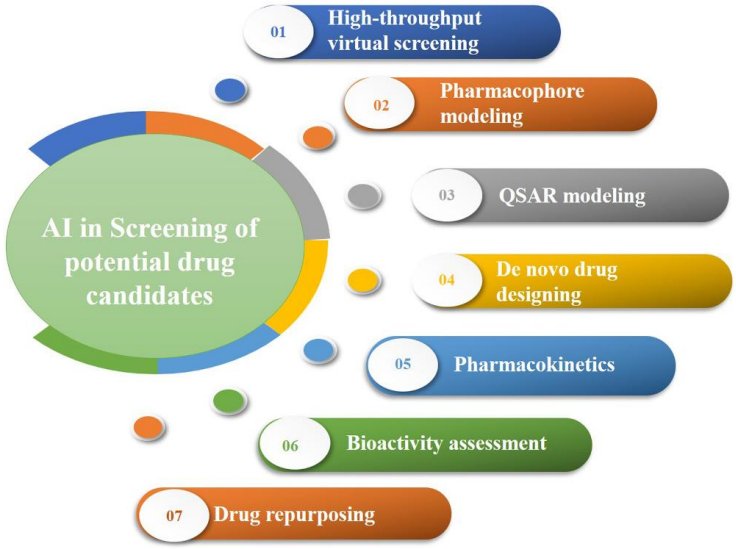

Fig .5 Application of AI in screening of potential drug candidate, and several approaches that can be selected for the identification of lead compound

AI in QSAR modeling

A computational method called quantitative structure-activity relationship (QSAR) modeling uses mathematical models to establish a connection between chemical structures, physical and chemical characteristics, and biological activities [34]. When identifying a variety of chemical structures from databases, mathematical models are useful. After selection, the lead molecule is subjected to wet laboratory tests. QSAR models are often regression and classification models. Gaussian processes (GPs) are a resilient type of regression model . AI-based QSAR techniques are the result of recent developments in QSAR technology. Decision trees, SVMs, RF, DL, gradient boosting methods (GBM), and linear discriminant analysis (LDA) are some of these techniques[35]. In order to create a model that combined multitask DNNs, GBMs, and gaussian process regression (GPR), Dahl et al., 2015 [36] were successful in winning the 2012 Merck Kaggle Molecular Activity Challenge. Additionally, a unique ANN-based technique dubbed Transformer-CNN based on SMILES chemical representation was created by Karpov et al., 2020[35]. The Cloud 3D-QSAR technique was used to generate monoamine oxidase B (MAO-B) inhibitors in a work by Jin et al., 2016[37]. Furthermore, Kim and Cho, 2018 used Jupyter Notebook and machine learning to create the PyQSAR algorithm, which is available as a stand-alone Python package. Geoffrey et al., 2020 [38] used PyQSAR to find possible therapeutic targets for new coronavirus. Zuvela et al., 2018 created an ANN-based QSAR that forecasts the antioxidant properties of flavonoids. Ding et al., 2020 [38] created VISAR, another online application that uses the DNN QSAR approach to distinguish chemical characteristics. Optimizing several parameters is the foundation of the drug development process [24]. Numerous studies have demonstrated the efficacy of multitask QSAR models in the identification of drug candidates with a wide range of biological activity [24]. Tenorio-Borroto et al., 2013 used artificial neural networks (ANN) to create a multitask QSAR model that successfully classified data from multiplexed assays with an astounding 92% precision. Research on the impact of many tasks in the multitask paradigm was conducted by Ramsundar et al., 2015. Once over 300 targets and 1.6 million molecules were taken into account, they were compiled and classified as extended-connectivity fingerprints (ECFP). Compared to RF, logistic regression, and single-task neural networks, the multitask network fared better. Multitask-based learning models outperformed single-task learning-based models, according to Zhao et al., 2020 [39]. Consequently, correlated data may be used more effectively by multitask learning when the entire library is not thoroughly filtered while taking into account all endpoints of interest[34].

AI in drug Synergism/Antagonism prediction

Drug combination effects fall into two categories: antagonistic and synergistic effects. While the latter lessens the efficiency of medications, the former can overcome primary and secondary drug resistance and is useful in the treatment of bacterial infections,[40] AIDS, and malignancie There are an enormous number of conceivable drug combinations due to the growing number of pharmaceuticals. As a result, studying the effects of medication combinations experimentally is expensive and time-consuming. The development of AI tools has allowed for the more effective and economical exploration of potential medication combinations. A Bayesian network model for investigating and evaluating medication combinations was presented by Li et al. in 2015.[41] A random forest-based model for forecasting compound synergism from chemical-genetic interactions was created by Wildenhain et al. the same year.[42] Preuer et al. recently proposed DeepSynergy,[42] a deep learning model for predicting anticancer medication synergism. DeepSynergy received as inputs the genetic and chemical makeup of illnesses as well as medicine formulations, which were subsequently transmitted throughout the network to the output device.

AI in Pharmacophore modeling

pharmacophore is a molecular structure that contains the fundamental characteristics that enable a chemical to exert its biological function. The purpose of pharmacophore models is to advance our understanding of ligand-protein interactions. They find new compounds that are thought to be biologically active since they satisfy the pharmacophore requirements. Pharmacophores convert functional groups sharing the same interaction profile into conceptual characteristics of a molecule associated with a particular kind of non-bonding contact, like a π stacking interaction or an H-bond donor/acceptor interaction[43]. The initial stage in structure-based pharmacophore modeling involves choosing and prepping the target protein structure. Subsequently, the next step is to predict the location of the binding. Afterward, a thorough examination is conducted to determine the complementary chemical features and structural configurations of the amino acids in the binding site. Finally, the pharmacophore features obtained are fine-tuned using various tools and algorithms. Pharmacophore modeling and mapping can be carried out with the help of the user-friendly open-source tool called Enhanced Ligand Exploration and Interaction Recognition Algorithm (ELIXIR-A)[43]. This program, based on Python, is capable of refining pharmacophores in multiple ligands and receptors by utilizing a structural geometry technique to identify similarities between protein-ligand binding sites. It has the ability to import models produced by visual molecular dynamics (VMD) and manual coordinate input. ELIXIR-A is designed to create pharmacophores that can be easily utilized in the screening process.The tool Pharmacoprint [34] demonstrates the presence, variety, and relationships among the pharmacophore features of a molecule.. In comparison to other well-known molecular fingerprint tools, Pharmacoprint has shown superior performance. It employs machine learning algorithms such as SVM, NN, and logistic regression (LR) to generate a pharmacophoric features fingerprint classification and is widely used in cheminformatics.

Table.1 List of AI-based software for drug discovery.[10]

Tools	Details	Website URL	Refs
DeepNeuralNetQSAR	Python-based system driven by computational tools that aid detection of the molecular activity of compounds	https://github.com/Merck/Deep- NeuralNet-QSAR	34
DeepChem	MLP model that uses a python-based AI system to find a suitable candidate in drug discovery	https://github.com/deepchem/deepchem	1
ORGANIC	A molecular generation tool that helps to create molecules with desired properties https://github.com/aspuru-guzik-group/ORGANIC	https://github.com/aspuru-guzik- group/ORGANIC	1
PotentialNet	Uses NNs to predict binding affinity of ligands https://pubs.acs.org/doi/full/ 10.1021/acscentsci.8b00507	https://pubs.acs.org/doi/full/ 10.1021/acscentsci.8b00507	1
Hit Dexter	ML technique to predict molecules that might respond to biochemical assays hitdexter2.zbh.uni-hamburg.de/	http://hitdexter2.zbh.uni-hamburg.de	1
DeltaVina	A scoring function for rescoring drug–ligand binding affinity https://github.com/ chengwang88/deltavina	https://github.com/chengwang88/ deltavina	1
Neural graph finger- print	Helps to predict properties of novel molecules https://github.com/HIPS/neural- fingerprint	https://github.com/HIPS/neural- fingerprint	1
DeepTox	Software that predicts the toxicity of total of 12 000 drugs www.bioinf.jku.at/ research/DeepTox	www.bioinf.jku.at/research/Deep- Tox	34
AlphaFold	Predicts 3D structures of proteins https://deepmind.com/blog/alphafold	https://deepmind.com/blog/alpha- fold	1

Chemputer	Helps to report procedure for chemical synthesis in standardized format https:// zenodo.org/record/1481731	https://zenodo.org/record/ 1481731	[1]

AI in Pharmacovigilance

The monitoring, identification, and avoidance of medication side effects are the main concerns of pharmacovigilance. Owing to the exponential growth of health records from clinical datasets to genetic databases, this industry needs cutting-edge and creative applications of AI techniques, such as deep learning and machine learning. To create new chemical molecules with desirable properties, generative adversarial networks are employed. The FDA recently released regulatory guidelines [34] for the safe use of AI-powered medical devices. Artificial Intelligence has a significant influence on the advancement of medication safety specialists. Professionals in drug safety evaluate and review patient safety data to protect the general public's health from harmful drug effects. These experts work with the extraction of medical records from individual case safety reports (ISCRs) in relation to the processing of medication side effects. The application of AI, particularly machine learning techniques, has demonstrated a useful function in leveraging the effectiveness of learning new skills and proficiency for drug safety experts. ML techniques enhance decision-making procedures to enable more accurate and quick processing and evaluation of drug safety-related issues. Consequently, using AI in pharmacovigilance may enhance operational tasks in handling huge datasets and their statistical analysis to further access drug toxicity and safety [34,44].

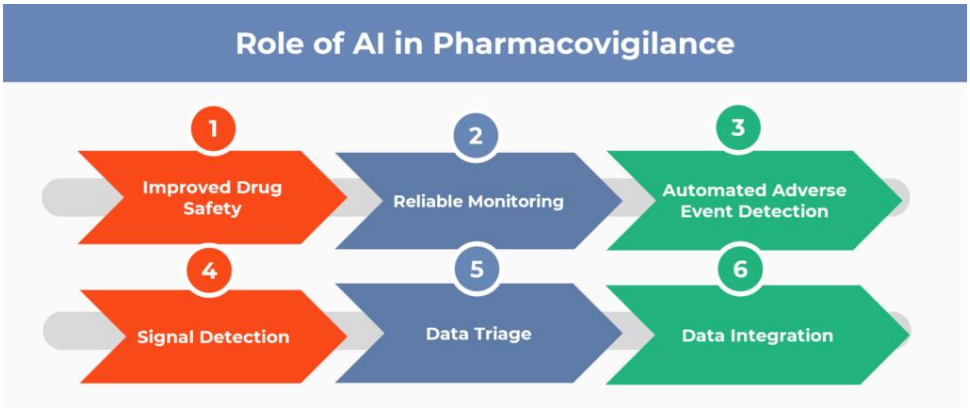

Fig.6 Shows Role AI in Pharmacovigilance

Challenges And Limitation of AI in Drug Discovery

Even though AI holds great promise to revolutionize the drug discovery process, there are still a number of extremely difficult obstacles that need to be overcome before its full potential can be achieved. Securing data quality and accessibility is a major concern. Because AI models are data-driven, the volume and variety of the training data determines how effective the models will be. Obtaining high-quality biological data is challenging because of data dispersion among multiple organizations and privacy constraints. Furthermore, gathering the required data can be costly and time-consuming, particularly for smaller research groups. Access to comprehensive and diverse datasets is thus made possible through cooperation and data-sharing initiatives. [45] There are also important constraints related to data bias and generalizability. When trained on skewed data, AI models may generate predictions that are not accurate. These biases may result from differences between healthcare professionals, regional discrepancies in data sources, or the underrepresentation of particular groups in clinical trials. Furthermore, false positives or ineffective medication candidates may be identified as a result of overfitting, which is defined as a model's proficiency with training data but difficulty with unseen data.[46]

When training AI models, researchers can use bias correction approaches to lessen the impact of biases on model outputs. For instance, data bias in an AI-powered drug discovery study can be addressed by applying the SMOTE (Synthetic Minority Oversampling Technique) bias correction technique.[46] SMOTE balances the dataset and reduces the impact of bias by creating artificial data points for underrepresented groups. There is ongoing research on bias correction approaches, but no one-size-fits-all solution has been found. Nevertheless, researchers can lessen the influence of data bias in AI applications by using careful dataset selection, processing, and bias correction strategies. Significant obstacles may include processing power and resource intensity, particularly for deep learning models. Smaller pharmaceutical businesses and academic research teams with low funding may find it difficult to use these models since they demand significant computer resources for both training and inference. Reduced computational costs and improved accessibility are achieved through the use of cloud-based AI services and partnerships with AI technology providers. Furthermore, two essential processes that AI models in drug research must go through are regulatory approval and validation. Getting regulatory approval and establishing trust in the pharmaceutical industry require proving the safety, efficacy, and reproducibility of AI-generated results. To create validation protocols and recommendations, cooperation between pharmaceutical companies, regulatory authorities, and AI researchers is essential.[9]

Conclusions

Artificial intelligence (AI) has proven to have a substantial impact on improving the quality and efficacy of therapeutic interventions, and its integration into drug development and discovery has marked a key milestone in the pharmaceutical sector. Through a wide range of applications, artificial intelligence (AI) has not only sped up the drug development process but also created new opportunities for target identification, medication repurposing, and the prediction of potential therapeutic uses. AI's critical role in repurposing expands its potential to challenge established drug discovery paradigms, making it a vital tool in the search for novel therapies. Artificial Intelligence (AI) is being used to improve drug development processes, as demonstrated by its careful drug design and virtual screening applications. Researchers can precisely evaluate medication candidates by using AI's computational power to locate and classify target cells. This computational efficiency demonstrates the transformative potential of AI in enhancing global healthcare outcomes in polypharmacology, chemical synthesis, and drug repurposing. AI-based models can mimic medication distribution and clearance in the body, predict pharmacokinetic parameters, and optimize drug dosage and administration methods. Animal research and human clinical trials may be avoided by using AI-based computational techniques for PBPK models, which can streamline the creation of such models and optimize their parameters. All things considered, the incorporation of AI technology has enormous potential to expedite medication discovery, enhance patient outcomes, and transform the pharmaceutical sector, hence spurring its advancement.

References:

1. Paul D, Sanap G, Shenoy S, Kalyane D, Kalia K, Tekade RK. Artificial intelligence in drug discovery and development. *Drug Discov Today*. 2021;26(1):80-93. doi:10.1016/j.drudis.2020.10.010

2. Mak KK, Pichika MR. Artificial intelligence in drug development: present status and future prospects. *Drug Discov Today*. 2019;24(3):773-780. doi:10.1016/j.drudis.2018.11.014

3. Sellwood MA, Ahmed M, Segler MHS, Brown N. Artificial intelligence in drug discovery. *Future Med Chem*. 2018;10(17):2025-2028. doi:10.4155/fmc-2018-0212

4. Hughes JP, Rees SS, Kalindjian SB, Philpott KL. Principles of early drug discovery. *Br J Pharmacol*. 2011;162(6):1239-1249. doi:10.1111/j.1476-5381.2010.01127.x

5. Chan HCS, Shan H, Dahoun T, Vogel H, Yuan S. Advancing Drug Discovery via Artificial Intelligence. *Trends Pharmacol Sci*. 2019;40(8):592-604. doi:10.1016/j.tips.2019.06.004

6. Shen C, Ding J, Wang Z, Cao D, Ding X, Hou T. From machine learning to deep learning: Advances in scoring functions for protein–ligand docking. *Wiley Interdiscip Rev Comput Mol Sci*. 2020;10(1). doi:10.1002/wcms.1429

7. Firth NC, Atrash B, Brown N, Blagg J. MOARF, an Integrated Workflow for Multiobjective Optimization: Implementation, Synthesis, and Biological Evaluation. *J Chem Inf Model*. 2015;55(6):1169-1180. doi:10.1021/acs.jcim.5b00073

8. Zang Q, Mansouri K, Williams AJ, et al. In Silico Prediction of Physicochemical Properties of Environmental Chemicals Using Molecular Fingerprints and Machine Learning. *J Chem Inf Model*. 2017;57(1):36-49. doi:10.1021/acs.jcim.6b00625

9. Yang X, Wang Y, Byrne R, Schneider G, Yang S. Concepts of Artificial Intelligence for Computer-Assisted Drug Discovery. *Chem Rev*. 2019;119(18):10520-10594. doi:10.1021/acs.chemrev.8b00728

10. Hessler G, Baringhaus KH. Artificial intelligence in drug design. *Molecules*. 2018;23(10). doi:10.3390/molecules23102520

11. Lusci A, Pollastri G, Baldi P. Deep architectures and deep learning in chemoinformatics: The prediction of aqueous solubility for drug-like molecules. *J Chem Inf Model*. 2013;53(7):1563-1575. doi:10.1021/ci400187y

12. Kumar R, Sharma A, Siddiqui MH, Tiwari RK. Prediction of Human Intestinal Absorption of Compounds Using Artificial Intelligence Techniques. *Curr Drug Discov Technol*. 2017;14(4). doi:10.2174/1570163814666170404160911

13. Chai S, Liu Q, Liang X, et al. A grand product design model for crystallization solvent design. *Comput Chem Eng*. 2020;135. doi:10.1016/j.compchemeng.2020.106764

14. Thafar M, Raies A Bin, Albaradei S, Essack M, Bajic VB. Comparison Study of Computational Prediction Tools for Drug-Target Binding Affinities. *Front Chem*. 2019;7. doi:10.3389/fchem.2019.00782

15. Lounkine E, Keiser MJ, Whitebread S, et al. Large-scale prediction and testing of drug activity on side-effect targets. *Nature*. 2012;486(7403):361-367. doi:10.1038/nature11159

16. Mahmud SMH, Chen W, Jahan H, Liu Y, Sujan NI, Ahmed S. IDTi-CSsmoteB: Identification of Drug-Target Interaction Based on Drug Chemical Structure and Protein Sequence Using XGBoost with Over-Sampling Technique SMOTE. *IEEE Access*. 2019;7:48699-48714. doi:10.1109/ACCESS.2019.2910277

17. Gao KY, Fokoue A, Luo H, Iyengar A, Dey S, Zhang P. *Interpretable Drug Target Prediction Using Deep Neural Representation*.; 2018. http://www.rdkit.org/

18. Mayr A, Klambauer G, Unterthiner T, Hochreiter S. DeepTox: Toxicity prediction using deep learning. *Front Environ Sci*. 2016;3(FEB). doi:10.3389/fenvs.2015.00080

19. Pu L, Naderi M, Liu T, Wu HC, Mukhopadhyay S, Brylinski M. EToxPred: A machine learning-based approach to estimate the toxicity of drug candidates 03 Chemical Sciences 0305 Organic Chemistry 03 Chemical Sciences 0304 Medicinal and Biomolecular Chemistry. *BMC Pharmacol Toxicol*. 2019;20(1). doi:10.1186/s40360-018-0282-6

20. Basile AO, Yahi A, Tatonetti NP. Artificial Intelligence for Drug Toxicity and Safety. *Trends Pharmacol Sci*. 2019;40(9):624-635. doi:10.1016/j.tips.2019.07.005

21. Lysenko A, Sharma A, Boroevich KA, Tsunoda T. An integrative machine learning approach for prediction of toxicity-related drug safety. *Life Sci Alliance*. 2018;1(6). doi:10.26508/lsa.201800098

22. Li T, Tong W, Roberts R, Liu Z, Thakkar S. DeepCarc: Deep Learning-Powered Carcinogenicity Prediction Using Model-Level Representation. *Front Artif Intell*. 2021;4. doi:10.3389/frai.2021.757780

23. Rai S, Raj U, Tichkule S, et al. Recent trends in In silico drug discovery. *Int J Comput Biol*. 2016;5(1):54-76. http://www.ijcb.in

24. Shultz MD. Two Decades under the Influence of the Rule of Five and the Changing Properties of Approved Oral Drugs. *J Med Chem*. 2019;62(4):1701-1714. doi:10.1021/acs.jmedchem.8b00686

25. Zhavoronkov A, Vanhaelen Q, Oprea TI. Will Artificial Intelligence for Drug Discovery Impact Clinical Pharmacology? *Clin Pharmacol Ther*. 2020;107(4):780-785. doi:10.1002/cpt.1795

26. Feijoo F, Palopoli M, Bernstein J, Siddiqui S, Albright TE. Key indicators of phase transition for clinical trials through machine learning. *Drug Discov Today*. 2020;25(2):414-421. doi:10.1016/j.drudis.2019.12.014

27. Ménard T, Koneswarakantha B, Rolo D, Barmaz Y, Popko L, Bowling R. Follow-Up on the Use of Machine Learning in Clinical Quality Assurance: Can We Detect Adverse Event Under-Reporting in Oncology Trials? *Drug Saf*. 2020;43(3):295-296. doi:10.1007/s40264-019-00894-3

28. Karlafti E, Anagnostis A, Kotzakioulafi E, et al. Does COVID-19 clinical status associate with outcome severity? An unsupervised machine learning approach for knowledge extraction. *J Pers Med*. 2021;11(12). doi:10.3390/jpm11121380

29. Levin JM, Oprea TI, Davidovich S, et al. Artificial intelligence, drug repurposing and peer review. *Nat Biotechnol*. 2020;38(10):1127-1131. doi:10.1038/s41587-020-0686-x

30. Yu L, Qiu W, Lin W, Cheng X, Xiao X, Dai J. HGDTI: predicting drug–target interaction by using information aggregation based on heterogeneous graph neural network. *BMC Bioinformatics*. 2022;23(1). doi:10.1186/s12859-022-04655-5

31. Luo Y, Zhao X, Zhou J, et al. A network integration approach for drug-target interaction prediction and computational drug repositioning from heterogeneous information. *Nat Commun*. 2017;8(1). doi:10.1038/s41467-017-00680-8

32. Zeng X, Zhu S, Liu X, Zhou Y, Nussinov R, Cheng F. DeepDR: A network-based deep learning approach to in silico drug repositioning. *Bioinformatics*. 2019;35(24):5191-5198. doi:10.1093/bioinformatics/btz418

33. Luo H, Li M, Yang M, Wu FX, Li Y, Wang J. Biomedical data and computational models for drug repositioning: A comprehensive review. *Brief Bioinform*. 2021;22(2):1604-1619. doi:10.1093/bib/bbz176

34. Singh S, Gupta H, Sharma P, Sahi S. Advances in Artificial Intelligence (AI)-assisted approaches in drug screening. *Artif Intell Chem*. 2024;2(1):100039. doi:10.1016/j.aichem.2023.100039

35. Karpov P, Godin G, Tetko I V. Transformer-CNN: Fast and Reliable tool for QSAR. Published online October 21, 2019. doi:10.1186/s13321-020-00423-w

36. Ma J, Sheridan RP, Liaw A, Dahl GE, Svetnik V. Deep neural nets as a method for quantitative structure-activity relationships. *J Chem Inf Model*. 2015;55(2):263-274. doi:10.1021/ci500747n

37. Reis J, Cagide F, Chavarria D, et al. Discovery of New Chemical Entities for Old Targets: Insights on the Lead Optimization of Chromone-Based Monoamine Oxidase B (MAO-B) Inhibitors. *J Med Chem*. 2016;59(12):5879-5893. doi:10.1021/acs.jmedchem.6b00527

38. Geoffrey A S B, Madaj R, Sanker A, et al. Automated In Silico Identification of Drug Candidates for Coronavirus Through a Novel Programmatic Tool and Extensive Computational (MD, DFT) Studies of Select Drug Candidates. Published online August 17, 2020. doi:10.26434/chemrxiv.12423638.v3

39. Zhao Z, Qin J, Gou Z, Zhang Y, Yang Y. Multi-task learning models for predicting active compounds. *J Biomed Inform*. 2020;108. doi:10.1016/j.jbi.2020.103484

40. Tamma PD, Cosgrove SE, Maragakis LL. Combination therapy for treatment of infections with gram-negative bacteria. *Clin Microbiol Rev*. 2012;25(3):450-470. doi:10.1128/CMR.05041-11

41. Li P, Huang C, Fu Y, et al. Large-scale exploration and analysis of drug combinations. *Bioinformatics*. 2015;31(12):2007-2016. doi:10.1093/bioinformatics/btv080

42. Wildenhain J, Spitzer M, Dolma S, et al. Prediction of Synergism from Chemical-Genetic Interactions by Machine Learning. *Cell Syst*. 2015;1(6):383-395. doi:10.1016/j.cels.2015.12.003

43. Tilahun Muhammed M, Aki-Yalcin E. Pharmacophore Modeling in Drug Discovery: Methodology and Current Status. 2021;8(3):749-762. doi:10.18596/jotcsa

44. Wu J, Zhang Q, Wu W, et al. WDL-RF: Predicting bioactivities of ligand molecules acting with G protein-coupled receptors by combining weighted deep learning and random forest. In: *Bioinformatics*. Vol 34. Oxford University Press; 2018:2271-2282. doi:10.1093/bioinformatics/bty070

45. Bannigan P, Aldeghi M, Bao Z, Häse F, Aspuru-Guzik A, Allen C. Machine learning directed drug formulation development. *Adv Drug Deliv Rev*. 2021;175. doi:10.1016/j.addr.2021.05.016

46. Jiménez-Luna J, Grisoni F, Schneider G. Drug discovery with explainable artificial intelligence. *Nat Mach Intell*. 2020;2(10):573-584. doi:10.1038/s42256-020-00236-4