German J Pharm Biomaterials 2022; 1(2): 1-3
https://doi.org/10.5530/gjpb.2022.2.6

GJPB

Editorial

Machine learning empowered drug discovery

Thirumoorthy Durai Ananda Kumar[1]* and Naraparaju Swathi[2]

[1]Department of Pharmaceutical Chemistry, JSS College of Pharmacy, Mysuru, JSS Academy of Higher Education & Research, Mysuru, India.
[2]Department of Pharmaceutical Chemistry, Gokaraju Rangaraju College of Pharmacy, Hyderabad, India.

*Correspondence: anandurai78@gmail.com, swa.pharma@gmail.com

Traditional drug discovery strategies include lead molecule identification, lead optimization, preclinical studies and clinical trials [1]. The pharmaceutical and biotechnology research and development (R&D) department spends more than 10 years and $1 billion to bring the molecule to market [2] successfully. About 90% of drug candidates fail in the drug development due to safety and efficacy issues [3]. The lack of technologies is the main limitation for identifying potential candidates from the available chemical space (>10^{60} molecules).

De Novo design methods explore chemical space through pharmacophore (ligand-based), and docking (structure-based) approaches [4]. Structure-based drug discovery approaches use the insights gained from biological data of target structures. Schrödinger, AutoDock and Biovia (Accelrys) pioneered the development of structure-based tools to improve drug discovery. Libraries of molecules can be screened for their target suitability, known as virtual screening [5]. The structure-based drug discovery approach uses the three-dimensional (3D) details of the target structure and explains the intermolecular interactions (biophysical simulations) [6].

Ligand-based drug discovery approaches are based on in vitro test data from several hundred molecules. The availability of an extensive database of experimental biological activity facilitates the development of ligand-based predictive models (QSAR models) for drug discovery. The inefficiency of computational tools, high cost and inaccessibility of the target's three-dimensional (3D) details are the main disadvantages and drastically reduce the success rate. With the advancement in data science and artificial intelligence techniques, QSAR-based analysis is enhanced by adopting artificial intelligence methods.

Artificial intelligence

Artificial intelligence (AI) and data science applications in drug discovery are reducing research timeframes and costs. AI technologies support the rapid validation of target molecules and the optimization of lead molecules [7]. The adoption of AI in drug discovery programs has increased during the COVID-19 pandemic [8]. It is predicted that AI enabled computer modelling and simulation will dominate drug discovery strategies by 2030. The size of the global AI market has been increasing in recent years and is expected to reach $9.1 billion by 2030. The expected compound annual growth rate (CAGR) for the period 2022 to 2030 is 29.4% [9]. More than 120 pharmaceutical companies are using AI solutions for their drug discovery programs.

Computer algorithms speed up research results and reduce the multi-year process to just a few years for drug discovery [10]. The quick results bring in potential candidates in less time and help reduce the cost of the products making the product more affordable. AI has multiple applications in drug discovery, ranging from drug design, drug screening, chemical synthesis, and polypharmacology to drug repurposing [11].

Machine learning is the subset of AI and has significantly improved drug discovery programs [12]. A computational ability to analyze without explicit programming is called machine learning (ML). ML model identifies the active molecules from the chemical library. ML applications are based on supervised, unsupervised, semi-supervised and reinforcement learning algorithms. ML in drug discovery can facilitate bringing the molecules to market much earlier at a lower price. BenevolentAI, a

machine learning-based computational and experimental platform, supports the drug discovery process from target identification to clinical trials [13]. BioIVT, an ML-based clinical database mining tool, supports ADMET predictions, biomarker discovery and precision medicine development [14].

Applications

AI is used extensively for drug discovery, diagnosis, clinical trial prediction and patient adherence monitoring.

Molecular pharmacokinetics

Molecular pharmacokinetic (ADMET) properties determine their tissue selectivity. AI predicts off-target effects, such as effects on ion channels (e.g., hERG), well in advance during lead optimization [15]. Advanced ML-based models (RF, SVM, Naïve Bayesian and ANN) provide insights into molecular dynamics and form the basis for developing the branch of so-called quantitative systems pharmacology [16]. The RF algorithm is considered the gold standard for ML-supported ADMET predictions [17]. eToxPred, an ML-based algorithm, helps predict molecules' synthetic feasibility and toxicity with high accuracy [18].

Molecular polymorphism

Genome polymorphism (race, ethnicity) has profound implications for candidate efficacy. AI models predict the effect of polymorphism and provide data on possible intended/unintended effects [19]. AI tools have been developed to select clinical trial participants based on the genomic markers, which helps overcome genomic heterogeneity.

Drug discovery

BenevolentAI, an ML-based application, is used in drug discovery for amyotrophic lateral sclerosis (ALS), parkinsonism, ulcerative colitis and sarcopenia. BenevolentAI algorithm identifies lead molecules to treat ALS [20]. The Sheffield Institute for Translational Neuroscience (SITraN) investigated the therapeutic potential of identified molecules in the gold standard disease model. BenevolentAI established collaboration with AstraZeneca and Novartis Pharma AG on molecule discovery for chronic kidney disease (CKD), idiopathic pulmonary fibrosis (IPF) and cancer. Self-organizing maps (SOM) and SVM models have been used in the validation of EGFR (ERbB-1/HER1) kinase inhibitory activity [21]. Exscientia, in collaboration with Celgene, GlaxoSmithKline plc (GSK), Roche, Sanofi and Shanghai Biotech Company GT Apeiron, is working on drugability and target specificity and phenotype drug design through AI tools. Exscientia-Sanofi collaborative AI-assisted research identified a novel molecule for fibrosis and is in clinical trials [22]. Affinity2Vec, a KronRLS (Kronecker-regularized least squares) based algorithm developed to assess the drug-target binding affinity (DTBA) [23].

Outlook

The merging of ligand-based ML techniques with structure-based modelling should further accelerate drug discovery. The research outcomes will significantly impact the discovery of effective therapeutic tools for the diseases that remain untreated for decades. ML-based research facilitates effective decisions for researchers and accelerates the drug discovery process. The standards of ML algorithms are insufficient and require a critical evaluation of an ML model. The lack of quality and data with high dimensionality are the main challenges.

References

1. Hughes JP, Rees S, Kalindjian SB, Philpott KL. Principles of early drug discovery. Br J Pharmacol. 2011;162(6):1239-49.
2. Hingorani AD, Kuan V, Finan C, Kruger FA, Gaulton A, Chopade S, et al. Improving the odds of drug development success through human genomics: modelling study. Sci Rep. 2019;9(1):18911.
3. Sun D, Gao W, Hu H, Zhou S. Why 90% of clinical drug development fails and how to improve it? Acta Pharm Sin B. Forthcoming 2022.

4. Vázquez J, López M, Gibert E, Herrero E, Luque FJ. Merging ligand-based and structure-based methods in drug discovery: an overview of combined virtual screening approaches. Molecules. 2020;25(20):4723.
5. Stumpfe D, Ripphausen P, Bajorath J. Virtual compound screening in drug discovery. Future Med Chem. 2012;4(5):593-602.
6. Maia EHB, Assis LC, de Oliveira TA, da Silva AM, Taranto AG. Structure-based virtual screening: from classical to artificial intelligence. Front Chem. 2020;8. 343.
7. Paul D, Sanap G, Shenoy S, Kalyane D, Kalia K, Tekade RK. Artificial intelligence in drug discovery and development. Drug Discov Today. 2021;26(1):80-93.
8. Suryanarayanan P, Tsou CH, Poddar A, Mahajan D, Dandala B, Madan P, et al. AI-assisted tracking of worldwide non-pharmaceutical interventions for COVID-19. Sci Data. 2021;8(1):94.
9. Research and Markets [Internet]. Artificial intelligence in drug discovery market size, share & trends analysis report by application (drug optimization & repurposing, preclinical testing), by therapeutic area, by region, and segment forecasts, 2022-2030. [cited 2022 June 11]. Available from: https://www.researchandmarkets.com/reports/5595828/artificial-intelligence-in-drug-discovery-market?
10. Dara S, Dhamercherla S, Jadav SS, Babu CHM, Ahsan MJ. Machine learning in drug discovery: A Review. Artif Intell Rev. 2022;55(3):1947-99.
11. Selvaraj G, Kaliamurthi S, Peslherbe HG, Wei D-Q. Application of artificial intelligence in drug repurposing: A mini-review. Curr Chinese Sci. 2021;1(3):333-45.
12. Sarker IH. Machine Learning: Algorithms, Real-World Applications and Research Directions. SN Comput Sci. 2021;2(3):160.
13. BenevolentAI [Internet]. BenevolentAI Drug Programmes. [cited 2022 June 11]. Available from: https://www.benevolent.com/benevolentai-drug-programmes
14. BioIVT [Internet]. Applications. [cited 2022 June 11]. Available from: https://bioivt.com/applications
15. Ietswaart R, Arat S, Chen AX, Farahmand S, Kim B, DuMouchel W, et al. Machine learning guided association of adverse drug reactions with in vitro target-based pharmacology. EBioMedicine. 2020;57:102837.
16. Aghamiri SS, Amin R, Helikar T. Recent applications of quantitative systems pharmacology and machine learning models across diseases. J Pharmacokinet Pharmacodyn. 2022;49(1):19-37.
17. Vatansever S, Schlessinger A, Wacker D, Kaniskan HÜ, Jin J, Zhou MM, et al. Artificial intelligence and machine learning-aided drug discovery in central nervous system diseases: State-of-the-arts and future directions. Med Res Rev. 2021;41(3):1427-73.
18. Pu L, Naderi M, Liu T, Wu HC, Mukhopadhyay S, Brylinski M. eToxPred: a machine learning-based approach to estimate the toxicity of drug candidates. BMC Pharmacol Toxicol. 2019;20(1):2.
19. Heng T, Yang D, Wang R, Zhang L, Lu Y, Du G. Progress in Research on Artificial Intelligence Applied to Polymorphism and Cocrystal Prediction. ACS Omega. 2021;6(24):15543-50.
20. BenevolentAI [Internet]. BenevolentAI Drug Pipeline Amyotrophic Lateral Sclerosis. [cited 2022 June 12]. Available from: https://www.benevolent.com/amyotrophic-lateral-sclerosis
21. Kong Y, Qu D, Chen X, Gong YN, Yan A. Self-organizing map (SOM) and support vector machine (SVM) models for the prediction of human epidermal growth factor receptor (EGFR/ ErbB-1) inhibitors. Comb Chem High Throughput Screen. 2016;19(5):400-11.
22. Kirkpatrick P. Artificial intelligence makes a splash in small-molecule drug discovery. Biopharma dealmakers. 2022 May 17. Available from: https://www.nature.com/articles/d43747-022-00104-7
23. Thafar MA, Alshahrani M, Albaradei S, Gojobori T, Essack M, Gao X. Affinity2Vec: drug-target binding affinity prediction through representation learning, graph mining, and machine learning. Sci Rep. 2022;12(1):4751

How to cite this article:
Kumar TDA, Swathi N. Machine learning empowered drug discovery. German J Pharm Biomaterials. 2022;1(2):1-3.