Computational Insights into Drug Repurposing: A Brief Review

Surajit Bal, Rathi Dasgupta[*]

Biology Department, Embibe Individual Learning Limited, Bangalore, India

ABSTRACT

The COVID-19 pandemic has triggered a surge in the reutilization of established medications, albeit the underlying evidentiary foundation varies. Drug reutilization involves repurposing known drugs or drug combinations for unforeseen medical contexts. Consequently, it plays a crucial role in expediting the pre-clinical phase of developing novel drugs, offering time and cost savings compared to traditional drug discovery. Given that drug reuse relies on extensive data from existing drugs and diseases, the substantial proliferation of publicly accessible, large-scale machine learning techniques represents cutting-edge data science applications for disease identification, therapeutics and target discovery, minimizing errors. In this write-up, we have explored different computational databases, techniques and methodologies, as well as possibilities, for harnessing machine learning methods to expedite the process of drug repurposing.

Keywords: Disease identification; Computational databases; Drug repurposing; Drug reutilization

INTRODUCTION

The drug discovery process, before the advent of Computer-Aided Drug Design (CADD), was a lengthy, arduous and expensive endeavor. It often took decades and billions of dollars to bring a single drug to market. This traditional process relied heavily on serendipity, animal testing and trial and error, resulting in a low success rate and high attrition. The development of CADD in the mid-1970s revolutionized the drug discovery process, offering a more efficient and rational approach to identifying and developing new drugs. Drug repurposing, also known as drug repositioning, is another strategy that has gained traction in recent years. It involves identifying new uses for existing drugs that have already been approved for other indications. The integration of CADD and drug repurposing strategies has significantly accelerated the drug discovery process, leading to the development of new and more effective therapies for various diseases.

LITERATURE REVIEW

Computer-aided drug designing

The process of drug development and design is not only resource-intensive but also costly. To mitigate these challenges, computers are harnessed for drug design, a practice known as *in silico* drug design.

Numerous tools and software applications have been created to conduct docking studies, utilizing parameters to fine-tune the drug and target structures for desired outcomes [1-4]. CADD represents an interdisciplinary domain that combines molecular biology, chemistry, biochemistry, immunology, pharmacology, nanotechnology, data science and informatics.

Drug repurposing

Drug repurposing, also known as drug repositioning, involves a mathematical approach to pinpointing established, archived or discontinued drugs, as well as those under clinical investigation, to offer treatment solutions for various diseases [5-7]. *De novo* strategy in drug discovery comes with prohibitive costs and time constraints. In contrast, drugs with established mechanisms of action and pharmacokinetics provide a valuable domain-specific knowledge base. Identifying the potential benefits of these known drugs, which are both effective and safe, eliminates the need to start from scratch. In such instances, drug repurposing incurs significantly reduced time and economic costs. Drug repositioning proves to be a valuable and efficient method, allowing the application of a drug approved for one medical condition to potentially treat another ailment, provided that there is a degree of structural, side chain or side effect similarity [8]. AI-supported drug repositioning further contributes to time and cost reduction.

The databases, tools and their contributions to drug designing and repurposing

To date, numerous databases have been created, facilitating the repurposing of drugs and playing a pivotal role in this endeavor. Beyond these databases, tools have been developed, leveraging the data and are instrumental in the drug repurposing process. This discussion will delve into these tools, examining their features and contents. In this paper, we present data containing a comprehensive list of databases and tools utilized, their web links, brief descriptions and references (Table 1,2) [9-61].

Table 1: The investigated databases for the repurposing of drugs.

Databases	Weblink	Description	Reference
Compound Muscle Action Potential (CMAP)	http://www.complement.us/cmap	CMAP includes experimental data curated by experts.	[9], [10]
Absorption, Distribution, Metabolism and Excretion (ADME)	https://www.fujitsu.com/global/solutions/	An online database for pharmacokinetic information.	[42], [45]
Drug repurposing hub	https://clue.io/repurposing	It has created a huge library including repurposed drugs, news, tools, etc.	[9], [11]
The Health Improvement Network (THIN)	https://www.ucl.ac.uk/	THIN has collected results related to drugs and diseases which are reported in different studies.	[9], [12]
Protein Data Bank (PDB)	https://www.rcsb.org/	The protein data bank archive provides information about the 3D structure of proteins, nucleic acids and complex assemblies.	[42], [50]
UniProt	https://www.uniprot.org/	An open resource of protein sequences and functional information.	[42], [51]
Atom3D	https://github.com/drorlab/atom3d	A benchmark of existing datasets of 3D molecules, spanning several types.	[42]
MoleculeNet	https://moleculenet.org/	A benchmark of datasets for molecular machine learning.	[42], [52]
Drug Signatures Database (DSigDB)	http://tanlab.ucdenver.edu/DSigDB	It has been developed to hold drugs or compounds and their gene targets.	[9], [13]
Drug2Gene	http://www.drug2gene.com	Drug2Gene integrates drug-target information from 19 popular databases.	[9], [14]
DrugBank	https://www.drugbank.ca/	DrugBank provides a comprehensive database of drugs, their different targets, their 3D structure them and other useful information.	[9], [15]
Drug-path	http://www.cuilab.cn/drugpath	DPTH contains pathways that are induced by drugs.	[9], [16]
ZINC	https://zinc.docking.org/	An open resource for virtual screening of compounds.	[42], [43]
Binding Database	https://www.bindingdb.org/bind/index.jsp	A database of measuring binding affinity between the target and the drug.	[42], [44]
GeneSet Database	http://genesetdb.auckland.ac.nz/haeremai.html	GSDB is an integrated meta-database that interprets a list of genes.	[9], [17]
Potential Drug Target Database (PDTD)	http://www.dddc.ac.cn/pdtd/	PDTD is an integrated database for the identification of protein targets.	[9], [18]
Search Tool for Interactions of Chemicals (STITCH)	http://stitch.embl.de/	An integrated database of chemical-protein interactions.	[42], [46]
Side Effect Resource (SIDER)	http://sideeffects.embl.de	SIDER contains the main targets of drugs and their side effects.	[9], [19]
The Genomics of Drug Sensitivity in Cancer (GDSC)	https://www.cancerrxgene.org/	Drug response data and genomic biomarkers.	[42], [47]
Protein Data Bank Bind Database (PDBBind)	http://www.pdbbind.org.cn/	A comprehensive collection of binding affinities for the protein–ligand complexes in the Protein Data Bank (PDB).	[42], [48]

CanSar	https://cansarblack.icr.ac.uk/	Cancer translational research and drug discovery knowledgebase.	[42], [49]
Therapeutic target database	http://bidd.nus.edu.sg/group/ttd/ttd.asp	It has been developed for drug discovery based on target information.	[9], [20]
Tuberculosis drug resistance	http://www.tbdb.org	For the repurposing applications, TBDB including gene expression data and their annotations has been extended.	[9], [21]
DRUGSURV	http://www.bioprofiling.de/drugsurv	DRUGSURV holds information on experimentally repurposed drugs for oncology.	[9], [22]
HIVDB	https://hivdb.stanford.edu/DR/	HIVDB covers information on the drug resistance of Human Immunodeficiency Virus (HIV).	[9], [23]
Ontario	www.healthinfo.moh.gov.on.ca	Ontario Drug Benefit (ODB) involves described drugs and their claimed indications in Ontario.	[9], [24]
SuperCYP	http://bioinformatics.charite.de/supercyp	Super CYP is a comprehensive database for Cytochrome P450 (CYP) along with a tool to predict drugs that can interact with CYP.	[9], [25]
Traditional Chinese Medicine (TCM)	http://tcm.cmu.edu.tw/	TCM has been developed in Taiwan and consists of in-silicon drug screening of traditional medicine.	[9], [26]
Traditional Chinese Medicine Systems Pharmacology Database and Analysis Platform (TCMSP)	http://sm.nwsuaf.edu.cn/lsp/tcmsp.php	TCMSP is a database for drug discovery from herbal medicine.	[9], [27]
Tropical Disease Research (TDR) Targets Database	http://tdrtargets.org	TDR comprises tropical disease information and can offer a list of genes for repositioning applications.	[9], [28]
Cancer HSP	http://lsp.nwsuaf.edu.cn/CancerHSP.php	Anti-cancer herbs along with their molecular information are available in CHP.	[9], [29]
PubChem	https://pubchem.ncbi.nlm.nih.gov/	The largest collection of freely accessible chemical and bio-activity information.	[9], [41]
CheMBL	https://www.ebi.ac.uk/chembldb/	CheMBL consists of a large number of drug-like compounds for drug discovery applications.	[9], [30]
Drug-directionality Map (DMAP)	http://bio.informatics.iupui.edu/cmaps	To obviate the data limitation of CMAP, DMAP has been designed.	[9], [31]
Swiss BIOisostere	http://www.swissbioisostere.ch	Swiss BIOisostere incorporates information on molecular replacement and their biological effects.	[9], [33]
DRAR-CPI	http://cpi.bio-x.cn/drar/	Using the adverse reaction of drugs on proteins, DRAR predicts their interaction based on the collection of drug molecules.	[9], [34]
DrugNet	http://genome2.ugr.es/drugnet/	DrugNet integrates heterogeneous data and prioritizes the effects of drugs on diseases.	[9], [36]
e-Drug3D	http://chemoinfo.ipmc.cnrs.fr/e-drug3d	It contains the 3D chemical structure of drugs and fragment-based information.	[9], [39]
NFFinder	http://nffinder.cnb.csic.es	NNFIN searches similar transcriptome data in different contexts.	[8], [9]

Table 2: The investigated tools for the repurposing of drugs.

Tools	Weblink	Description	Reference
Drug-Target Interactome (DTome)	http://bioinfo.mc.vanderbilt.edu/DTome	It is a tool that helps to understand the molecular mechanism of drug bioactivities.	[9], [37]
KsRepo	http://github.com/adam-sam-brown/ksRepo	KSRPO encompasses a tool for predicting drug-target interactions at the gene level.	[9], [40]
Drug-Map Central (DMC)	http://r2d2drug.org/DMC.aspx	DMC includes a tool that integrates multi data from various sources and proposes a list for drug repositioning studies.	[9], [35]
Database Technology Web (DT-Web)	http://alpha.dmi.unict.it/dtweb	Based on DrugBank data, it provides a network-based drug-target prediction tool.	[9], [38]
Assisted Model Building with Energy Refinement (AMBER)	https://ambermd.org/	It is a package for molecular dynamics simulation.	[42]
Accelerated Molecular Dynamics (ACEMD)	https://www.acellera.com/	An accelerated platform for faster and longer bio-molecular simulations.	[42], [53]
AutoDock Vina	https://vina.scripps.edu/	A program for molecular docking and screening.	[42], [54]
DeePMD	https://github.com/deepmodeling/deepmd-kit/	It is a deep-learning package for molecular docking simulation and energy representation.	[42], [55]
RBio3D	http://thegrantlab.org/bio3d/	R package for the analysis of molecular docking trajectories.	[42], [56]
Pymol	https://pymol.org/2/	An interactive platform for visualization of molecules.	[42], [57]
Rosetta Commons	https://www.rosettacommons.org/	A tool for predicting the mutant structure.	[42], [58]
Anni 2.0	https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2481428/	It interprets differentially expressed genes and literature-based knowledge discovery.	[1]
Balestra Web	http://balestra.csb.pitt.edu/	Predicts small molecule binding affinity to proteins using a protein-ligand complex deep learning model.	[1]
DrugQuest	https://bmcbioinformatics.biomedcentral.com/articles/10.1186/s12859-016-1041-6	It is a text-mining tool for knowledge discovery in order to find new associations between drugs.	[59]
DeepCodex	https://deepcodex.org/	It detects correlation between small molecules functional similarity based on gene expression data.	[1]
HitPicK	https://doi.org/10.1093/bioinformatics/btt303	Prediction of hits in High-Throughput Screening (HTS) and their targets.	[60]
PROMISCUOUS	http://bioinformatics.charite.de/promiscuous/	PROMISCUOUS has gathered various drug repositioning drugs using text mining methods.	[9], [32]
PolySearch	https://bio.tools/polysearch	Finds three-dimensional chemical structures based on similarity to a query structure.	[61]
SuperPred	https://prediction.charite.de/index.php	Predicts the structure of proteins from their amino acid sequence.	[1]

Artificial intelligence and machine learning in drug repurposing

Artificial Intelligence (AI) and Machine Learning (ML) technologies play a pivotal role, particularly with the emergence of advanced deep learning and neural network algorithms. Machine learning algorithms can learn from data and replicate cognitive human behavior [62]. In contemporary times, biological researchers have incorporated machine learning into the processes of drug repurposing to enhance overall outcomes [63]. Machine learning helps drug design in two ways. Firstly, a statistical survey of the effectiveness of a group of existing drugs in a very focused way. Secondly, the design of repurposed drugs takes suitable feedback.

Support Vector Machine (SVM)

Support Vector Machine (SVM) is a machine learning algorithm that has found application in various domains, including drug repurposing. The data points closer to the hyperplane, influencing its position and orientation, are termed support vectors. Utilizing these vectors, the largest margin is established, dividing the dataset

into two groups and determining the category to which new data belongs. The SVM kernel transforms the input space into higher dimensions, showcasing the distinction between the new data and the support vectors. The key element of SVM lies in identifying the optimal hyperplane to effectively partition the data into distinct clusters [64]. The performance of SVM in drug repurposing depends on the quality and representativeness of the training data, the choice of features and the specific characteristics of the problem at hand.

Random Forest (RF)

Random Forest (RF) is a frequently employed algorithm specifically crafted for extensive datasets with numerous features. It streamlines tasks by eliminating outliers, classifying data and assigning datasets based on relative features tailored for the specific algorithm. Typically trained for substantial inputs and variables, it gains accessibility through data collection from various databases. RF proves advantageous in various scenarios, including handling missing data, addressing outliers and estimating characteristics for classification [65]. In the realms of drug discovery and drug repurposing, RFs find primary applications in feature selection, classification or regression.

K- Nearest Neighbor (K-NN)

The K-NN algorithm involves storing all data points during the training phase and the classification of new data points is determined by their similarity to existing data. As new instances are introduced to the established categories, the K-NN algorithm assesses these new cases against previous instances and assigns them to the same category as the new data [64]. In the K-NN algorithm, the Minkowski Parameter (p) is a configurable parameter that influences the choice of distance metric. Choosing different values of p can impact the performance of the algorithm based on the characteristics of the dataset. Common choices include p=1 for Manhattan distance and p=2 for Euclidean distance. K-Nearest Neighbor offers a straightforward yet powerful approach in drug repurposing by leveraging the concept of similarity. Its capacity to identify drugs with analogous characteristics provides a valuable tool for predicting potential novel therapeutic uses for existing medications.

Logistic Regression (LR)

Logistic Regression finds application in drug repurposing as a predictive modeling tool. It is primarily employed for categorizing observations into discrete classes. This technique adjusts its output utilizing the logistic sigmoid function, converting probability values into at least two distinct classes [64].

Ensemble model

An ensemble strategy merges two or more base classifiers, elevating the outcomes of machine learning. It delivers superior predictive performance compared to a single model [64].

Homogeneous model

A homogeneous model integrates similar types of base classifiers, resulting in improved performance. Random Forest and AdaBoost are two distinct homogeneous models utilized for training the dataset. The Random forest method employs numerous decision trees as base classifiers in a parallel manner whereas the AdaBoost classifier amalgamates multiple weak classifiers to generate a robust classifier with a high level of accuracy [64].

Heterogeneous model

This diverse model integrates various base classifiers to achieve enhanced accuracy. The heterogeneous ensemble comprises five distinct types of base classifiers: K-NN, LR, SVM, RF and AdaBoost. Hard voting and soft voting techniques are applied to aggregate the results. The voting classifier receives training data from multiple machine learning models and predicts an output based on the class with the highest probability, combining the results from each base estimator [64].
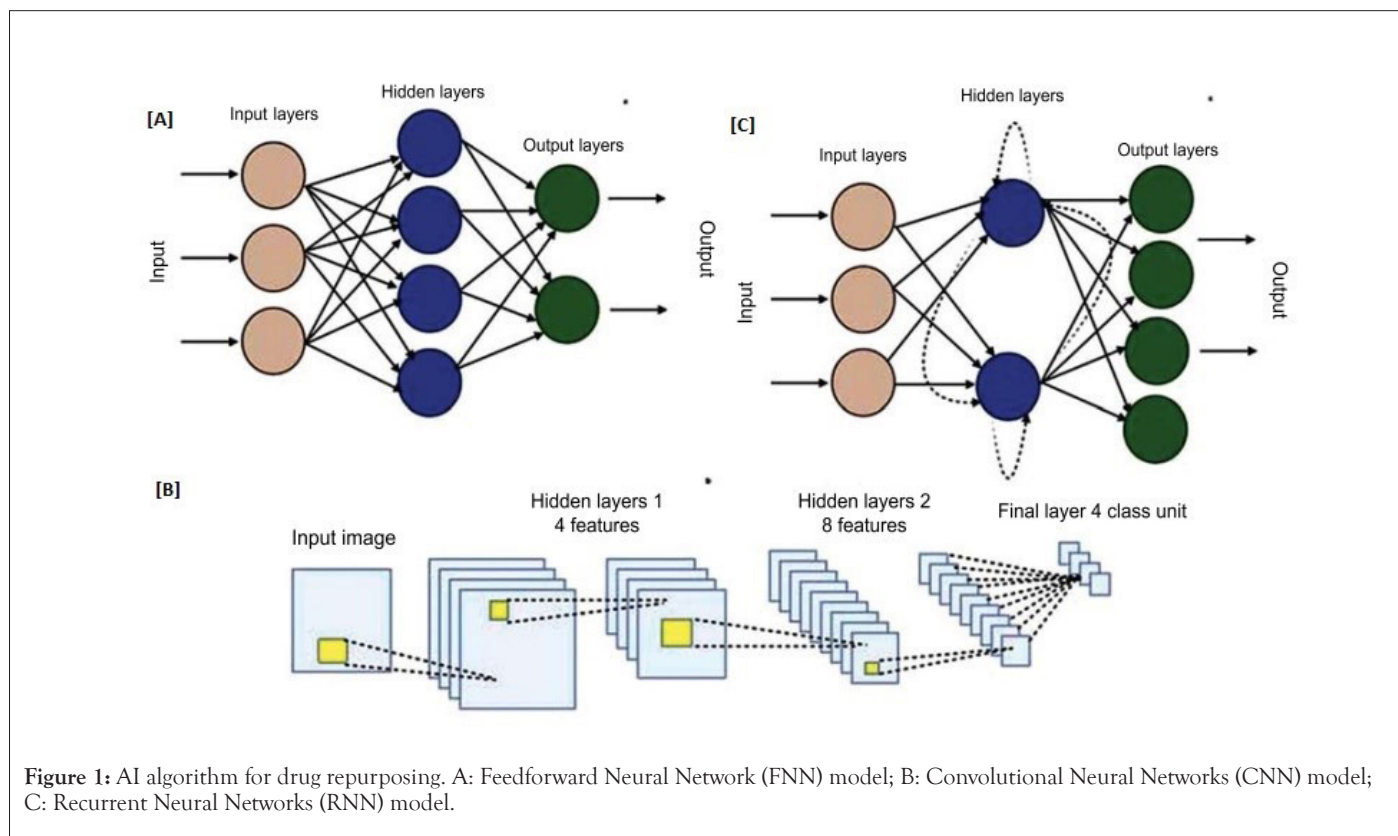
Deep learning in drug repurposing

Deep learning, a subfield of machine learning, has revolutionized various fields, including drug discovery and repurposing. Its ability to extract meaningful patterns from large datasets makes it well-suited for identifying novel therapeutic applications for existing drugs, potentially reducing the time and costs associated with traditional drug discovery processes.

DISCUSSION

One of the primary applications of deep learning in drug repurposing lies in predicting Drug-Target Interactions (DTIs). Deep learning models can analyze vast biological data, including protein sequences, gene expression profiles and drug structures, to uncover potential interactions between known drugs and novel targets. This capability enables researchers to explore new therapeutic opportunities for existing drugs that may have been overlooked using traditional methods [66]. Three different algorithms widely used in deep learning namely, Feedforward Neural Network (FNN), Convolutional Neural Network (CNN) and Recurrent Neural Network (RNN) (Figure 1). The FNN algorithm is applied for the vector representation of samples in the dataset. It is connected through artificial neurons layer-by-layer from input variables to output targets [67-69]. Techniques such as CNNs are employed to analyze biological images, such as those from high-throughput screenings or medical imaging. This facilitates the identification of patterns indicative of drug efficacy or safety, contributing to the discovery of new therapeutic applications [70,71]. As per the RNN, its primary application lies in investigating features for drug repurposing within biological sequences. These models play a key role in producing extensive libraries of molecules for drug development, where each molecule is represented as a sequence using straightforward computational codes for molecular input line entry [71-74].

Deep learning also plays a crucial role in identifying drug combinations, where multiple drugs are used together to achieve enhanced therapeutic effects. By analyzing patterns in drug response data and patient outcomes, deep-learning models can predict synergistic drug interactions, leading to the development of more effective treatment regimens [75].

Figure 1: AI algorithm for drug repurposing. A: Feedforward Neural Network (FNN) model; B: Convolutional Neural Networks (CNN) model; C: Recurrent Neural Networks (RNN) model.

CONCLUSION

In conclusion, drug repurposing has emerged as a promising strategy to accelerate the discovery and development of new therapeutic agents. The availability of comprehensive databases and powerful artificial intelligence techniques has significantly enhanced the drug repurposing process by enabling researchers to identify new uses for existing drugs, predict drug-target interactions and optimize drug combinations. Databases such as DrugBank, BindingDB, PubChem, etc., provide valuable information on drug properties, targets and biological activities, facilitating the exploration of potential repurposing opportunities. Additionally, tools like DrugMap, DrugQuest, etc., offer computational methods for analyzing and interpreting drug-related data, further aiding in the identification of promising repurposing candidates. Artificial intelligence techniques, particularly deep learning, have revolutionized drug repurposing by enabling the extraction of complex patterns from large datasets. Deep learning models can predict drug-target interactions, identify synergistic drug combinations and optimize repurposing strategies, significantly accelerating the drug repurposing process. The integration of comprehensive databases and powerful artificial intelligence techniques has transformed drug repurposing into a data-driven and efficient approach for expanding the therapeutic landscape. By leveraging these tools, researchers can accelerate the discovery and development of new treatments, addressing unmet medical needs and improving patient outcomes. As drug repurposing continues to gain traction, the role of databases and artificial intelligence will only grow in importance, paving the way for a more efficient and effective drug discovery process.

REFERENCES

1. Selvaraj G, Kaliamurthi S, Peslherbe GH, Wei DQ. Application of artificial intelligence in drug repurposing: A mini-review. Curr Chin Sci. 2021;1(3):333-345.

2. Álvarez-Machancoses Ó, Fernández-Martínez JL. Using artificial intelligence methods to speed up drug discovery. Expert Opin Drug Discov. 2019;14(8):769-777.

3. Chinnasamy S, Selvaraj G, Kaushik AC, Kaliamurthi S, Chandrabose S, Singh SK, et al. Molecular docking and molecular dynamics simulation studies to identify potent AURKA inhibitors: assessing the performance of density functional theory, MM-GBSA and mass action kinetics calculations. J Biomol Struct Dyn. 2020;38(14):4325-4335.

4. Chinnasamy S, Selvaraj G, Selvaraj C, Kaushik AC, Kaliamurthi S, Khan A, et al. Combining *insilico* and *invitro* approaches to identification of potent inhibitor against phospholipase A2 (PLA2). Int J Biol Macromol. 2020;144:53-66.

5. Xue H, Li J, Xie H, Wang Y. Review of drug repositioning approaches and resources. Int J Biol Sci. 2018;14(10):1232.

6. Sabale AS, Patel UM, Gorhe AA, Bodhankar VR, Shinde MN, Sonawane GA, et al. Prospective Anti-cancer Agents: A Recent Update on Drug Repurposing. Anticancer. Agents Med. Chem. 2020;20.

7. Martorana A, Perricone U, Lauria A. The repurposing of old drugs or unsuccessful lead compounds by *insilico* approaches: new advances and perspectives. Curr Top Med Chem. 2016;16(19):2088-2106.

8. Setoain J, Franch M, Martínez M, Tabas-Madrid D, Sorzano CO, Bakker A, et al. NFFinder: an online bioinformatics tool for searching similar transcriptomics experiments in the context of drug repositioning. Nucleic Acids Res. 2015;43(W1):W193-199.

9. Masoudi-Sobhanzadeh Y, Omidi Y, Amanlou M, Masoudi-Nejad A. Drug databases and their contributions to drug repurposing. Genomics. 2020;112(2):1087-1095.

10. Yang K, Dinasarapu AR, Reis ES, DeAngelis RA, Ricklin D, Subramaniam S, et al. CMAP: complement map database. Bioinformatics. 2013;29(14):1832-1833.

11. Corsello SM, Bittker JA, Liu Z, Gould J, McCarren P, Hirschman JE, et al. The Drug Repurposing Hub: a next-generation drug library and information resource. Nat Med. 2017;23(4):405-408.

12. Lewis JD, Schinnar R, Bilker WB, Wang X, Strom BL. Validation studies of the health improvement network (THIN) database for pharmacoepidemiology research. Pharmacoepidemiol Drug Saf. 2007;16(4):393-401.

13. Yoo M, Shin J, Kim J, Ryall KA, Lee K, Lee S, et al. DSigDB: drug signatures database for gene set analysis. Bioinformatics. 2015;31(18):3069-3071.

14. Roider HG, Pavlova N, Kirov I, Slavov S, Slavov T, Uzunov Z, et al. Drug2Gene: an exhaustive resource to explore effectively the drug-target relation network. BMC Bioinformatics. 2014;15:1-11.

15. Wishart DS, Knox C, Guo AC, Shrivastava S, Hassanali M, Stothard P, et al. DrugBank: a comprehensive resource for insilico drug discovery and exploration. Nucleic Acids Res. 2006;34:D668-672.

16. Zeng H, Qiu C, Cui Q. Drug-path: a database for drug-induced pathways. Database. 2015;bav061.

17. Araki H, Knapp C, Tsai P, Print C. GeneSetDB: a comprehensive meta-database, statistical and visualisation framework for gene set analysis. FEBS Open Bio. 2012;2:76-82.

18. Gao Z, Li H, Zhang H, Liu X, Kang L, Luo X, et al. PDTD: a web-accessible protein database for drug target identification. BMC Bioinformatics. 2008;9:1-7.

19. Kuhn M, Letunic I, Jensen LJ, Bork P. The SIDER database of drugs and side effects. Nucleic Acids Res. 2016;44(D1):D1075-1079.

20. Zhu F, Shi Z, Qin C, Tao L, Liu X, Xu F, et al. Therapeutic target database update 2012: a resource for facilitating target-oriented drug discovery. Nucleic Acids Res. 2012;40(D1):D1128-1136.

21. Sandgren A, Strong M, Muthukrishnan P, Weiner BK, Church GM, Murray MB. Tuberculosis drug resistance mutation database. PLoS Med. 2009;6(2):e1000002.

22. Amelio I, Gostev M, Knight RA, Willis AE, Melino G, Antonov AV. DRUGSURV: a resource for repositioning of approved and experimental drugs in oncology based on patient survival information. Cell Death Dis. 2014;5(2):e1051.

23. Shafer RW. Rationale and uses of a public HIV drug-resistance database. J Infect Dis. 2006;194:S51-58.

24. Levy AR, O'Brien BJ, Sellors C, Grootendorst P, Willison D. Coding accuracy of administrative drug claims in the Ontario Drug Benefit database. Can J Clin Pharmacol. 2003;10(2):67-71.

25. Preissner S, Kroll K, Dunkel M, Senger C, Goldsobel G, Kuzman D, et al. SuperCYP: a comprehensive database on Cytochrome P450 enzymes including a tool for analysis of CYP-drug interactions. Nucleic Acids Res. 2010;38:D237-243.

26. Chen CY. TCM Database@ Taiwan: the world's largest traditional Chinese medicine database for drug screening insilico. PloS One. 2011;6(1):e15939.

27. Ru J, Li P, Wang J, Zhou W, Li B, Huang C, et al. TCMSP: a database of systems pharmacology for drug discovery from herbal medicines. J Cheminform. 2014;6:1-6.

28. Agüero F, Al-Lazikani B, Aslett M, Berriman M, Buckner FS, Campbell RK, et al. Genomic-scale prioritization of drug targets: the TDR Targets database. Nat Rev Drug Discov. 2008;7(11):900-907.

29. Tao W, Li B, Gao S, Bai Y, Shar PA, Zhang W, et al. CancerHSP: anticancer herbs database of systems pharmacology. Sci Rep. 2015;5(1):11481.

30. Gaulton A, Bellis LJ, Bento AP, Chambers J, Davies M, Hersey A, et al. ChEMBL: a large-scale bioactivity database for drug discovery. Nucleic Acids Res. 2012;40(D1):D1100-1107.

31. Huang H, Nguyen T, Ibrahim S, Shantharam S, Yue Z, Chen JY. DMAP: a connectivity map database to enable identification of novel drug repositioning candidates. BMC Bioinformatics 2015;16(13):1-11.

32. Von Eichborn J, Murgueitio MS, Dunkel M, Koerner S, Bourne PE, Preissner R. PROMISCUOUS: a database for network-based drug-repositioning. Nucleic Acids Res. 2010;39:D1060-1066.

33. Wirth M, Zoete V, Michielin O, Sauer WH. SwissBioisostere: a database of molecular replacements for ligand design. Nucleic Acids Res. 2013;41(D1):D1137-1143.

34. Luo H, Chen J, Shi L, Mikailov M, Zhu H, Wang K, et al. DRAR-CPI: a server for identifying drug repositioning potential and adverse drug reactions via the chemical–protein interactome. Nucleic Acids Res. 2011;39:W492-498.

35. Fu C, Jin G, Gao J, Zhu R, Ballesteros-Villagrana E, Wong ST. DrugMap Central: an on-line query and visualization tool to facilitate drug repositioning studies. Bioinformatics. 2013;29(14):1834-1836.

36. Martinez V, Navarro C, Cano C, Fajardo W, Blanco A. DrugNet: network-based drug–disease prioritization by integrating heterogeneous data. Artif Intell Med. 2015;63(1):41-49.

37. Sun J, Wu Y, Xu H, Zhao Z. DTome: a web-based tool for drug-target interactome construction. BMC Bioinformatics 2012;13(9):1-10.

38. Alaimo S, Bonnici V, Cancemi D, Ferro A, Giugno R, Pulvirenti A. DT-Web: a web-based application for drug-target interaction and drug combination prediction through domain-tuned network-based inference. BMC Syst Biol. 2015;9(3):1-11.

39. Pihan E, Colliandre L, Guichou JF, Douguet D. e-Drug3D: 3D structure collections dedicated to drug repurposing and fragment-based drug design. Bioinformatics. 2012;28(11):1540-1541.

40. Brown AS, Kong SW, Kohane IS, Patel CJ. ksRepo: a generalized platform for computational drug repositioning. BMC bioinformatics. 2016;17(1):1-5.

41. Kim S, Thiessen PA, Bolton EE, Chen J, Fu G, Gindulyte A, et al. PubChem substance and compound databases. Nucleic Acids Res. 2016;44(D1):D1202-1213.

42. Qureshi R, Irfan M, Gondal TM, Khan S, Wu J, Hadi MU, et al. AI in Drug Discovery and its Clinical Relevance. Heliyon. 2023; 9(7):e17575.

43. Irwin JJ, Shoichet BK. ZINC a free database of commercially available compounds for virtual screening. J Chem Inf Model. 2005;45(1):177-182.

44. Gilson MK, Liu T, Baitaluk M, Nicola G, Hwang L, Chong J. BindingDB in 2015: a public database for medicinal chemistry, computational chemistry and systems pharmacology. Nucleic Acids Res. 2016;44(D1):D1045-1053.

45. Di L, Kerns EH. Drug-like properties: concepts, structure design and methods from ADME to toxicity optimization. 2015.

46. Szklarczyk D, Santos A, Von Mering C, Jensen LJ, Bork P, Kuhn M. STITCH 5: augmenting protein–chemical interaction networks with tissue and affinity data. Nucleic Acids Res. 2016;44(D1):D380-384.

47. Yang W, Soares J, Greninger P, Edelman EJ, Lightfoot H, Forbes S, et al. Genomics of Drug Sensitivity in Cancer (GDSC): a resource for therapeutic biomarker discovery in cancer cells. Nucleic Acids Res. 2012;41(D1):D955-961.

48. Wang R, Fang X, Lu Y, Yang CY, Wang S. The PDBbind database: methodologies and updates. J Med Chem. 2005;48(12):4111-4119.

49. Coker EA, Mitsopoulos C, Tym JE, Komianou A, Kannas C, Di Micco P, et al. canSAR: update to the cancer translational research and drug discovery knowledgebase. Nucleic Acids Res. 2019;47(D1):D917-922.

50. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, et al. The protein data bank. Nucleic Acids Res. 2000;28(1):235-242.

51. UniProt Consortium. UniProt: a hub for protein information. Nucleic Acids Res. 2015;43(D1):D204-212.

52. Wu Z, Ramsundar B, Feinberg EN, Gomes J, Geniesse C, Pappu AS, et al. MoleculeNet: a benchmark for molecular machine learning. Chem Sci. 2018;9(2):513-530.

53. Harvey MJ, Giupponi G, Fabritiis GD. ACEMD: accelerating biomolecular dynamics in the microsecond time scale. J Chem Theory Comput. 2009;5(6):1632-1639.

54. Trott O, Olson AJ. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. J Comput Chem. 2010;31(2):455-461.

55. Wang H, Zhang L, Han J, Weinan E. DeePMD-kit: A deep learning package for many-body potential energy representation and molecular dynamics. Comput Phys Commun. 2018;228:178-184.

56. Skjærven L, Yao XQ, Scarabelli G, Grant BJ. Integrating protein structural dynamics and evolutionary analysis with Bio3D. BMC Bioinformatics. 2014;15(1):1-11.

57. DeLano WL. Pymol: An open-source molecular graphics tool. CCP4 Newsl. Protein Crystallogr. 2002;40(1):82-92.

58. Lyskov S, Chou FC, Conchuir SO, Der BS, Drew K, Kuroda D, et al. Serverification of molecular modeling applications: the Rosetta Online Server that Includes Everyone (ROSIE). PloS One. 2013;8(5):e63906.

59. Papanikolaou N, Pavlopoulos GA, Theodosiou T, Vizirianakis IS, Iliopoulos I. DrugQuest-a text mining workflow for drug association discovery. BMC Bioinformatics. 2016;17:333-341.

60. Liu X, Vogt I, Haque T, Campillos M. HitPick: a web server for hit identification and target prediction of chemical screenings. Bioinformatics. 2013;29(15):1910-1912.

61. Cheng D, Knox C, Young N, Stothard P, Damaraju S, Wishart DS. PolySearch: a web-based text mining system for extracting relationships between human diseases, genes, mutations, drugs and metabolites. Nucleic Acids Res. 2008;36:W399-405.

62. Peyvandipour A, Saberian N, Shafi A, Donato M, Draghici S. A novel computational approach for drug repurposing using systems biology. Bioinformatics. 2018 Aug 15;34(16):2817-2825.

63. Ahmed F, Soomro AM, Salih AR, Samantasinghar A, Asif A, Kang IS, et al. A comprehensive review of artificial intelligence and network based approaches to drug repurposing in Covid-19. Biomed Pharmacother. 2022;153:113350.

64. Simon E, Bankapur S. A machine learning based study for the prediction of drug-target interaction using protein and drug molecule descriptors. AIP Conference Proceedings. 2023.

65. Patel L, Shukla T, Huang X, Ussery DW, Wang S. Machine learning methods in drug discovery. Molecules. 2020;25(22):5277.

66. Rustandi T, Prihandiwati E, Nugroho F, Hayati F, Afriani N, Alfian R, et al. Application of Artificial Intelligence (AI) in the development of Jamu"traditional Indonesian medicine"as a more effective drug. Front Artif Intel. 2023;6:1274975.

67. Rumelhart DE, Hinton GE, Williams RJ. Learning representations by back-propagating errors. Nature. 1986;323(6088):533-536.

68. Aliper A, Plis S, Artemov A, Ulloa A, Mamoshina P, Zhavoronkov A. Deep learning applications for predicting pharmacological properties of drugs and drug repurposing using transcriptomic data. Mol Pharm. 2016;13(7):2524-2530.

69. Wood T. Convolutional Neural Network Definition-DeepAI.

70. Meyer JG, Liu S, Miller IJ, Coon JJ, Gitter A. Learning drug functions from chemical structures with convolutional neural networks and random forests. J Chem Inf Model. 2019;59(10):4438-4449.

71. Wang X, Liu Y, Lu F, Li H, Gao P, Wei D. Dipeptide frequency of word frequency and graph convolutional networks for DTA prediction. Front Bioeng Biotechnol. 2020;8:267.

72. Segler MH, Kogej T, Tyrchan C, Waller MP. Generating focused molecule libraries for drug discovery with recurrent neural networks. ACS Cent Sci. 2018;4(1):120-131.

73. Wang YB, You ZH, Yang S, Yi HC, Chen ZH, Zheng K. A deep learning-based method for drug-target interaction prediction based on long short-term memory neural network. BMC Med Inform Decis Mak. 2020;20(2):1-9.

74. Chu Y, Kaushik AC, Wang X, Wang W, Zhang Y, Shan X, et al. DTI-CDF: a cascade deep forest model towards the prediction of drug-target interactions based on hybrid features. Brief Bioinform. 2021;22(1):451-462.

75. Liu R, Wei L, Zhang P. A deep learning framework for drug repurposing *via* emulating clinical trials on real-world patient data. Nat Mach Intell. 2021;3(1):68-75.