ASIAN JOURNAL OF PHARMACEUTICAL AND CLINICAL RESEARCH

INNOVARE ACADEMIC SCIENCES
Knowledge to Innovation

Vol 15, Issue 4, 2022

Online - 2455-3891
Print - 0974-2441

Review Article

ARTIFICIAL INTELLIGENCE IN PHARMACY DRUG DESIGN

NISHA V KALAYIL*, SHONA S D'SOUZA, SHOWKHIYA Y KHAN, PALLAVI PAUL

Department of Pharmaceutics, St. John College of Pharmacy and Research, Palghar, Maharashtra, India. Email: nishak@sjipr.edu.in

Received: 14 December 2021, Revised and Accepted: 22 January 2022

ABSTRACT

Drug discovery is said to be a multi-dimensional issue in which different properties of drug candidates including efficacy, pharmacokinetics, and safety need to be improved with respect to giving the final drug product. Current advances in fields such as artificial intelligence (AI) systems that refine the design thesis through report investigation, microfluidics-assisted chemical synthesis, and biological testing are now giving a cornerstone for the establishment of greater automation into detail of this process. AI has stimulated computer-aided drug discovery. This could likely speed up time duration for compound discovery and enhancement and authorize more productive hunts of related chemicals. However, such optimization also increases substantial theories, technical, and organizational queries, as well as suspicion about the ongoing boost around them. Machine learning, in particular deep learning, in multiple scientific disciplines, and the development in computing hardware and software, among other factors, continue to power this development worldwide.

Keywords: Drug discovery, Artificial intelligence systems, Computer-aided drug discovery, Quantitative structure-activity relationship approach.

 DOI: http://dx.doi.org/10.22159/ajpcr.2022v15i4.43890. Journal homepage: https://innovareacademics.in/journals/index.php/ajpcr

INTRODUCTION

Drug discovery

The process of Drug Discovery affects the whole Pharmaceutical sector including all the phases such as a preliminary phase of research from target Discovery and validation to the discerning of molecules. A variety of streams can be used to initiate the identification of small therapeutic molecules [1]. New research can lead to awareness of different diseases with different routes of drug administration and can be developed to take part in pharmaceutical companies to run large-scale trials and other waste programs, in order to identify the targeted molecular compounds. This process is mostly performed at the beginning of lead Discovery, with the prospect of taking identified compounds in the right way through preclinical and clinical trials [2,3].

The steps involved in the process of drug discovery

Step 1 – Target identification and validation

The first step involves target identification and validation as it starts the whole drug discovery process. The basic targets for therapeutic use are cellular on modular structures which materialize naturally and take part in prime roles in pathogenicity [3]. The target molecule should be conveniently safe and efficacious and should meet the clinical demands of the patient to get the desired drug molecule after identification of the drug target system approach should be performed in the mode of action of the API to be qualified for efficacy (Fig. 1).

Step 2 – hit identification and validation

The next step is to recognize whether the small molecule leads have the desired effect against the identified targets. Hits can be identified using several techniques, such as high-throughput screening, knowledge-based approaches, and virtual screening. Validation of hits is important during the initial stage of screening.

Step 3 – moving from a hit to a lead validation

The aim now is to clarify each hit series to develop more selective compounds based on the various series of hits introduced so far. It is particularly important to work on multiple series in random order since some successful series will fail due to unique characteristics. Focusing on multiple series structurally, different sets of hit series will help to offset this possibility.

Step 4 – lead optimization

To develop a preclinical drug candidate, it is necessary to improve on the deficiencies of lead compounds and maintain their desired properties. Using this step, you can determine whether your drug metabolizes in the right area of your body, or whether you have any current side effects of concern. An integrated approach is recommended for the same. Bringing together experts in computational chemistry, medical chemistry, drug metabolism, and other fields will allow them to provide unique insight into this late stage of the process [4-7].

Step 5 – late lead optimization

Before progression to preclinical and clinical trials, late-stage optimization, in which further pharmacological safety of a lead compound is assessed, is a vital step. A drug's efficacy, pharmacokinetics, and safety will be more likely to be compromised later in development if this phase is ignored. As part of safety optimization, the aim is to identify and progress leads with the best overall safety profile, remove the most toxic leads, and establish a well-defined hazard and translational risk profile necessary for further *in vitro* testing. By thoroughly calculating the risks at this point, more opportunities can be taken when investments are made into leads. Further, if the drug is approved for development it is passed to perform preclinical and clinical trials [8].

Artificial intelligence (AI)

AI and the area of machine learning (ML) are the study of the processes and viability of sanctioning machines to cleverly execute intelligent tasks, without obviously being programmed for those tasks [9,10]. At prresent, AI systems have approached or excelled human performance in multiple tasks, such as game playing and image recognition, but these have generally been quite limited and focused domains. However, AI in its varied forms is today effectively used across several domains and for complex tasks, such as robotics, speech synthesis, image analysis, and logistics, as well as its application in the design of molecules (Fig. 2).

Since 1960, medicinal chemistry has been applied to AI in different forms and with an assorted degree of progress in designing compounds. It involves supervised learning, in which labeled training datasets are used to train models broadly. One example is the quantitative structure-activity relationship (QSAR) approach, which is frequently

used to predict properties, such as log P and solubility, for chemical structures [4]. Unlike unsupervised learning, which is not based on labels, this is also popular within medicinal chemistry, with examples such as hierarchical clustering, algorithms, and principal components analysis used broadly to examine and break down large molecular libraries into smaller collections of indistinguishable compounds [11].

AI neural networks

AI in drug discovery

The Research and Development (R and D) cycle that encapsulates: Problem analysis, grounding, design, implementation, and evaluation, for innovative small-molecule drugs as an example are required to overcome obstacles, such as high marketing cost, limited success rates in clinical trials, and overall prolonged R and D cycle time. The R and D productivity of drugs (for drugs of small-molecule) in the medical care industry continues to dwindle, despite expenditure records. In addition, partially because of the saturation of the market, it has been challenging to bring novel chemical matter through a complex approval process, and partly because people are willing to pay in both developed and developing markets.

In recent times, scientists have more information than ever before on an assortment of topics that pertain to the subject, outperforming the ability of most to properly parse and amalgamate into their own workflows and research objectives [12,13]. One solution to this kind of complication is to outsource our reasoning to machine intelligence when it comes to the analysis of multisource and multidimensional data. ML is a type of AI that enhances the accuracy of predictive outcomes and offers fresh opportunities for small-molecule drug discovery [16]. ML approaches that might be considered instances of weak AI have made remarkable progress, with blossoming in their fundamental algorithms and application [14-17].

AI has recently emerged as a hot topic in the pharmaceutical industry. AI in Biopharmaceutical Company's initiative subsistence rose at a summer workshop in 1956 at Dartmouth College, The Dartmouth Summer Research Project on AI was a groundbreaking event for AI. During the 1970s, a number of AI-based diagnostic tools were developed, such as PIP, Internist-1, MYCIN (an early expert system, or AI program, for treating blood infections), and Common Attribute Support Network. In 1985, the R1 (eXpert CONfigurer, for expert configure) program was a production rule-based system that achieved 95–98% accuracy and was blown to a million-dollar industry [18]. In the course of the early 1990s, AI exponentially grew in data communication (Internet), cloud technologies (Apps, Elastic Computing cloud, etc.), high-performance cloud computing, and big data storage. In the chess match between Garry Kasparov and an IBM supercomputer called Deep Blue, Deep Blue defeated Kasparov in 1997 which put AI in the limelight [19]. Over the past decade, AI has expanded exponentially, and there has been a significant increase in investments in AI for drug discovery. A firm in Silicon Valley called Andréessen Horowitz recently announced a 450 million dollar bio investment fund, with AI as a major focus area.

APPLICATIONS OF AI IN OTHER PHARMACEUTICAL SUBFIELDS

AI in molecular designing technique

In drug discovery, computer-aided drug design is becoming ever more important for the cost-effective identification of promising drug candidates [20,21]. The scientific methods presented here are of relevance to limiting the use of animals in pharmacological research, for aiding the rational design of novel and safe drug candidates, and for repositioning marketed drugs, supporting medicinal chemists and pharmacologists during the discovery phase of new drugs. Here are some of the commonly used molecular designing techniques (Fig. 3).

QSAR/quantitative structure-property relationship (QSPR) and structure-based modeling with AI

Over the past 50 years, QSAR/QSPR modeling has improved greatly. Statistical models such as absorption, distribution, metabolism, excretion, and toxicity (ADMET) provide accurate predictions of biological activity as well as pharmacokinetic parameters like ADMET [22,23]. When modeling ligand-based QSAR/QSPR, the structural characteristics of molecules (such as pharmacophore distribution, physicochemical properties, and functional groups) are commonly converted into machine-readable numbers using molecular descriptors. The range of hand-made molecular descriptors is wide, which aims to capture different features of the key chemical structure [24-26]. Conventionally, QSAR/QSPR have focused on simplified models, such as linear regression and k-nearest neighbors. However, in the last decade, ML techniques have become widely available, such as support vector machines (SVMs) and gradient boosting methods, which aim to reveal complex and potentially nonlinear relationships between the chemical structure and its physicochemical/biological properties, often at a charge of accountability. Over time, deep learning has grown to become a popular part of ML [27]. Autoencoders, recurrent systems for sequence, deep and adaptive network architectures, involving Chemo informatics Artificial Neural Networks reached their peak during the 1990s, which was when most of the current concepts were pioneered. It was not until deep networks won the 2012 Merck Molecular Activity Challenge that they achieved their ultimate breakthrough. Though it has not been proven that the latest type of model is superior to other techniques (such as gradient boosting machines), the same set of descriptors are used. Using the same set of descriptors, deep learning has several advantages. However, the most important thing during training is that deep networks can extract features automatically. To build internal context-specific representations of molecular structure, message-passing techniques are used, for example, Graph Neural Networks or Recurrent Neural Networks, in more detail [28]. The latent representations of atoms and bonds are learned during the training process for graph neural networks. The advantages of deep learning approaches are that they can be applied to modeling tasks for which conventional descriptors were not originally designed. Several examples include the modeling of peptides, macrocycles, and proteolysis-targeting chimeras. Furthermore, multiple-layer architectures can also be applied to multitask learning, which has the goal of identifying a common internal representation in order to address a set of related endpoints. Multitask learning based on AI could be a more efficient and effective way to use correlated data than prior imputation, especially in scenarios where the molecular library is not tested exhaustively on all endpoints [29,30].

In the years before deep learning became popular, multi-output QSAR modeling aimed at linking a set of predefined chemical descriptors to observable endpoints was being explored. In spite of the promise of multitask learning, performance improvements over single-task models have only been modest so far [31,32].

In contrast to traditional QSAR, structure-based prediction of protein-ligand interactions for the same targets has made remarkable progress. The classical method of modeling a protein-ligand complex with partial

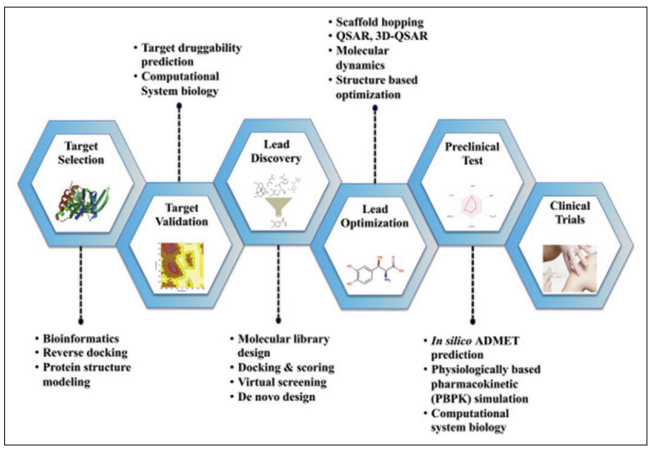

Fig. 1: Steps involved in drug discovery [3]

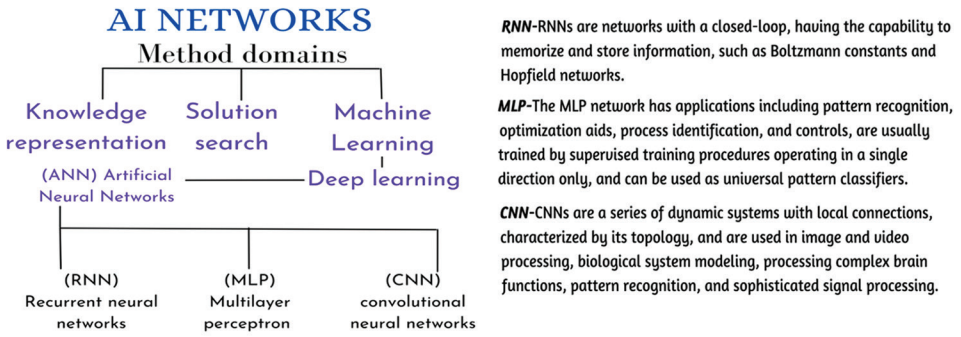

Fig. 2: Artificial networks including method domains and classification of ANNs-Artificial Neural networks [11]

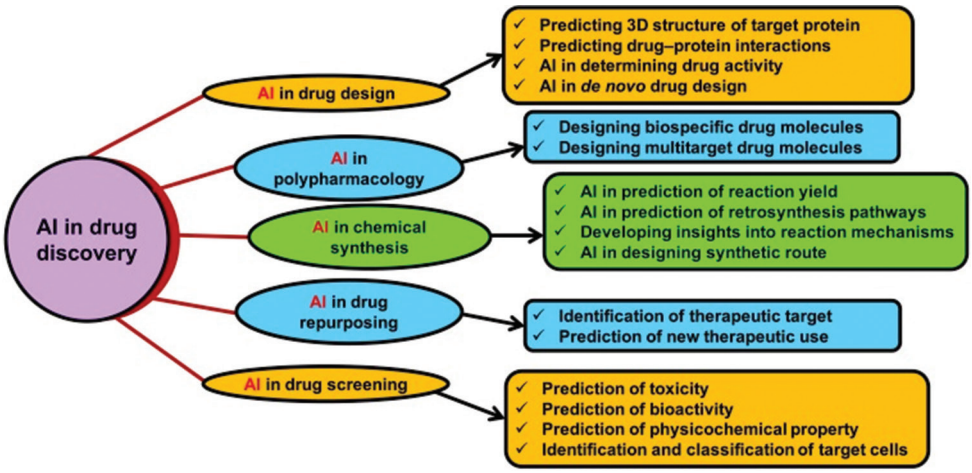

Fig. 3: Artificial intelligence applications in various subfields of the pharmaceutical industry, ranging from drug design to drug screening [19]

least squares or multiple linear regressions takes into consideration the contribution of individual descriptors (e.g., physicochemical properties) for a target property [33]. Random forests or SVMs emerged in 2000, becoming popular by the early 2010s. These models are often used with an array of descriptors such as protein-ligand atom pairs, shape data, or basic atomic interactions, which are often combined with nonlinear QSAR models that are more flexible. In the same manner, as ligand-based research, this subfield also benefited from deep learning. Using a convolutional neural network, the techniques before were developed for bioactivity prediction based on advances in computer vision and image recognition [34]. Many other approaches use graph-based approaches in conjunction with distance and angle-based feature selection. Tests based on sets from a database drawn pseudo-randomly may result in overconfident results regarding model evaluation. Other options such as scaffold-based or time-based splits are potentially more informative as they try to resemble the course of a lead optimization project [35].

***De novo* drug design with AI**

In drug discovery, one of the most challenging computer-aided tasks is the creation of new molecules with desired pharmacological properties from scratch. This is due to the unique chemical space of molecules with such properties [36,37]. In *de novo* molecule generation, there is a danger of a series of combinative explosions because of the enormous amount of different atomic types and molecular topologies a researcher could explore. As a result, ligand-based approaches, structure-based approaches, or a combination of both approaches may be used based on the information aided to guide the *de novo* design process.

The ligand-based methodology can be grouped into two main categories:

(i) A rule-based approach consists of a set of rules for building molecules from reagents or molecular fragments [38]. For the process of stepping through analogs of active leads to their maximum potency, the Topliss scheme is one of the forerunners of modern rule-based *de novo* design [10].

Modern approaches to optimization use a set of molecular transformations, such as matched pairs of molecules, or rules-of-thumb for functional groups, and molecular framework modification. Synthesis-oriented approaches include rules for building block assembly and ligand generation in particular [39]. These approaches can be used to design synthetically accessible libraries, such as BI CLAIM and CHIPMUNK. Hybrid approaches such as Testing Operations Provisioning and Administration System, Difference of Gaussian, and DINGOES, have been developed to guide the generation of novel compounds, since the late 1990s, by concurrently maximizing both, their similarity to known bioactive ligands and the chemical synthesizability of the designs [40-44].

(ii) Rule-free approaches have no rigid construction rules, but rather aim to generate molecules with desired properties without the need for molecular construction rules. Molecular representations are learned by learning generative deep-learning models, which sample new molecules from the latent molecular representation. Although these approaches have gained popularity within a previous couple of years, the concept of sampling from a numerical presentation of molecules for *de novo* design goes back to the "inverse QSAR" problem formulated in the pioneering work of Skvortsova and Zefirov in the early 1990s [45]. The present QSAR model, anchored by inverse QSAR, not only detects the descriptive values which correlate to the desired property but also utilizes this information for molecule production. In contrast, the latter method poses several obstacles, such as the existence of multiple solutions for each property and the difficulty of reverse-decoding cogent structures from molecular descriptors. To address a few of these problems, deep learning can be used to generate new compounds by selecting molecules from a known distribution and then designing the

fundamental distribution [46]. The most common implementation is Simplified Molecular Input Line Entry System (SMILES) paired with generative models borrowed from natural language processing. It teaches them how to use the SMILE "syntax" (i.e., create a chemically valid string), based on chosen "semantics" (e.g., bioactivity or other desired molecular property). They generally rely on recurrent neural networks, combined with transfer or reinforcement learning [47]. There have also been a variety of other deep-learning-based generative learning models reported, such as variational autoencoders and generative adversarial networks, as well as graph-convolution models. Some recent papers have proposed conditional generative design approaches, which take advantage of supplementary factors steering the design, such as three-dimensional shape, drug-likeness, synthesizability, and molecular descriptor values [48]. Elucidating balanced objective functions that enable intricate and constrained multi-parameter optimizations will be a major challenge in the future, analogous to those found in Pareto or desirability-based approaches that are classic in drug discovery [50,51].

As a result of the rapid development of generative neural networks, the number of ligand-based design methods has risen dramatically. In addition to neural generative models, GuacaMol and MOSES provide a set of metrics that can compare them, as well as more traditional models (like genetic algorithms)[52]. *de novo* design tools are generally more difficult to evaluate retrospectively than predictive methods, but some of the most commonly used metrics include:

1. Rationality of the created molecular representations and novelty of the simultaneous molecules
2. Chemical and biological resemblance with the known compounds, and
3. Scaffold and fragment diversity.

The advantages of rule-based and rule-free approaches differ. The rule-based methods use preexisting knowledge, such as building blocks and reaction rules, to generate molecules that are often freely synthesizable and have the desired physicochemical properties. Hard-coded rules and library selections, however, play a role in influencing the chemical diversity of the designs. Theoretically, rule-free approaches learn straight from the data without requiring hard-coded design or similarity rules, thereby permitting a wide exploration of the chemical space. But to the disadvantage, this freedom of exploration endangers the generation of compounds that are more difficult, if not impossible, to synthesize. Lately, a combined strategy was demonstrated to be feasible for designing bioactive molecules in a rule-free fashion, while simultaneously retaining synthesizability within a microfluidics system, using predefined virtual reactions [53].

The majority of deep-learning-based *de novo* design studies, till now, have concentrated on ligand-based approaches. Research directed at orphan receptors and hitherto unexplored macromolecules through structure-based generative design has great potential. Based on the best knowledge known, deep learning has not yet heavily invaded these approaches that utilize information about the ligand-binding site (e.g., fragment linking or growing) [54]. In the meantime, prime developments in ligand design have been made based on the shape and properties of the binding pocket.

Automated synthesis planning with AI

The fragments of today, the drugs that treat diseases, the fertilizers that safeguard our crops, the products that make life accessible are becoming rapidly worldly-wise and refined thanks to advancements in chemical synthesis. As tools for synthesis upgrades, molecular designers can be valiant and innovative in the way they outline and put together the molecules [55,56]. A variety of devices combining chemical synthesis and data science have come to the forefront in recent years, including robotics for autonomous or high-throughput synthesis, as well as algorithms for retrosynthesis and reaction prediction. Recently, progress has been made in retro-synthesis logic, predicting reactivity, and chemistry automata, with assurance from contributors in several fields. Chemical synthesis in the information era guarantees to increase the quality of the molecules of the future with help of data-harnessing and automation. Automation molecule synthesis involves three main components: Retrosynthesis, reaction prediction, and automated synthesis. The following topics describe how the strategy of multistep synthesis can be distilled into a logic that can be taught to a computer. Various tools and models are implemented in the part on reaction prediction, including those for developing reaction conditions, catalysts, and even the most recent alterations based on the abundance of data sets and census tools such as ML [57-60].

Synthesis planning with AI has a well-of past, in 1970 in the field of computer-aided retrosynthetic prediction. Improved computational facilities, the surfacing of big data, and the advancement of novel algorithms for deep learning and expansion have arisen in a revival of AI for synthetic organic chemistry [61]. In retrosynthesis, where the main aim is to periodically design efficient synthetic pathways for a molecule of the concert, rule-based techniques are undoubtedly beneficial. These aim to stipulate retrosynthetic pathways with the help of reaction mechanism encoding and skeletal building. One of their main curbs is their reliance on direct chemical changes/reactions. These generally demand physical establishment and administration. The latest drawn inspiration of the field is from natural languages processing techniques, such as sequence-to-sequence models and transformer models [62,63].

Forward synthesis planning differentiates itself from retrosynthesis. While the final might be explainable with the help of existing reaction databases, forward synthesis would be in need of data from reactions that give no product though. The present chemical reaction databases are heavily distorted regarding productive reaction data. There is an essential call for additional data, like experimental conditions (e.g., solvent and temperature) or side-product information. With the goal of addressing some of these disadvantages, several measures have been taken to develop known reaction databases with negative reaction results, and therefore, design the latest customized data collection for automated synthesis planning [64].

THE RISE OF AI IN THE DRUG DISCOVERY MARKET

The biopharmaceutical sector offers stimulating opportunities for advancements in AI. It is no secret that the biopharmaceutical industry is working hard to leverage AI to improve drug discovery, decrease costs for R and D, reduce failure rates, and ultimately develop superior drugs [65]. Recent years have seen an evolution of AI-based start-up companies devoted to drug discovery as a result of the fast development of ML algorithms and the availability of immense statistics in life sciences. Several notable AIbio pharmaceutical alliances were announced in 2016-2017, such as Pfizer and IBM Watson, AstraZeneca, Abbvie, Merck, Novartis, GSK Sanofi Genzyme, and Recursion Pharmaceuticals, and Exscientia [66].

The top biopharmaceutical company's current AI initiatives include:

a. Mobile platforms - the caliber to recommend patients, by the means of real-time data collection and thus improving patient outcomes
b. Personalized medicine - the capability to evaluate a big database of patients so as to recognize cure alternatives using a cloud-based system as personalized medicine
c. Acquisitions galore - new startups are combining AI and healthcare to foster the innovation requirements of large biotech firms
d. Drug Discovery - in conjunction with software companies, pharma companies are trying to implement the most cutting-edge technologies in the costly and extensive process of drug discovery [9,62].

Based on the growth of deep learning applications in drug discovery and the fact that these methods avail from large training sets, conscientious data curation and proper benchmarking of newly developed models are mandatory [67]. The availability and size of chemical compound libraries have improved over the past few years, with databases such as ZINC and European Molecular Biology Laboratory representing a commonly-used starting point for ligand-based projects. A similar trend was observed for structure-based modeling, for which databases

such as protein database bind and binding database provide highly detailed structural information on protein-ligand complexes, as well as their associated bioactivity data [68,69].

CHALLENGES OF AI IN DRUG DISCOVERY

The drug discovery and development process are very lengthy, highly expensive, and extremely complex in nature. Conventional methods involve expensive techniques and take long years to launch a new drug in the market. With the dawn of new tools and technologies in the field, the major challenge is to reduce the time and cost required for the development of a new drug. These composite problems involve extremely high computations and can be addressed with the help of techniques based on AI [70].

A well-known drawback of deep learning is its poor performance in medium-to-low data scenarios. Further insight might be provided into these scenarios by chemogenomic-based approaches alongside exploiting additional genomic or biological interactome data sources [71]. In addition, recent advances in "few-shots" learning (i.e., a set of approaches that can use prior knowledge to obtain better generalization when data is scarce) and meta-learning (i.e., a branch of metacognition that aims to develop a set of learnable frameworks that can quickly adapt to new, unseen tasks) hold promise in alleviating some of these issues [72,73]. Data-driven approaches for molecular predictions of properties are in contradiction to techniques that are physics-based and are also fundamentally limited in their ability to extrapolate and make reliable predictions for the unseen classes of compounds. ML approaches are physics-inspired and additional active learning strategies (i.e., approaches where for improvised generalization the model has a role in requesting specific training data) provide supplementary tools to overcome these limitations [74].

AI techniques could offer partial solutions to these problems by providing intelligible interpretations of the decision-making process undertaken by deep learning approaches. Continued development of feature attribution techniques (i.e., approaches aiming to highlight the comprehensive importance of an input) instance-based elucidations (e.g., counterfactuals, model-generated examples that are conditioned on user-defined queries), and attention-based networks will help narrow the gap between deep learning and drug-discovery specialists. Hence, close collaboration between these fields is crucial [75].

Another commonly-claimed drawback of deep-learning approaches is their high computational cost. Deep learning typically entails longer training and evaluation times than many other machine-learning approaches without specialized hardware such as consumer-grade graphical processing or tensor-processing units [76]. Deep-learning models can be learned in an online setting by naturally taking advantage of its most popular training strategy, that is, stochastic gradient-descent optimization. This is advantageously scaling linearly with respect to the size of the training dataset, and thus it does not require the latter to be entirely loaded into the system's memory. Many argue that the capability to train deep learning models stochastically on sequential, random, batches of data can make them more suitable than other alternatives in big data scenarios [77-80].

FUTURE BEHOLD

Reports have shown that a few companies are engaging *in silico* synthetic planning into their whole course, accessing target molecules through the use of AI and ML; this has proven to be a beneficial technique in predictive chemistry and synthetic planning of small molecules [93]. ML for Pharmaceutical Discovery and Synthesis (MLPDS) consortium, composed of MIT and 13 chemical and pharmaceutical companies, is developing and evaluating a data-driven synthesis planning program. The integration of predictive models into the medicinal chemistry synthesis workflow, the use of these models in MLPDS member companies, and the outlook for this field [81].

Predictive deep learning tends to require significantly more human expertise in many practical scenarios compared to other, more thoroughly tested approaches. While one can train a well-performing random forest model with a relatively small effort for hyperparameter tuning, our understanding of contemporary deep learning approaches is not yet at the level of reliable defaults, although recent theory suggests that this may change soon. Deep neural networks might provide the right answers for misleading reasons and tend to produce overly confident predictions, even when these are evidently wrong [82,83]. This is further exacerbated in the context of property prediction in drug discovery, as experiments under similar conditions can provide significantly different measurements. This drawback might be alleviated in the next few years with the wider adoption of uncertainty estimation techniques, either with deep learning approaches that have uncertainty directly embedded into their design, such as Bayesian neural networks, or *post hoc* techniques such as ensemble learning [84].

Chemical reaction prediction problems can be viewed as graph transformation problems as a way to directly rank products. In recent years, quantum mechanics has led to the development of approaches that make use of first-principle calculations (e.g., density functional theory) to assess the energy barrier of a reaction. The accurate prediction of energies and forces through quantum-mechanical ML might help bridge this gap soon. Concerning template-free forward synthesis prediction, natural language processing approaches based on the transformer or recurrent neural network architectures are also becoming popular [85-90]. These have reported a top-1 reactant accuracy above 90%. In terms of the future, Friedrich Rippmann sees the greatest opportunity as drug discovery and attrition rates being reduced. This would result in more novel drugs being available to patients faster. AI has great prospects for drug discovery, but for now, the cost of implementing it is prohibitive. Increasing competition will bring down costs, opening up exciting possibilities for discoveries in a variety of fields [91-100].

CONCLUSION

The newfound interest in explainable AI, with methodologies such as feature attribution, instance-based molecular counterfactual explanation, and uncertainty estimation, will increase the acceptance of AI-supported drug discovery. AI is still far from perfect, and there are still areas in which data cannot be analyzed, which makes combining human and machine intelligence an effective strategy. The current trend suggests that these methods will be increasingly accessible in the coming future, with the continued development of general high-level research and deployment software packages, as well as comprehensible documentation. Recently, AI has captured the attention of many people and has been extensively used in drug discovery. It is possible to expect improvement in the future if the chemical representation can be adapted to the problem at hand. AI for drug discovery is strongly benefited from open source implementations, which allow access to software libraries that allow neural networks to be implemented. The development of robotics will be particularly vital to this progress.

AUTHORS' CONTRIBUTIONS

We hereby declare that all the authors contributed equally in preparing and finalizing this review manuscript.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

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;26:80.
2. Fleming N. How artificial intelligence is changing drug discovery. Nature 2018;557:S55.
3. Moingeon P, Kuenemann M, Guedj M. Artificial intelligence-enhanced

drug design and development: Toward a computational precision medicine. Drug Discov Today 2021;27:215-22.
4. Vyas M, Thakur S, Riyaz B, Bansal KK, Tomar B, Mishra V. Artificial intelligence: The beginning of a new era in the pharmacy profession. Asian J Pharm 2018;12:72-6.
5. Yang X, Wang Y, Byrne R, Schneider G, Yang S. Concepts of artificial intelligence for computer-assisted drug discovery. Chem Rev 2019;119:10520-94.
6. Nelson SD, Walsh CG, Olsen CA, McLaughlin AJ, LeGrand JR, Schutz N, *et al*. Demystifying artificial intelligence in pharmacy. Am J Health Syst Pharm 2020;77:1556-70.
7. Piroozmand F, Mohammadipanah F, Sajedi H. Spectrum of deep learning algorithms in drug discovery. Chem Biol Drug Des. 2020;96(3):886-901.
8. Specht DF. IEEE Int Conf Neur Netk 1988;1:525-32.
9. Gupta J. The Power of Artificial Intelligence in Drug Discovery. INDIAai; 2021.
10. Domenico A, Nicola G, Daniela T, Fulvio C, Nicola A, Orazio N. *De novo* drug design of targeted chemical libraries based on artificial intelligence and pair-based multiobjective optimization. J Chem Inform Model 2020;60:4582-93.
11. Brown N, Ertl P, Lewis R, Luksch T, Reker D, Schneider N. Artificial intelligence in chemistry and drug design, J Comput Aided Mol Des. 2020;34(7):709-715.
12. Grisoni F, Huisman BJ, Button AL, Moret M, Atz K, Merk D, *et al*. Combining generative artificial intelligence and on-chip synthesis for *De novo* drug design. Sci Adv 2021;7:eabg3338.
13. Bai J, Cao L, Mosbach S, Akroyd J, Lapkin AA, Kraft M. From Platform to Knowledge Graph: Evolution of Laboratory Automation, JACS Au. 2022;2:292-309.
14. Todd MH. Computer-Aided Organic Synthesis. Chem Soc Rev. 2005:34:247–266.
15. Shen Y, Borowski JE, Hardy MA, Sarpong R, Doyle AG, Cernak T. Automation and computer-assisted planning for chemical synthesis. Nat Rev Methods Prim 2021;1:1-23.
16. Dara S, Dhamercherla S, Jadav SS, Babu CH, Ahsan MJ. Machine learning in drug discovery: A review. Artif Intell Rev 2021;1-53.
17. Brogi S, Ramalho TC, Kuca K, Medina-Franco JL, Valko M. *In silico* methods for drug design and discovery. Front Chem 2020;8:612.
18. Hessler G, Baringhaus KH. Artificial intelligence in drug design. Molecules 2018;23:2520.
19. Yee LC, Wei YC. Current modeling methods used in QSAR/QSPR. Assessment 2012;10:11.
20. Ramesh AN, Kambhampati C, Monson JR, Drew PJ. Artificial intelligence in medicine. Ann R Coll Surg Engl 2004;86:334.
21. Arkan E, Shahlaei M, Pourhossein A, Fakhri K, Fassihi A. Validated QSAR analysis of some diaryl substituted pyrazoles as CCR2 inhibitors by various linear and nonlinear multivariate chemometrics methods. Eur J Med Chem 2010;45:3394-406.
22. Deeb O, Goodarzi M. Exploring QSARs for inhibitory activity of non-peptide HIV-1 protease inhibitors by GA-PLS and GA-SVM. Chem Biol Drug Design 2010;75:506-14.
23. Mishra V. Artificial intelligence: The beginning of a new era in pharmacy profession. Asian J Pharm 2018;12:72-6.
24. Flynn A. Using artificial intelligence in health-system pharmacy practice: Finding new patterns that matter. Am J Health Syst Pharm 2019;76:622-7.
25. Hamet P, Tremblay J. Artificial intelligence in medicine. Metabolism 2017;69:S36-40.
26. Challen R, Denny J, Pitt M, Gompels L, Edwards T, Tsaneva-Atanasova K. Artificial intelligence, bias and clinical safety. BMJ Qual Saf 2019;28:231-7.
27. Kubinyi H. Strategies and recent technologies in drug discovery. Die Pharm 1995;50:647-62.
28. Chan HS, Shan H, Dahoun T, Vogel H, Yuan S. Advancing drug discovery via artificial intelligence. Trends Pharmacol Sci 2019;40:592-604.
29. Díaz Ó, Dalton JA, Giraldo J. Artificial intelligence: A novel approach for drug discovery. Trends Pharmacol Sci 2019;40:550-1.
30. Jiménez-Luna J, Grisoni F, Weskamp N, Schneider G. Artificial intelligence in drug discovery: Recent advances and future perspectives. Expert Opin Drug Discov 2021;16:949-59.
31. Mak KK, Pichika MR. Artificial intelligence in drug development: Present status and future prospects. Drug Discov Today 2019;24:773-80.
32. Kapetanovic IM. Computer-aided drug discovery and development (CADDD): *In silico*-chemico-biological approach. Chemicobiol Interact 2008;171:165-76.
33. Clark DE. What has computer-aided molecular design ever done for drug discovery? Expert Opin Drug Discov 2006;1:103-10.
34. Bajorath J. Computer-aided drug discovery. F1000Res 2015;4:630.
35. Zhao L, Ciallella HL, Aleksunes LM, Zhu H. Advancing computer-aided drug discovery (CADD) by big data and data-driven machine learning modeling. Drug Discov Today 2020;25:1624-38.
36. Tang Y, Zhu W, Chen K, Jiang H. New technologies in computer-aided drug design: Toward target identification and new chemical entity discovery. Drug Discov Today Technol 2006;3:307-13.
37. Xiang M, Cao Y, Fan W, Chen L, Mo Y. Computer-aided drug design: Lead discovery and optimization. Combinator Chem High Through Screen 2012;15:328-37.
38. Song CM, Lim SJ, Tong JC. Recent advances in computer-aided drug design. Brief Bioinform 2009;10:579-91.
39. Tropsha A. Recent trends in computer-aided drug discovery. Curr Opin Drug Discov Dev 2000;3:310-3.
40. Ebhohimen IE, Edemhanria L, Awojide S, Onyijen OH, Anywar G. Advances in computer-aided drug discovery. In: Phytochemicals as Lead Compounds for New Drug Discovery. Netherlands: Elsevier; 2020. p. 25-37.
41. Singh B, Mal G, Gautam SK, Mukesh M. Computer-aided drug discovery. In: Advances in Animal Biotechnology. Cham: Springer; 2019. p. 471-81.
42. Tantillo DJ, Siegel JB, Saunders CM, Palazzo TA, Painter PP, O'Brien TE, *et al*. Computer-aided drug design for undergraduates. J Chem Educ 2019;96:920-5.
43. Imam SS, Gilani SJ. Computer aided drug design: A novel loom to drug discovery. Organ Med Chem Int J 2017;1:113-8.
44. Mohs RC, Greig NH. Drug discovery and development: Role of basic biological research. Alzheimer's Dement 2017;3:651-7.
45. Lake F. Artificial intelligence in drug discovery: What is new, and what is next? Fut Drug Discov 2019;1:FDD19.
46. Torjesen I. Drug development: The journey of a medicine from lab to shelf. Pharm J 2015.
47. Xu Z, Wauchope OR, Frank AT. Navigating chemical space by interfacing generative artificial intelligence and molecular docking. J Chem Inform Model 2021;61:5589-600.
48. Jing Y, Bian Y, Hu Z, Wang L, Xie XQ. Deep learning for drug design: An artificial intelligence paradigm for drug discovery in the big data era. AAPS J 2018;20:1.
49. Schneider P, Walters WP, Plowright AT, Sieroka N, Listgarten J, Goodnow RA, *et al*. Rethinking drug design in the artificial intelligence era. Nat Rev Drug Discov 2020;19:353-64.
50. Méndez-Lucio O, Baillif B, Clevert DA, Rouquié D, Wichard J. *De novo* generation of hit-like molecules from gene expression signatures using artificial intelligence. Nat Commun 2020;11:1.
51. Saikin SK, Kreisbeck C, Sheberla D, Becker JS, Aspuru-Guzik A. Closed-loop discovery platform integration is needed for artificial intelligence to make an impact in drug discovery. Expert Opin Drug Discov 2019;14:1-4.
52. Hassanzadeh P, Atyabi F, Dinarvand R. The significance of artificial intelligence in drug delivery system design. Adv Drug Deliv Rev 2019;151:169-90.
53. Meyers J, Fabian B, Brown N. *De novo* molecular design and generative models. Drug Discov Today 2021;26:2707-15.
54. Cavasotto CN, Di Filippo JI. Artificial intelligence in the early stages of drug discovery. Arch Biochem Biophys 2021;698:108730.
55. Colombo S. Applications of artificial intelligence in drug delivery and pharmaceutical development. In: Artificial Intelligence in Healthcare. United States: Academic Press; 2020. p. 85-116.
56. Bohr H. Drug discovery and molecular modeling using artificial intelligence. In: Artificial Intelligence in Healthcare. United States: Academic Press; 2020. p. 61-83.
57. Lo YC, Ren G, Honda H, Davis KL. Artificial intelligence-based drug design and discovery. Cheminformatics and its Applications. India: Intechopen; 2019.
58. Duch W, Swaminathan K, Meller J. Artificial intelligence approaches for rational drug design and discovery. Curr Pharm Design 2007;13:1497-508.
59. Archer M, Germain S. The integration of artificial intelligence in drug discovery and development. Int J Digit Health 2021;1:31.
60. Mannan HA. Creating efficiencies in drug discovery and development using artificial intelligence. Int House Counsel J 2017;11:1.
61. Selvaraj G, Kaliamurthi S, Peslherbe GH, Wei DQ. Application of artificial intelligence in drug repurposing: A mini-review. Curr Chine Sci 2021;1:333-345.

62. Chen H, Engkvist O, Wang Y, Olivecrona M, Blaschke T. The rise of deep learning in drug discovery. Drug Discov Today 2018;23:1241-50.
63. Bajorath J, Kearnes S, Walters WP, Meanwell NA, Georg GI, Wang S. Artificial intelligence in drug discovery: Into the great wide open. J Med Chem 2020;63:8651-2.
64. Sandfort F, Strieth-Kalthoff F, Kühnemund M, Beecks C, Glorius F. A structure-based platform for predicting chemical reactivity. Chem 2020;6:1379-90.
65. Jacobs I, Maragoudakis M. *De novo* drug design using artificial intelligence applied on SARS-CoV-2 viral proteins ASYNT-GAN. BioChem 2021;1:36-48.
66. Jacobs I, Maragoudakis M. *De novo* Drug Design using Artificial Intelligence ASYNT-GAN.2020.
67. Wang Z, Zhao W, Hao G, Song B. Automated synthesis: Current platforms and further needs. Drug Discov Today 2020.
68. Struble TJ, Alvarez JC, Brown SP, Chytil M, Cisar J, DesJarlais RL, *et al*. Current and future roles of artificial intelligence in medicinal chemistry synthesis. J Med Chem 2020;63:8667-82.
69. Thakkar A, Johansson S, Jorner K, Buttar D, Reymond JL, Engkvist O. Artificial intelligence and automation in computer aided synthesis planning. React Chem Eng 2021;6:27-51.
70. Godfrey AG, Masquelin T, Hemmerle H. A remote-controlled adaptive medchem lab: An innovative approach to enable drug discovery in the 21st century. Drug Discov Today 2013;18:795-802.
71. Bijral RK, Singh I, Manhas J, Sharma V. Exploring artificial intelligence in drug discovery: A comprehensive review. Arch Comput Methods Eng 2021;25:1-7.
72. Mak KK, Balijepalli MK, Pichika MR. Success stories of AI in drug discovery-where do things stand? Expert Opin Drug Discov 2021;17:79-92.
73. Dana D, Gadhiya SV, St Surin LG, Li D, Naaz F, Ali Q, *et al*. Deep learning in drug discovery and medicine; scratching the surface. Molecules 2018;23:2384.
74. Batool M, Ahmad B, Choi S. A structure-based drug discovery paradigm. Int J Mol Sci 2019;20:2783.
75. Reddy AS, Zhang S. Polypharmacology: Drug discovery for the future. Expert Rev Clin Pharmacol 2013;6:41-7.
76. Boran AD, Iyengar R. Systems approaches to polypharmacology and drug discovery. Curr Opin Drug Discov Dev 2010;13:297.
77. Anighoro A, Bajorath J, Rastelli G. Polypharmacology: Challenges and opportunities in drug discovery: Miniperspective. J Med Chem 2014;57:7874-87.
78. Lavecchia A. Deep learning in drug discovery: Opportunities, challenges and future prospects. Drug Discov Today 2019;24:2017-32.
79. Malandraki-Miller S, Riley PR. Use of artificial intelligence to enhance phenotypic drug discovery. Drug Discov Today 2021;26:887-901.
80. Collins BN. AI and Automation-accelerating Synthetic Chemistry and Streamlining Drug Development. 2020.
81. Elbadawi M, Gaisford S, Basit AW. Advanced machine-learning techniques in drug discovery. Drug Discov Today 2021;26:769-77.
82. Gurgula O. AI-assisted inventions in the field of drug discovery: Readjusting the inventive step analysis. Int J Soc Sci Public Policy; 2020. Available from: https://www.ssrn.com/abstract=3683127
83. Patel L, Shukla T, Huang X, Ussery DW, Wang S. Machine learning methods in drug discovery. Molecules 2020;25:5277.
84. Eisenstein M. AI in drug discovery starts to live up to the hype: Machine learning-based techniques are finding a place in early-stage target discovery and in drug development workflows. Genet Eng Biotechnol News 2021;41:38-40.
85. Vamathevan J, Clark D, Czodrowski P, Dunham I, Ferran E, Lee G, *et al*. Applications of machine learning in drug discovery and development. Nat Rev Drug Discov 2019;18:463-77.
86. Smith JS, Roitberg AE, Isayev O. Transforming computational drug discovery with machine learning and AI. ACS Med Chem Lett 2018;9:1065-9.
87. Jiménez-Luna J, Grisoni F, Schneider G. Drug discovery with explainable artificial intelligence. Nat Mach Intell 2020;2:573-84.
88. Cai C, Wang S, Xu Y, Zhang W, Tang K, Ouyang Q, *et al*. Transfer learning for drug discovery. J Med Chem 2020;63:8683-94.
89. Hochreiter S, Klambauer G, Rarey M. Machine learning in drug discovery. J Chem Inform Model 2018;58:1723-4.
90. Zhavoronkov A. Artificial intelligence for drug discovery, biomarker development, and generation of novel chemistry. Mol Pharm 2018;15:4311-3.
91. Zhavoronkov A, Vanhaelen Q, Oprea TI. Will artificial intelligence for drug discovery impact clinical pharmacology? Clin Pharmacol Ther 2020;107:780-5.
92. Schneider G. Mind and machine in drug design. Nat Mach Intell 2019;1:128-30.
93. Vijayan RS, Kihlberg J, Cross JB, Poongavanam V. Enhancing preclinical drug discovery with artificial intelligence. Drug Discov Today 2021.
94. Chen Z, Liu X, Hogan W, Shenkman E, Bian J. Applications of artificial intelligence in drug development using real-world data. Drug Discov Today 2021;26:1256-64.
95. Lamberti MJ, Wilkinson M, Donzanti BA, Wohlhieter GE, Parikh S, Wilkins RG, *et al*. A study on the application and use of artificial intelligence to support drug development. Clin Ther 2019;41:1414-26.
96. Álvarez-Machancoses Ó, Fernández-Martínez JL. Using artificial intelligence methods to speed up drug discovery. Expert Opin Drug Discov 2019;14:769-77.
97. Liu Z, Roberts RA, Lal-Nag M, Chen X, Huang R, Tong W. AI-based language models powering drug discovery and development. Drug Discov Today 2021;26:2593-607.
98. Patel V, Shah M. A comprehensive study on artificial intelligence and machine learning in drug discovery and drug development. Intel Med 2021.
99. Jordan AM. Artificial intelligence in drug design-the storm before the calm? ACS Med Chem Lett 2018;9:1150-2.
100. Sellwood MA, Ahmed M, Segler MH, Brown N. Artificial intelligence in drug discovery. Fut Med Chem 2018;10:2025-8.