Asian Journal of Pharmaceutical Education and Research
Vol -11, Issue-1, January- March 2022
ISSN:2278 7496

REVIEW ARTICLE Impact Factor: 7.014

REVIEW ON ARTIFICIAL INTELLIGENCE PARADIGM FOR DRUG DISCOVERY

Vivek Jain*, Sunil Kumar Jain

Adina Institute of Pharmaceutical Science, NH86A, Lahdara, Sagar, MP 470001

*Corresponding Author's E mail: drvivekjainn@gmail.com

Received 12 Nov. 2021; Revised 18 Nov. 2021; Accepted 28 Nov. 2021, Available online 10 Jan 2022.

Cite this article as: JainV, JainSK. Review on Artificial Intelligence Paradigm for Drug Discovery. Asian Journal of Pharmaceutical Education and Research. 2022; 11(1): 23-45.

https://dx.doi.org/10.38164/AJPER/10.1.2022.23-45

ABSTRACT

Drug discovery can be through target identification, target verification, lead identification, and effectiveness of lead. 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 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. In this review, we provide an overview of current AI technologies and offer a glimpse of how AI is reimagining drug discovery by highlighting examples where AI has made a real impact. Considering the excitement and hyperbole surrounding AI in drug discovery, we present a realistic view by discussing both opportunities and challenges in adopting AI in drug discovery. 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.

Keywords: Drug discovery,Artificial intelligence,Computers,Innovation.

INTRODUCTION

Drug discovery is a long, complex and high-risk process. It typically takes a staggering 10-15 years and costs up to US$2.8 billion to develop a new drug, while an astonishing proportion (80-90%) of them fail in the clinic with Phase II proof-of- concept (PoC) trials accounting for the most significant number of clinical failures [1]. Although the number of new molecular

entities (NMEs) approved by regulatory agencies, such as the US Food and Drug Administration (FDA), has increased over the past decade (2010-2020) compared with the prior decade, the cost of bringing a new drug to market has risen precipitously because of regulatory incentives and the advent of new modalities [2,3].The key drivers contributing to the increased cost of pharmaceutical innovation include investment lost from late-stage clinical attrition, an increasingly stringent regulatory system

that sets a high bar for approval, and higher clinical trial costs, especially for pivotal trials. Given these realities, pharmaceutical and biotech companies are incentivized to innovate and adopt new technologies to improve productivity, cut costs, and ensure sustainability. Over the past few years, there has been a drastic increase in data digitalization in the pharmaceutical sector. However, this digitalization comes with the challenge of acquiring, scrutinizing, and applying that knowledge to solve complex clinical problems [4]. This motivates the use of AI, because it can handle large volumes of data with enhanced automation[5]. AI is a technology based system involving various advanced tools and networks that can mimic human intelligence.

At the same time, it does not threaten to replace human physical presence[6,7]completely.

AI utilizes systems and software that can interpret and learn from the input data to make independent decisions for accomplishing specific objectives. According to the McKinsey Global Institute, the rapid advances in AI-guided automation will be likely to completely change the work culture of society[8, 9]. AI involves several method domains, such as reasoning, knowledge representation, solution search, and, among them, a fundamental paradigm of machine learning (ML). ML uses algorithms that can recognize patterns within a set of data that has been further classified. A subfield of the ML is deep learning (DL), which engages artificial neural networks (ANNs). These comprise a set of interconnected sophisticated computing elements involving perceptons analogous to human biological neurons, mimicking the transmission of electrical impulses in the human brain[10]. ANNs constitute a set of nodes, each receiving a separate input, ultimately converting them to output, singly or multilinked using algorithms to solve problems[11]. ANNs involve various types, including multilayer perceptron (MLP) networks, recurrent neural networks (RNNs), and convolutional neural networks (CNNs), which utilize either supervised or unsupervised training procedures[12, 13]. The MLP network has applications including pattern recognition, optimization aids, process identification, and controls, are usually trained by supervised training procedures operating in a single direction only, and can be used as universal pattern classifiers[14]. RNNs are networks with a closed-loop, having the capability to memorize and store information, such as Boltzmann constants and Hopfield networks[15,16]. CNNs are a series of dynamic systems with local connections, characterized by its topology, and have use in image and video processing, biological system modeling, processing complex brain functions, pattern recognition, and sophisticated signal processing[17]. The more complex forms include Kohonen networks, RBF networks, LVQ networks, counter-propagation networks, and ADALINE networks. Several tools have been developed based on the networks that form the core architecture of AI systems. One such tool developed using AI technology is the International Business Machine (IBM) Watson supercomputer (IBM, New York, USA). It was designed to assist in the analysis of a patient's medical information and

its correlation with a vast database, resulting in suggesting treatment strategies for cancer. This system can also be used for the rapid detection of diseases. This was demonstrated by its ability to detect breast cancer in only 60 s[18,19]. The AI technologies used today in drug discovery have evolved from earlier machine learning (ML) and cheminformatics concepts. For example, the application of ML to the development of quantitative structure activity relationship (QSAR) models and expert systems for toxicity prediction has a longstanding history[20,21]. The widespread adoption of these techniques witnessed in recent times has been fueled by the advent of big data, advanced analytics, GPU-accelerated computing, cloud processing, algorithm development, and the democratization of AI toolkits. There are opportunities to apply AI technologies across the drug discovery and development continuum, starting from target identification through to preclinical development. Evidence suggests that lack of clinical efficacy has been the foremost cause of attrition in clinical Phase II studies[22], highlighting that target selection remains one of the most crucial decision points in drug discovery. Given this reality, there is a desire to improve the target selection process by applying AI techniques. AI-driven discovery platforms can extract and synthesize target-relevant information from a large volume of complex, disparate multiomics data, providing a better understanding of target biology, uncovering disease target associations, and identifying targets with a strong link to a disease. Target DB is one such example that integrates publicly available data on a given target and uses an ML-based classification system to categorize target tractability[23]. The approach and scoring system used within Target DB provides useful criteria for ligand ability assessment and prioritization of drug targets for development. Once a target of interest has been identified and validated, the next stage in drug discovery is to identify high-quality chemical start points (hits) that bind to, and modulate, the target. Although there is a range of hit-finding methods available, virtual screening (VS) is a cost-effective and resource-sparing approach used to prioritize a subset of compounds for evaluation in a primary assay. The use of AI-driven approaches to improve the performance of VS is increasing[24]. AI-powered VS campaigns have identified novel chemical hits against seemingly difficult drug targets [25, 26, thereby turning undruggable targets into tractable drug targets. To ensure that quality hits worthy of further consideration are progressed, computational methods have been used to identify, prioritize and select hit compounds, a process referred to as hit triage.15 ML models are now being used to automate and improve the efficiency of the hit triage process[27]. Fast, accurate, and reliable prediction of binding free energies to enable VS and structure-based design remains a significant challenge, including rank-ordering of compounds from a VS. In recent years, ML-based scoring functions trained on databases of protein ligand complexes have shown great promise in improving hit rates during VS[28]. Unlike traditional scoring functions, ML-based scoring functions can implicitly account for binding

interactions that are difficult to model, and are not constrained to any predefined functional form. With the advent of 'make-on-demand' libraries and the screening collections breaking the billion compound barriers, conventional docking methods have become impractical. Active learning methods integrated with molecular docking offer an elegant solution for efficient exploration of the chemical space through iterative screening[29,30]. The lead optimization (LO) phase is the most expensive and time-consuming phase in preclinical drug discovery[31]. It is inherently a multiparameter optimization (MPO) problem [32], with the goal of identifying compounds with an optimal balance of drug like properties while maintaining sufficient potency. Hitting this sweet spot is a challenge, because it involves simultaneous optimization of multiple and often competing objectives, such as safety, specificity, efficacy and pharmacokinetics (PK) properties, while maintaining potency[33]. LO involves iterative rounds of the design–make-test-analyze (DMTA) cycle and shortening these cycle times is crucial for accelerating the LO process. Generative chemistry that relies on AI-guided generative modeling for compound design has demonstrated success in shortening cycle times and designing compounds that meet the defined LO criteria[34]. Generative modeling platforms also integrate various predictive models for absorption, distribution, metabolism, excretion, and toxicity (ADMET) endpoints to guide the design and selection of compounds with favorable properties that satisfy the defined LO criteria. In this way, generative chemistry can automate and shorten the design phase of the DMTA cycle and offset individual cognitive biases during molecule ideation. AI is also making headway in computer-aided synthesis planning (CASP), which is valuable in both hit identification and improving DMTA cycle efficiency[35]. AI-assisted synthesis planning helps chemists to objectively choose the most efficient and cost-effective synthetic route for a target molecule, thus accelerating the make phase of the DMTA cycle. Automated continuous flow chemical synthesis is another emerging technology poised to revolutionize organic synthesis[36]. This technology opens new avenues by integrating smart automation and intelligent synthesis, thereby enabling fully autonomous synthesis. Closing the loop of the DMTA cycle is the 'analysis' phase. To improve DMTA cycle efficiency, the data must be turned into knowledge to make better design suggestions for the next iteration. Given the sparse and non-uniform nature of the data encountered in drug discovery, the incorporation of sparse data AI methods, such as few-shot learning, for data analysis allows extracting valuable insights to inform the next round of the design cycle. Another practical application of AI is using deep imputation methods to handle the noisy, sparse, missing, and truncated data often generated in drug discovery[37-39]. Deep imputation methods combine DL and statistical imputation methods to learn correlations between experimental endpoints and gain valuable information, even from minimal experimental data, to more accurately fill in missing experimental values[40]. Such techniques can help establish assay correlations and build multi target

QSAR models, which can be used for in silico off-target profiling against protein target families, such as kinases. Translating preclinical discoveries into clinical practice in the form of new therapeutics is one of the biggest challenges in clinical development and, too often, clinical candidates are lost during translation. To bridge this translational gap, translational 1 strategies are increasingly being integrated as early as LO to improve Phase II and Phase III clinical success rates, more evident in oncology drug discovery programs[41]. To that end, the use of translational biomarkers that provide information on target modulation, target engagement, confirm proof of mechanism (POM), and for designing stratified clinical trials are used for derisking clinical development. The ability of AI techniques to learn hidden and meaningful patterns by integrating large amounts of heterogeneous and high-dimensional omics data sets provides valuable insights for translational biomarker discovery[42]. As innovations in AI technologies continue, the use of AI in drug discovery will also continue to grow.

AI in Drug Discovery

The vast chemical space, comprising >1060 molecules, fosters the development of a large number of drug molecules. However, the lack of advanced technologies limits the drug development process, making it a time-consuming and expensive task, which can be addressed by using AI. AI can recognize hit and lead compounds, and provide a quicker validation of the drug target and optimization of the drug structure design[43,44]. Despite its advantages, AI faces some significant data challenges, such as the scale, growth, diversity, and uncertainty of the data. The data sets available for drug development in pharmaceutical companies can involve millions of compounds, and traditional ML tools might not be able to deal with these types of data. Quantitative structure-activity relationship (QSAR)-based computational model can quickly predict large numbers of compounds or simple physicochemical parameters, such as log P or log D. However, these models are some way from the predictions of complex biological properties, such as the efficacy and adverse effects of compounds. In addition, QSAR-based models also face problems such as small training sets, experimental data error in training sets, and lack of experimental validations. To overcome these challenges, recently developed AI approaches, such as DL and relevant modeling studies, can be implemented for safety and efficacy evaluations of drug molecules based on big data modeling and analysis. In 2012, Merck supported a QSAR ML challenge to observe the advantages of DL in the drug discovery process in the pharmaceutical industry. DL models showed significant predictivity compared with traditional ML approaches for 15 absorption, distribution, metabolism, excretion, and toxicity (ADMET) data sets of drug candidates [45,46].The virtual chemical space is enormous and suggests a geographical map of molecules by illustrating the distributions of molecules and their properties. The idea behind the illustration of chemical space is to collect positional information about molecules within the space to

search for bioactive compounds and, thus, virtual screening (VS) helps to select appropriate molecules for further testing. Several chemical spaces are open access, including Pub Chem, Chem Bank, Drug Bank, and Chem DB. Numerous in silico methods to virtual screen compounds from virtual chemical spaces along with structure and ligand-based approaches, provide a better profile analysis, faster elimination of non lead compounds and selection of drug molecules, with reduced expenditure. Drug design algorithms, such as coulomb matrices and molecular fingerprint recognition, consider the physical, chemical, and toxicological profiles to select a lead compound[47]. Various parameters, such as predictive models, the similarity of molecules, the molecule generation process, and the application of in silico approaches can be used to predict the desired chemical structure of a compound[48]. Pereira et al. presented a new system, Deep VS, for the docking of 40 receptors and 2950 ligands, which showed exceptional performance when 95 000 decoys were tested against these receptors[49]. Another approach applied a multiobjective automated replacement algorithm to optimize the potency profile of a cyclin-dependent kinase-2 inhibitor by assessing its shape similarity, biochemical activity, and physicochemical properties[50]. QSAR modeling tools have been utilized for the identification of potential drug candidates and have evolved into AI-based QSAR approaches, such as linear discriminate analysis (LDA), support vector machines (SVMs), random forest (RF) and decision trees, which can be applied to speed up QSAR analysis[51-53]. King et al. found a negligible statistical difference when the ability of six AI algorithms to rank anonymous compounds in terms of biological activity was compared with that of traditional approaches[54].

AI in Drug Screening

The process of discovering and developing a drug can take over a decade and costs US$2.8 billion on average. Even then, nine out of ten therapeutic molecules fail Phase II clinical trials and regulatory approval[55,56]. Algorithms, such as Nearest-Neighbour classifiers, RF, extreme learning machines, SVMs, and deep neural networks (DNNs), are used for VS based on synthesis feasibility and can also predict in vivo activity and toxicity[57]. Several biopharmaceutical companies, such as Bayer, Roche, and Pfizer, have teamed up with IT companies to develop a platform for the discovery of therapies in areas such as immuno oncology and cardiovascular diseases. The aspects of VS to which AI has been applied are discussed below.

Prediction of the physicochemical properties

Physicochemical properties, such as solubility, partition coefficient (logP), degree of ionization, and intrinsic permeability of the drug, indirectly affect its pharmacokinetics properties and its target receptor family and, hence, must be considered when designing a new drug[58]. Different AI-based tools can be used to predict physicochemical properties. For example, ML uses large data sets produced

during compound optimization done previously to train the program[59]. Algorithms for drug design include molecular descriptors, such as SMILES strings, potential energy measurements, electron density around the molecule, and coordinates of atoms in 3D, to generate feasible molecules via DNN and thereby predict its properties[60]. Zang et al. created a quantitative structure–property relationship (QSPR) workflow to determine the six physicochemical properties of environmental chemicals obtained from the Environmental Protection Agency (EPA) called the Estimation Program Interface (EPI) Suite. Neural networks based on the ADMET predictor and ALGOPS program have been used to predict the lipophilicity and solubility of various compounds[61]. DL methods, such as undirected graph recursive neural networks and graph-based convolutional neural networks (CVNN), have been used to predict the solubility of molecules[62]. In several instances, ANN-based models, graph kernels, and kernel ridge-based models were developed to predict the acid dissociation constant of compounds[63]. Similarly, cell lines, such as Madin-Darby canine kidney cells and human colon adenocarcinoma (Caco-2) cells have been utilized to generate cellular permeability data of a diverse class of molecules, which are subsequently fed to AI-assisted predictors. Kumar et al. developed six predictive models [SVMs, ANNs, knearestneighbor algorithms, LDAs, probabilistic neural network algorithms, and partial least square (PLS)] utilizing 745 compounds for training; these were used later on 497 compounds to predict their intestinal absorptivity based on parameters including molecular surface area, molecular mass, total hydrogen count, molecular refractivity, molecular volume, logP, total polar surface area, the sum of E- states indices, solubility index (log S), and rotatable bonds[64]. On similar lines, RF and DNN-based in silico models were developed to determine human intestinal absorption of a variety of chemical compounds[65]. Thus, AI has a significant role in the development of a drug, to predict not only its desired physicochemical properties, but also the desired bioactivity.

Prediction of bioactivity

The efficacy of drug molecules depends on their affinity for the target protein or receptor. Drug molecules that do not show any interaction or affinity towards the targeted protein will not be able to deliver the therapeutic response. In some instances, it might also be possible that developed drug molecules interact with unintended proteins or receptors, leading to toxicity. Hence, drug target binding affinity (DTBA) is vital to predict drug–target interactions. AI-based methods can measure the binding affinity of a drug by considering either the features or similarities of the drug and its target. Feature-based interactions recognize the chemical moieties of the drug and that of the target to determine the feature vectors. By contrast, in similarity-based interaction, the similarity between drug and target is considered, and it is assumed that similar drugs will interact with the same targets[66] .Web applications, such as Chem Mapper and the similarity ensemble approach (SEA), are available for predicting drug–

target interactions[67]. Many strategies involving ML and DL have been used to determine DTBA, such as Kron RLS, Sim Boost, Deep DTA, and PADME. ML-based approaches, such as Kronecker-regularized least squares (Kron RLS), evaluate the similarity between drugs and protein molecules to determine DTBA. Similarly, Sim Boost utilized regression trees to predict DTBA, and considers both feature-based and similarity-based interactions. Drug features from SMILES, ligand maximum common substructure (LMCS), extended connectivity fingerprint, or a combination thereof can also be considered. DL approaches have shown improved performance compared with ML because they apply network-based methods that do not depend on the availability of the 3D protein structure. Deep DTA, PADME, Wide DTA, and Deep Affinity are some DL methods used to measure DTBA. Deep DTA accepts drug data in the form of SMILES, whereby, the amino acid sequence is entered for protein input data and for the 1D representation of the drug structure[68] . Wide DTA is CVNN DL method that incorporates

ligand SMILES (LS), amino acid sequences, LMCS, and protein domains and motifs as input data for assessing the binding affinity[69]. Deep Affinity and Protein and Drug Molecule interaction prediction (PADME) are similar to the approaches described earlier[70]. Deep Affinity is an interpretable DL model that uses both RNN and CNN and both unlabeled and labeled data. It takes into account the compound in the SMILES format and protein sequences in the structural and physicochemical properties[71]. PADME is a DL-based platform that utilizes feed-forward neural networks for predicting drug target interactions (DTIs). It considers the combination of the features of the drug and target protein as input data and forecasts the interaction strength between the two. For the drug and the target, the SMILES representation and the protein sequence composition (PSC) are used for illustration, respectively. Unsupervised ML techniques, such as MANTRA and PREDICT, can be used to forecast the therapeutic efficacy of drugs and target proteins of known and unknown pharmaceuticals, which can also be extrapolated to the application of drug repurposing and interpreting the molecular mechanism of the therapeutics. MANTRA groups compound based on similar gene expression profiles using a C Map data set and clusters those compounds predicted to have a common mechanism of action and common biological pathway. The bioactivity of a drug also includes ADME data. AI-based tools, such as Xeno Site, FAME, and SMART Cyp, are involved in determining the sites of metabolism of the drug. In addition, software such as Cyp Rules, Meta Site, Meta Pred, SMART Cyp, and Which Cyp were used to identify specific isoforms of CYP450 that mediate a particular drug metabolism. The clearance pathway of 141 approved drugs was done by SVM-based predictors with high accuracy[72].

Prediction of toxicity

The prediction of the toxicity of any drug molecule is vital to avoid toxic effects. Cell-based in vitro assays are often used as preliminary studies, followed by animal studies to identify the toxicity of a compound, increasing the expense of drug discovery. Several web-based tools, such as Lim Tox, pkCSM, admet SAR, and Tox tree, are available to help reduce the cost. Advanced AI-based approaches look for similarities among compounds or project the toxicity of the compound based on input features. The Tox21 Data Challenge organized by the National Institutes of Health, Environmental Protection Agency (EPA), and US Food

and Drug Administration (FDA) was an initiative to evaluate several computational techniques to forecast the toxicity of 12 707 environmental compounds and drugs; an ML algorithm named DeepTox outperformed all methods by identifying static and dynamic features within the chemical descriptors of the molecules, such as molecular weight (MW) and Van der Waals volume, and could efficiently predict the toxicity of a molecule based on predefined 2500 toxicophore features[73]. The different AI tools used in drug discovery are listed in Table 1. SEA was used to evaluate the safety target prediction of 656 marketed drugs against 73 unintended targets that might produceadverse effects. Developed using an ML-based approach, eToxPred was applied to estimate the toxicity and synthesis feasibility of small organic molecules and showed accuracy as high as 72%. Similarly, open-source tools, such as TargeTox and PrOCTOR, are also used in toxicity prediction[74]. TargeTox is biological network target-based drug toxicity risk prediction method that uses the guilt-by-association principle whereby entities that have similar functional properties share similarities in biological networks[75]. It can produce protein network data and unite pharmacological and functional properties in a ML classifier to predict drug toxicity[76]. PrOCTOR was trained using a RF model and took into account drug-likeliness properties, molecular features, target-based features, and properties of the protein targets to generate a 'PrOCTOR score', which forecasted whether a drug would fail in clinical trials owing to its toxicity. It also recognized FDA-approved drugs that later reported adverse drug events[77]. In another approach, Tox_(R)CNN involving a deep CVNN method evaluated the cytotoxicity of drugs that had been exposed to DAPI-stained cells [78].

AI IN DESIGNING DRUG MOLECULES

Prediction of the target protein structure

While developing a drug molecule, it is essential to assign the correct target for successful treatment. Numerous proteins are involved in the development of the disease and, in some cases, they are overexpressed. Hence, for selective targeting of disease, it is vital to predict the structure of the target protein to design the drug molecule. AI can assist in structure-based drug discovery by predicting the

3D protein structure because the design is in accordance with the chemical environment of the target protein site, thus helping to predict the effect of a compound on the target along with safety considerations before their synthesis or production[79]. The AI tool, AlphaFold, which is based on DNNs, was used to analyze the distance between the adjacent amino acids and the corresponding angles of the peptide bonds to predict the 3Dtarget protein structure and demonstrated excellent results by correctly predicting 25 out of 43 structures. In a study by AlQurashi, RNN was used to predict the protein structure. The author considered three stages (i.e., computation, geometry, and assessment) termed a recurrent geometric network (RGN). Here, the primary protein sequence was encoded, and the torsional angles for a given residue and a partially completed backbone obtained from the geometric unit upstream of this were then considered as input and provided a new backbone as output. The final unit produced the 3D structure as the output. Assessment of the deviation of predicted and experimental structures was done using the distance-based root mean square deviation (dRMSD) metric. The parameters in RGN were optimized to keep the dRMSD low between the experimental and predicted structures [80]. AlQurashi predicted that his AI method would be quicker than AlphaFold in terms of the time taken to predict the protein structure. However, AlphaFold is likely to have better accuracy in predicting protein structures with sequences similar to the reference structures[81]. A study was conducted to predict the 2D structure of a protein using MATLAB assisted by a nonlinear three-layered NN toolbox based on a feed-forward supervised learning and back propagation error algorithm. MATLAB was used to train input and output data sets, and the NNs were learning algorithms and performance evaluators. The accuracy in predicting the 2D structure was 62.72%[82].

Predicting drug-protein interactions

Drug–protein interactions have a vital role in the success of a therapy. The prediction of the interaction of a drug with a receptor or protein is essential to understand its efficacy and effectiveness, allows the repurposing of drugs, and prevents polypharmacology. Various AI methods have been useful in the accurate prediction of ligand–protein interactions, ensuring better therapeutic efficacy[83]. Wang et al. reported a model using the SVM approach, trained on 15 000 protein-ligand interactions, which were developed based on primary protein sequences and structural characteristics of small molecules to discover nine new compounds and their interaction with four crucial targets[84]. Yu et al. exploited two RF models to predict possible drug– protein interactions by the integration of pharmacological and chemical data and validating them against known platforms, such as SVM, with high sensitivity and specificity. Also, these modes were capable of predicting drug-target associations that could be further extended to target-disease and target-target associations, thereby speeding up the drug discovery process[85]. Xiao et al. adopted the Synthetic Minority Over-Sampling Technique and the Neighborhood

Cleaning Rule to obtain optimized data for the subsequent development of iDrugTarget. This is a combination of

four subpredictors (iDrug-GPCR, iDrug-Chl, iDrug-Enz, and iDrug- NR) for identifying interactions between a drug and G-protein coupled receptors (GPCRs), ion channels, enzymes, and nuclear receptors (NR) respectively. When this predictor was compared with existing predictors through target-jackknife tests, the former surpassed the latter in terms of both prediction accuracy and consistency[86]. The ability of AI to predict drug–target interactions was also used to assist the repurposing of existing drugs and avoiding polypharmacology. Repurposing an existing drug qualifies it directly for Phase II clinical trials. This also reduces expenditure because relaunching an existing drug costs US$8.4 million compared with the launch of a new drug entity (US$41.3 million) [87]. The 'Guilt by association' approach can be utilized to forecast the innovative association of a drug and disease, which is either a knowledge-based or computationally driven network[88]. In a computationally driven network, the ML approach is widely used, which utilizes techniques such as SVM, NN, logistic regression, and DL. Logistic regression platforms, such as PREDICT, SPACE, and other ML approaches, consider drug–drug, disease-disease similarity, the similarity between target molecules, chemical structure, and gene expression profiles while repurposing a drug[89]. Cellular network-based deep learning technology (deepDTnet) has been explored to predict the therapeutic use of topotecan, currently used as a topoisomerase inhibitor. It can also be used for the therapy of multiple sclerosis by inhibiting human retinoic acid receptor-related orphan receptor-gamma t (ROR-gt) [90]. This platform is currently under a provisional US patent. Self-organizing maps (SOMs) are in the unsupervised category of ML and are used in drug repurposing. They use a ligand-based approach to search novel off-targets for a set of drug molecules by training the system on a defined number of compounds with recognized biological activities, which is later used for the analysis of different compounds[91]. In a recent study, DNN was used to repurpose existing drugs with proven activity against SARS-CoV, HIV, influenza virus, and drugs that are 3C-like protease inhibitors. In this, extended connectivity fingerprint (ECFP), functional-class fingerprints (FCFPs), and an octanol-water partition coefficient were considered to train the AI platform. From the results, it was concluded that 13 of the screened drugs could be carried toward further development based on their cytotoxicity and viral inhibition[92]. Drug–protein interactions can also predict the chances of polypharmacology, which is the tendency of a drug molecule to interact with multiple receptors producing off-target adverse effects[93]. AI can design a new molecule based on the rationale of polypharmacology and aid in the generation of safer drug molecules[94]. AI platforms such as SOM, along with the vast databases available, can be used to link several compounds to numerous targets and off-targets. Bayesian classifiers and SEA algorithms can be used to establish

links between the pharmacological profiles of drugs and their possible targets[95]. Li et al. demonstrated the use of KinomeX, an AI-based online medium using DNNs for the detection of polypharmacology of kinases based on their chemical structures. This platform uses DNN trained with 14 000 bioactivity data points developed based on >300 kinases. Thus, it has practical application in studying the overall selectivity of a drug towards the kinase family and particular subfamilies of kinases, thus helping to design novel chemical modifiers. This study used NVP-BHG712 as a model compound to predict its primary targets and also its off-targets with reasonable accuracy[96]. One prominent instance is Cyclica's cloud-based proteome-screening AI platform, Ligand Express, which is used to find receptors that can interact with a particular small molecule (the molecular description of which is in SMILE string) and produce on and off-target interactions. This helps in understanding the possible adverse effects of the drug[97]. AI in de novo drug design over the past few years, the de novo drug design approach has been widely used to design drug molecules. The traditional method of de novo drug design is being replaced by evolving DL methods, the former having shortcomings of complicated synthesis routes and difficult prediction of the bioactivity of the novel molecule . Computer-aided synthesis planning can also suggest millions of structures that can be synthesized and also predicts several different synthesis routes for them[98]. Grzybowski et al. developed the Chematica program[99], now renamed Synthia, which has the ability to encode a set of rules into the machine and propose possible synthesizing routes for eight medicinally essential targets. This program has proven to be efficient both in terms of improving the yield and reducing expenses. It is also capable of providing alternate synthesizing strategies for patented products and is said to be helpful in the synthesis of

compounds that have not yet been synthesized. Similarly, DNN focuses on rules of organic chemistry and retrosynthesis, which, with the aid of Monte-Carlo tree searches and symbolic AI, help in reaction prediction and the process of drug discovery and design, which is much faster than traditional methods[100,101]. Coley et al. developed a framework in which a rigid forward reaction template was applied to a group of reactants to synthesize chemically feasible products with a significant rate of reaction. ML was used to determine the dominant product based on a score given by the NNs. Putin et al. explored a DNN architecture called the reinforced adversarial neural computer (RANC) based on RL for de novo design of small organic molecules. This platform was trained with molecules represented as SMILES strings. It then generated molecules with predefined chemical descriptors in terms of MW, logP, and topological polar surface area (TPSA). RANC was compared with another platform, ORGANIC, where the former outperformed in generating unique structures without sufficient loss of their structure length[102]. Even RNN was based on the long short-term memory (LSTM) relating to molecules obtained from the ChEMBL database and fed as SMILES strings. This was used to

generate a diverse library of molecules for VS. This approach was extended to procure novel molecules toward a particular target, such as targets for the 5- HT2A receptor, Staphylococcus aureus, and Plasmodium falciparum[103]. Popova et al. developed the Reinforcement Learning for Structural Evolution strategy for de novo drug synthesis, which involves generative and predictive DNNs to develop new compounds. In this, the generative model produces more unique molecules in terms of SMILE strings based on a stack memory, whereas the predictive models are used to forecast the properties of the developed compound[104]. Merk et al. also exploited the generative AI model to design retinoid X and PPAR agonist molecules, with desired therapeutic effects without requiring complex rules. The authors successfully designed five molecules, four out of which have shown good modulatory activity in cell assays, thereby emphasizing the use of generative AI in new molecule synthesis[105]. The involvement of AI in the de novo design of molecules can be beneficial to the pharmaceutical sector because of its various advantages, such as providing online learning and simultaneous optimization of the already-learned data as well as suggesting possible synthesis routes for compounds leading to swift lead design and development[106].

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[107].

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[108]. 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[109].

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 poly-pharmacology and hence this AI works well toward polypharmacology to better understand the desired target of diseases resulting in best results[110].

CONCLUSION AND FUTURE PROSPECTS

The advancement of AI, along with its remarkable tools, continuously aims to reduce challenges faced by pharmaceutical companies, impacting the drug development process along with the overall lifecycle

of the product, which could explain the increase in the number of start-ups in this sector. The current healthcare sector is facing several complex challenges, such as the increased cost of drugs and therapies, and society needs specific significant changes in this area. With the inclusion of AI in the manufacturing of pharmaceutical products, personalized medications with the desired dose, release parameters, and other required aspects can be manufactured according to individual patient need. Using the latest AI-based technologies will not only speed up the time needed for the products to come to the market, but will also improve the quality of products and the overall safety of the production process, and provide better utilization of available resources along with being cost-effective, thereby increasing the importance of automation. The most significant worry regarding the incorporation of these technologies is the job losses that would follow and the strict regulations needed for the implementation of AI. However, these systems are intended only to make work easier and not to completely replace humans 111-113. AI can not only aid quick and hassle-free hit compound identification, but also contribute

to suggestions of synthesis routes of these molecules along with the prediction of the desired chemical structure and an understanding of drug–target interactions and its SAR. AI can also make major contributions to the further incorporation of the developed drug in its correct dosage form as well as its optimization, in addition to aiding quick decision-making, leading

to faster manufacturing of better-quality products along with assurance of batch-to-batch consistency. AI can also contribute to establishing the safety and efficacy of the product in clinical trials, as well as ensuring proper positioning and costing in the market through comprehensive market analysis and prediction. Although there are no drugs currently on the market developed withAI-based approaches and specific challenges remain with regards to the implementation of this technology, it is likely that AI will become an invaluable tool in the pharmaceutical industry in the near future.

Table 1 Examples of AI tools used in drug discovery

Tools	Details	Website URL
DeepChem	MLP model that uses a python-based AI system to find a suitable candidate in drug discovery	https://github.com/deepchem/deepchem
DeepTox	Software that predicts the toxicity of total of 12 000 drugs	www.bioinf.jku.at/research/DeepTox
DeepNeural NetQSAR	Python-based system driven by computational tools that aid detection of the molecular activity of compounds	https://github.com/Merck/DeepNeuralNet-QSAR
ORGANIC	A molecular generation tool	https://github.com/aspuru-guzik-

	that helps to create molecules with desired properties	group/ORGANIC
PotentialNet	Uses NNs to predict binding affinity of ligands	https://pubs.acs.org/doi/full/10.1021/acscentsci.8b00507
Hit Dexter	ML technique to predict molecules that might respond to biochemical assays	http://hitdexter2.zbh.uni-hamburg.de
DeltaVina	A scoring function for rescoring drug–ligand binding affinity	https://github.com/chengwang88/deltavina
Neural graph fingerprint	Helps to predict properties of novel molecules	https://github.com/HIPS/neural-fingerprint
AlphaFold	Predicts 3D structures of proteins	https://deepmind.com/blog/alphafold
Chemputer	Helps to report procedure for chemical synthesis in standardized format	https://zenodo.org/record/1481731

REFERENCES

1. Wouters OJ, McKee M, Luyten J. Research and Development Costs of New Drugs—Reply. Jama. 2020;324(5):518.
2. Olivier T, Haslam A, Prasad V. Anticancer Drugs Approved by the US Food and Drug Administration From 2009 to 2020 According to Their Mechanism of Action. JAMA network open. 2021;4(12):e2138793-.
3. Brown DG, Wobst HJ. A Decade of FDA-Approved Drugs (2010–2019): Trends and Future Directions. Journal of Medicinal Chemistry. 2021;64(5):2312-38.
4. Ramesh AN, Kambhampati C, Monson JR, Drew PJ. Artificial intelligence in medicine. Annals of the Royal College of Surgeons of England. 2004;86(5):334.
5. Miles JC, Walker AJ. The potential application of artificial intelligence in transport. InIEE Proceedings-Intelligent Transport Systems 2006; 153(3): 183-198.
6. Yang Y, Siau K. A qualitative research on marketing and sales in the artificial intelligence age. Midwest United States Association for Information Systems (MWAIS) 2018 proceedings. 2018.
7. Wirtz BW, Weyerer JC, Geyer C. Artificial intelligence and the public sector—applications and challenges. International Journal of Public Administration. 2019;42(7):596-615.
8. Smith RG, Farquhar A. The road ahead for knowledge management: an AI perspective. AI magazine. 2000 ;21(4):17-.

9. Lamberti MJ, Wilkinson M, Donzanti BA, Wohlhieter GE, Parikh S, Wilkins RG, Getz K. A study on the application and use of artificial intelligence to support drug development. Clinical therapeutics. 2019;41(8):1414-26.
10. Beneke F, Mackenrodt MO. Artificial intelligence and collusion. IIC-International Review of Intellectual Property and Competition Law. 2019;50(1):109-34.
11. Steels L, Brooks R, editors. The artificial life route to artificial intelligence: Building embodied, situated agents. Routledge; 2018.
12. Bielecki A. Models of neurons and perceptrons: selected problems and challenges. Springer International Publishing; 2019.
13. Kalyane D, Sanap G, Paul D, Shenoy S, Anup N, Polaka S, Tambe V, Tekade RK. Artificial intelligence in the pharmaceutical sector: current scene and future prospect. InThe Future of Pharmaceutical Product Development and Research. Academic Press. 2020; 73-107.
14. da Nobrega RV, Reboucas Filho PP, Rodrigues MB, da Silva SP, Dourado Junior CM, de Albuquerque VH. Lung nodule malignancy classification in chest computed tomography images using transfer learning and convolutional neural networks. Neural Computing and Applications. 2020;32(15):11065-82.
15. Medsker L, Jain LC, editors. Recurrent neural networks: design and applications. CRC press; 1999.
16. Hänggi M, Moschytz GS, editors. Cellular neural networks: analysis, design and optimization. Springer Science & Business Media; 2000.
17. Paul D, Sanap G, Shenoy S, Kalyane D, Kalia K, Tekade RK. Artificial intelligence in drug discovery and development. Drug Discovery Today. 2021;26(1):80.
18. https://searchenterpriseai.techtarget.com/definition/ IBM-Watson-supercomputer
19. Vyas M, Thakur S, Riyaz B, Bansal KK, Tomar B, Mishra V. Artificial intelligence: the beginning of a new era in pharmacy profession. Asian J Pharm. 2018;12(2):72-6.
20. Aoyama T, Suzuki Y, Ichikawa H. Neural networks applied to pharmaceutical problems. III. Neural networks applied to quantitative structure-activity relationship (QSAR) analysis. Journal of medicinal chemistry. 1990;33(9):2583-90.
21. Klingler DE. Expert systems in the pharmaceutical industry. Drug Information Journal. 1988; 22(2):249-58.
22. Diogo D, Tian C, Franklin CS, Alanne-Kinnunen M, March M, Spencer CC, Vangjeli C, Weale ME, Mattsson H, Kilpeläinen E, Sleiman PM. Phenome-wide association studies across large population cohorts support drug target validation. Nature communications. 2018;9(1):1-3.

23. De Cesco S, Davis JB, Brennan PE. TargetDB: A target information aggregation tool and tractability predictor. PloS one. 2020;15(9):e0232644.
24. Jiménez-Luna J, Grisoni F, Schneider G. Drug discovery with explainable artificial intelligence. Nature Machine Intelligence. 2020;2(10):573-84.
25. Vignaux PA, Minerali E, Foil DH, Puhl AC, Ekins S. Machine Learning for Discovery of GSK3β Inhibitors. ACS omega. 2020;5(41):26551-61.
26. McCloskey K, Sigel EA, Kearnes S, Xue L, Tian X, Moccia D, Gikunju D, Bazzaz S, Chan B, Clark MA, Cuozzo JW. Machine learning on DNA-encoded libraries: a new paradigm for hit finding. Journal of Medicinal Chemistry. 2020;63(16):8857-66.
27. Kitchen DB, Decornez HY. Computational teChniquesto support hit triage. Small Molecule Medicinal Chemistry: Strategies and Technologies. 2015; 191.
28. Stork C, Chen Y, Sicho M, Kirchmair J. Hit Dexter 2.0: machine-learning models for the prediction of frequent hitters. Journal of chemical information and modeling. 2019;59(3):1030-43.
29. Ain QU, Aleksandrova A, Roessler FD, Ballester PJ. Machine-learning scoring functions to improve structure-based binding affinity prediction and virtual screening. Wiley Interdisciplinary Reviews: Computational Molecular Science. 2015;5(6):405-24.
30. Graff DE, Shakhnovich EI, Coley CW. Accelerating high-throughput virtual screening through molecular pool-based active learning. Chemical Science. 2021.
31. Yang Y, Yao K, Repasky MP, Leswing K, Abel R, Shoichet B, Jerome S. Efficient Exploration of Chemical Space with Docking and Deep-Learning. 2021; 17(11):7106-7119.
32. Paul SM, Mytelka DS, Dunwiddie CT, Persinger CC, Munos BH, Lindborg SR, Schacht AL. How to improve R&D productivity: the pharmaceutical industry's grand challenge. Nature reviews Drug discovery. 2010;9(3):203-14.
33. D Segall M. Multi-parameter optimization: identifying high quality compounds with a balance of properties. Current pharmaceutical design. 2012;18(9):1292-310.
34. Hann MM, Keserü GM. Finding the sweet spot: the role of nature and nurture in medicinal chemistry. Nature reviews Drug discovery. 2012;11(5):355-65.
35. Bian Y, Xie XQ. Generative chemistry: drug discovery with deep learning generative models. Journal of Molecular Modeling. 2021;27(3):1-8.
36. Coley CW, Green WH, Jensen KF. Machine learning in computer-aided synthesis planning. Accounts of chemical research. 2018;51(5):1281-9.

37. Webb D, Jamison TF. Continuous flow multi-step organic synthesis. Chemical Science. 2010;1(6):675-80.
38. B.W.J. Irwin, T.M. Whitehead, S. Rowland, S.Y. Mahmoud, G.J. Conduit, M.D. 1134
39. Irwin BW, Whitehead TM, Rowland S, Mahmoud SY, Conduit GJ, Segall MD. Deep Imputation on Large-Scale Drug Discovery Data. Applied AI Letters. 2021:e31.
40. Whitehead TM, Irwin BW, Hunt P, Segall MD, Conduit GJ. Imputation of assay bioactivity data using deep learning. Journal of chemical information and modeling. 2019;59(3):1197-204.
41. Irwin BW, Levell JR, Whitehead TM, Segall MD, Conduit GJ. Practical applications of deep learning to impute heterogeneous drug discovery data. Journal of Chemical Information and Modeling. 2020;60(6):2848-57.
42. Dolgos H, Trusheim M, Gross D, Halle JP, Ogden J, Osterwalder B, Sedman E, Rossetti L. Translational Medicine Guide transforms drug development processes: the recent Merck experience. Drug discovery today. 2016;21(3):517-26.
43. Toh TS, Dondelinger F, Wang D. Looking beyond the hype: Applied AI and machine learning in translational medicine. EBioMedicine. 2019; 47:607-15.
44. Mak KK, Pichika MR. Artificial intelligence in drug development: present status and future prospects. Drug discovery today. 2019;24(3):773-80.
45. Sellwood MA, Ahmed M, Segler MH, Brown N. Artificial intelligence in drug discovery. Future medicinal chemistry. 2018;10(17):2025-8.
46. Zhu H. Big data and artificial intelligence modeling for drug discovery. Annual review of pharmacology and toxicology. 2020;60:573-89.
47. Ciallella HL, Zhu H. Advancing computational toxicology in the big data era by artificial intelligence: Data-driven and mechanism-driven modeling for chemical toxicity. Chemical research in toxicology. 2019;32(4):536-47.
48. Chan HS, Shan H, Dahoun T, Vogel H, Yuan S. Advancing drug discovery via artificial intelligence. Trends in pharmacological sciences. 2019;40(8):592-604.
49. Brown N. In silico medicinal chemistry: computational methods to support drug design. Royal Society of Chemistry; 2015.
50. Pereira JC, Caffarena ER, Dos Santos CN. Boosting docking-based virtual screening with deep learning. Journal of chemical information and modeling. 2016;56(12):2495-506.
51. Firth NC, Atrash B, Brown N, Blagg J. MOARF, an integrated workflow for multiobjective optimization: implementation, synthesis, and biological evaluation. Journal of chemical information and modeling. 2015;55(6):1169-80.

52. Zhang L, Tan J, Han D, Zhu H. From machine learning to deep learning: progress in machine intelligence for rational drug discovery. Drug discovery today. 2017;22(11):1680-5.
53. Jain N, Gupta S, Sapre N, Sapre NS. In silico de novo design of novel NNRTIs: a bio-molecular modelling approach. RSC Advances. 2015;5(19):14814-27.
54. Wang Y, Guo Y, Kuang Q, Pu X, Ji Y, Zhang Z, Li M. A comparative study of family-specific protein–ligand complex affinity prediction based on random forest approach. Journal of computer-aided molecular design. 2015;29(4):349-60.
55. King RD, Hirst JD, Sternberg MJ. Comparison of artificial intelligence methods for modeling pharmaceutical QSARS. Applied Artificial Intelligence an International Journal. 1995;9(2):213-33.
56. Álvarez-Machancoses Ó, Fernández-Martínez JL. Using artificial intelligence methods to speed up drug discovery. Expert opinion on drug discovery. 2019;14(8):769-77.
57. Fleming N. How artificial intelligence is changing drug discovery. Nature. 2018;557(7706):S55.
58. Dana D, Gadhiya SV, St Surin LG, Li D, Naaz F, Ali Q, Paka L, Yamin MA, Narayan M, Goldberg ID, Narayan P. Deep learning in drug discovery and medicine; scratching the surface. Molecules. 2018;23(9):2384.
59. Zang Q, Mansouri K, Williams AJ, Judson RS, Allen DG, Casey WM, Kleinstreuer NC. In silico prediction of physicochemical properties of environmental chemicals using molecular fingerprints and machine learning. Journal of chemical information and modeling. 2017;57(1):36-49.
60. Yang X, Wang Y, Byrne R, Schneider G, Yang S. Concepts of artificial intelligence for computer-assisted drug discovery. Chemical reviews. 2019;119(18):10520-94.
61. Hessler G, Baringhaus KH. Artificial intelligence in drug design. Molecules. 2018;23(10):2520.
62. Lusci A, Pollastri G, Baldi P. Deep architectures and deep learning in chemoinformatics: the prediction of aqueous solubility for drug-like molecules. Journal of chemical information and modeling. 2013;53(7):1563-75.
63. Kumar R, Sharma A, Siddiqui MH, Tiwari RK. Prediction of human intestinal absorption of compounds using artificial intelligence techniques. Current drug discovery technologies. 2017;14(4):244-54.
64. Rupp M, Körner R, Tetko IV. Estimation of acid dissociation constants using graph kernels. Molecular Informatics. 2010;29(10):731-40.

65. Chai S, Liu Q, Liang X, Guo Y, Zhang S, Xu C, Du J, Yuan Z, Zhang L, Gani R. A grand product design model for crystallization solvent design. Computers & Chemical Engineering. 2020; 135:106764.
66. Thafar M, Raies AB, Albaradei S, Essack M, Bajic VB. Comparison study of computational prediction tools for drug-target binding affinities. Frontiers in chemistry. 2019;7:782.
67. Öztürk H, Özgür A, Ozkirimli E. DeepDTA: deep drug–target binding affinity prediction. Bioinformatics. 2018;34(17):i821-9.
68. Lounkine E, Keiser MJ, Whitebread S, Mikhailov D, Hamon J, Jenkins JL, Lavan P, Weber E, Doak AK, Côté S, Shoichet BK. Large-scale prediction and testing of drug activity on side-effect targets. Nature. 2012;486(7403):361-7.
69. Mahmud SH, 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; 7:48699-714.
70. Gao KY, Fokoue A, Luo H, Iyengar A, Dey S, Zhang P. Interpretable Drug Target Prediction Using Deep Neural Representation. InIJCAI 2018; 3371-3377.
71. Feng Q, Dueva E, Cherkasov A, Ester M. Padme: A deep learning-based framework for drug-target interaction prediction. arXiv preprint arXiv:1807.09741. 2018.
72. Karimi M, Wu D, Wang Z, Shen Y. DeepAffinity: interpretable deep learning of compound–protein affinity through unified recurrent and convolutional neural networks. Bioinformatics. 2019;35(18):3329-38.
73. Pu L, Naderi M, Liu T, Wu HC, Mukhopadhyay S, Brylinski M. e toxpred: A machine learning-based approach to estimate the toxicity of drug candidates. BMC Pharmacology and Toxicology. 2019;20(1):1-5.
74. Mayr A, Klambauer G, Unterthiner T, Hochreiter S. DeepTox: toxicity prediction using deep learning. Frontiers in Environmental Science. 2016; 3:80.
75. Basile AO, Yahi A, Tatonetti NP. Artificial intelligence for drug toxicity and safety. Trends in pharmacological sciences. 2019;40(9):624-35.
76. Lysenko A, Sharma A, Boroevich KA, Tsunoda T. An integrative machine learning approach for prediction of toxicity-related drug safety. Life science alliance. 2018;1(6).
77. Basile AO, Yahi A, Tatonetti NP. Artificial intelligence for drug toxicity and safety. Trends in pharmacological sciences. 2019;40(9):624-35.
78. Gayvert KM, Madhukar NS, Elemento O. A data-driven approach to predicting successes and failures of clinical trials. Cell chemical biology. 2016;23(10):1294-301.

79. Jimenez-Carretero D, Abrishami V, Fernandez-de-Manuel L, Palacios I, Quílez-Álvarez A, Díez-Sánchez A, Del Pozo MA, Montoya MC. Tox_ (R) CNN: Deep learning-based nuclei profiling tool for drug toxicity screening. PLoS computational biology. 2018;14(11):e1006238.
80. Wan F, Zeng JM. Deep learning with feature embedding for compound-protein interaction prediction. bioRxiv. 2016:086033.
81. AlQuraishi M. End-to-end differentiable learning of protein structure. Cell systems. 2019;8(4):292-301.
82. Hutson M. AI protein-folding algorithms solve structures faster than ever. Nature. 2019.
83. Avdagic Z, Purisevic E, Omanovic S, Coralic Z. Artificial intelligence in prediction of secondary protein structure using CB513 database. Summit on translational bioinformatics.2009;1.
84. Tian K, Shao M, Wang Y, Guan J, Zhou S. Boosting compound-protein interaction prediction by deep learning. Methods. 2016; 110:64-72.
85. Wang F, Liu D, Wang H, Luo C, Zheng M, Liu H, Zhu W, Luo X, Zhang J, Jiang H. Computational screening for active compounds targeting protein sequences: methodology and experimental validation. Journal of chemical information and modeling. 2011;51(11):2821-8.
86. Yu H, Chen J, Xu X, Li Y, Zhao H, Fang Y, Li X, Zhou W, Wang W, Wang Y. A systematic prediction of multiple drug-target interactions from chemical, genomic, and pharmacological data. PloS one. 2012;7(5):e37608.
87. Xiao X, Min JL, Lin WZ, Liu Z, Cheng X, Chou KC. iDrug-Target: predicting the interactions between drug compounds and target proteins in cellular networking via benchmark dataset optimization approach. Journal of Biomolecular Structure and Dynamics. 2015;33(10):2221-33.
88. Persidis A. The benefits of drug repositioning. Drug Discov World. 2011;12:9-12.
89. Koromina M, Pandi MT, Patrinos GP. Rethinking drug repositioning and development with artificial intelligence, machine learning, and omics. Omics: a journal of integrative biology. 2019;23(11):539-48.
90. Park K. A review of computational drug repurposing. Translational and Clinical Pharmacology. 2019;27(2):59-63.
91. Zeng X, Zhu S, Lu W, Liu Z, Huang J, Zhou Y, Fang J, Huang Y, Guo H, Li L, Trapp BD. Target identification among known drugs by deep learning from heterogeneous networks. Chemical Science. 2020;11(7):1775-97.
92. Achenbach J, Tiikkainen P, Franke L, Proschak E. Computational tools for polypharmacology and repurposing. Future medicinal chemistry. 2011;3(8):961-8.

93. Ke YY, Peng TT, Yeh TK, Huang WZ, Chang SE, Wu SH, Hung HC, Hsu TA, Lee SJ, Song JS, Lin WH. Artificial intelligence approach fighting COVID-19 with repurposing drugs. Biomedical journal. 2020;43(4):355-62.
94. Li X, Xu Y, Cui H, Huang T, Wang D, Lian B, Li W, Qin G, Chen L, Xie L. Prediction of synergistic anti-cancer drug combinations based on drug target network and drug induced gene expression profiles. Artificial intelligence in medicine. 2017;83:35-43.
95. Reddy AS, Zhang S. Polypharmacology: drug discovery for the future. Expert review of clinical pharmacology. 2013;6(1):41-7.
96. Li Z, Li X, Liu X, Fu Z, Xiong Z, Wu X, Tan X, Zhao J, Zhong F, Wan X, Luo X. KinomeX: a web application for predicting kinome-wide polypharmacology effect of small molecules. Bioinformatics. 2019;35(24):5354-6.
97. Paul D, Sanap G, Shenoy S, Kalyane D, Kalia K, Tekade RK. Artificial intelligence in drug discovery and development. Drug Discovery Today. 2021;26(1):80.
98. Corey EJ, Wipke WT. Computer-assisted design of complex organic syntheses. Science. 1969;166(3902):178-92.
99. Grzybowski BA, Szymkuć S, Gajewska EP, Molga K, Dittwald P, Wołos A, Klucznik T. Chematica: a story of computer code that started to think like a chemist. Chem. 2018;4(3):390-8.
100. Klucznik T, Mikulak-Klucznik B, McCormack MP, Lima H, Szymkuć S, Bhowmick M, Molga K, Zhou Y, Rickershauser L, Gajewska EP, Toutchkine A. Efficient syntheses of diverse, medicinally relevant targets planned by computer and executed in the laboratory. Chem. 2018;4(3):522-32.
101. Segler MH, Preuss M, Waller MP. Planning chemical syntheses with deep neural networks and symbolic AI. Nature. 2018;555(7698):604-10.
102. Putin E, Asadulaev A, Ivanenkov Y, Aladinskiy V, Sanchez-Lengeling B, Aspuru-Guzik A, Zhavoronkov A. Reinforced adversarial neural computer for de novo molecular design. Journal of chemical information and modeling. 2018;58(6):1194-204.
103. Segler MH, Kogej T, Tyrchan C, Waller MP. Generating focused molecule libraries for drug discovery with recurrent neural networks. ACS central science. 2018;4(1):120-31.
104. Popova M, Isayev O, Tropsha A. Deep reinforcement learning for de novo drug design. Science advances. 2018;4(7):7885.
105. Merk D, Friedrich L, Grisoni F, Schneider G. De novo design of bioactive small molecules by artificial intelligence. Molecular informatics. 2018;37(1-2):1700153.

106. Schneider G, Clark DE. Automated de novo drug design: are we nearly there yet?. Angewandte Chemie International Edition. 2019;58(32):10792-803.
107. Mayr A, Klambauer G, Unterthiner T, Hochreiter S. DeepTox: toxicity prediction using deep learning. Frontiers in Environmental Science. 2016; 3:80.
108. Bain EE, Shafner L, Walling DP, Othman AA, Chuang-Stein C, Hinkle J, Hanina A. Use of a novel artificial intelligence platform on mobile devices to assess dosing compliance in a phase 2 clinical trial in subjects with schizophrenia. JMIR mHealth and uHealth. 2017;5(2):e7030.
109. Deliberato RO, Celi LA, Stone DJ. Clinical note creation, binning, and artificial intelligence. JMIR medical informatics. 2017;5(3):e24.
110. Reddy AS, Zhang S. Polypharmacology: drug discovery for the future. Expert review of clinical pharmacology. 2013;6(1):41-7.
111. Rantanen J, Khinast J. The future of pharmaceutical manufacturing sciences. Journal of pharmaceutical sciences. 2015;104(11):3612-38.
112. Jämsä-Jounela SL. Future trends in process automation. Annual Reviews in Control. 2007;31(2):211-20.
113 Davenport TH, Ronanki R. Artificial intelligence for the real world. Harvard business review. 2018;96(1):108-16.