

Review

Not peer-reviewed version

Computational Advancements in Cancer Combination Therapy Prediction

Victoria L. Flanary , [Jennifer L. Fisher](#) , Elizabeth J. Wilk , Timothy C. Howton , [Brittany N. Lasseigne](#) *

Posted Date: 20 July 2023

doi: 10.20944/preprints202305.1637v2

Keywords: computational; drug repurposing; drug repositioning; cancer; combination therapy; network biology; machine learning; deep learning; precision oncology



Preprints.org is a free multidiscipline platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Review

Computational Advancements in Cancer Combination Therapy Prediction

Victoria L. Flanary¹, Jennifer L. Fisher¹, Elizabeth J. Wilk¹, Timothy C. Howton¹
and Brittany N. Lasseigne^{1,*}

¹ Department of Cell, Developmental and Integrative Biology, Heersink School of Medicine, The University of Alabama at Birmingham, McCallum Bldg, 1900 University Blvd, Birmingham, AL 35233, USA

* Correspondence: bnp0001@uab.edu; Tel.: (205)-934-8156

Abstract: As cancer remains resistant to several modes of treatment, novel therapeutics are still under active investigation to overcome treatment inefficacy in cancer. Given the high attrition rate of *de novo* drug discovery, drug screening, and drug repurposing have offered time- and cost-effective alternative strategies for the identification of potentially effective therapeutics. In contrast to large-scale drug screens, computational approaches for drug repurposing leverage the increasing amounts of biomedical data to predict candidate therapeutic agents prior to testing in biological models. Current studies in drug repurposing for cancer therapy prediction have increasingly focused on the prediction of combination therapies, as combination therapies have numerous advantages over monotherapies. These include increased effect from synergistic interactions, reduced toxicity from lowered drug doses, and a reduced risk of resistance due to multiple non-overlapping mechanisms of action. This review provides a summary of several classes of computational methods used for drug combination therapy prediction in cancer research, including networks, regression-based machine learning, classifier machine learning models, and deep learning approaches, with the goal of presenting current progress in the field, particularly to non-computational cancer biologists. We conclude by discussing the need for further advancements in technologies that incorporate disease mechanisms, drug characteristics, multi-omics data, and clinical considerations to generate effective patient-specific drug combinations, as holistic data integration will inevitably result in optimal targeted therapeutics for cancer.

Keywords: computational; drug repurposing; drug repositioning; cancer; combination therapy; network biology; machine learning; deep learning; precision oncology

Introduction

Although novel therapeutic approaches have revolutionized the treatment of cancer, cancer remains highly resistant to current therapies, with drug resistance and resultant treatment inefficacy responsible for upward to 90% of cancer-related deaths (Borst, 2012; Holohan *et al.*, 2013; Housman *et al.*, 2014; Alfarouk *et al.*, 2015; Rueff and Rodrigues, 2016). Given the high attrition rate of *de novo* drug discovery, with approximately 90% of drugs failing to pass phase I clinical trials despite average investments of \$3 billion across 13–15 years per compound, many researchers have turned to phenotypic drug screening or drug repositioning as a time- and cost-effective alternatives (Booth and Zimmel, 2004; Scannell *et al.*, 2012; Sun *et al.*, 2022). While *in vitro* drug screening assays have the advantage of providing direct knowledge of drug effects across various dosages within a biological disease model, computational drug repurposing methods can leverage the ever-growing amount of available biomedical data to identify novel drug candidates (Wilkinson and Pritchard, 2015; Jarada, Rokne and Alhajj, 2020).

Most drug repurposing studies have focused on the identification of single agents to treat a target disease. However, for complex and heterogeneous diseases like cancer, it is unlikely that a single agent can target every cancer cell or relevant disease pathway, thus resulting in residual treatment-resistant cells that can lead to tumor recurrence. Therefore, recent advances in drug repositioning efforts for cancer have prioritized the prediction of multi-targeted combination therapies to both enhance treatment efficacy and reduce the risk of monotherapy resistance (Al-

Lazikani, Banerji and Workman, 2012; Ayoub, 2021). Combination therapies also reduce the toxicity of treatment regimens by maintaining therapeutic efficacy at lower doses of individual agents within the drug regimen (Ayoub, 2021). However, the majority of combination therapy prediction methods have prioritized identifying maximally effective therapeutic treatments while neglecting the impact of potential toxicity and tolerability of the drug combination to the patient. Toxicity thus remains an important clinical consideration by which these methods can be improved upon (Kong *et al.*, 2022).

Computational drug combination prediction approaches have generally been designed to achieve one of two goals: to target parallel disease pathways or to maximize computational scores of therapeutic efficacy (He *et al.*, 2023). As opposed to the multi-pathway approach, which aims to identify non-overlapping targets to maximize the efficacy of candidate drug treatments by targeting parallel disease or treatment-resistance pathways, computational score-based approaches aim to maximize either the synergy or sensitivity scores of a drug combination. Sensitivity refers to the degree of treatment response measured by the percent inhibition of cell viability or growth in *in vitro* experiments (Güvenç Paltun, Kaski and Mamitsuka, 2021). Synergy is a type of drug-drug interaction in which the effect of a drug combination is greater than the additive effect of individual drugs in the combination (Güvenç Paltun, Kaski and Mamitsuka, 2021). This is the most common goal for drug combination prediction approaches, as achieving synergistic interactions may maximize the efficacy of drug treatments. Synergy is measured by several metrics, including Loewe additivity, Bliss independence, highest single agent, and the Chou-Talalay method, all of which have been concisely reviewed by Pemovska *et al.* (Pemovska, Bigenzahn and Superti-Furga, 2018).

This review will discuss computational drug combination prediction methods that aim to identify effective multi-drug regimens for cancer across both of these goals (Figure 1 and Table 1). While combination therapy prediction methods have been previously reviewed (Güvenç Paltun, Kaski and Mamitsuka, 2021; Kong *et al.*, 2022; He *et al.*, 2023), we categorize prediction methods by the machine learning model used in their design and provide explanations for each method class to introduce these computational modeling approaches. Similar reviews that primarily summarize monotherapy prediction approaches for cancer have been provided by Rafique *et al.* and Firoozbakht *et al.* (Rafique, Islam and Kazi, 2021; Firoozbakht, Yousefi and Schwikowski, 2022). This review expands upon the existing literature by providing an introduction to computational modeling in the context of the most recently developed and impactful findings for cancer combination therapy prediction. First, we describe various networks-based methods that examine the relationships between biological entities, such as genes, proteins, pathways, and phenotypes, to design combinatorial therapies that target optimal disease- and drug-related elements. Then, we review regression-based machine learning models that predict missing dose response values (i.e. sensitivity) for synergy calculations in drug pairs and multi-drug cocktails. This is followed by a review of classifier-based machine learning models capable of predicting new drug combinations through both drug targets and synergy calculations. The last category of computational approaches described in this review is deep learning methods, which build upon the previously described machine learning models to tackle larger and more complex data. We conclude this review with suggestions for future directions by which computational drug combination prediction methods may be improved to enhance their utility and translation to the clinic.

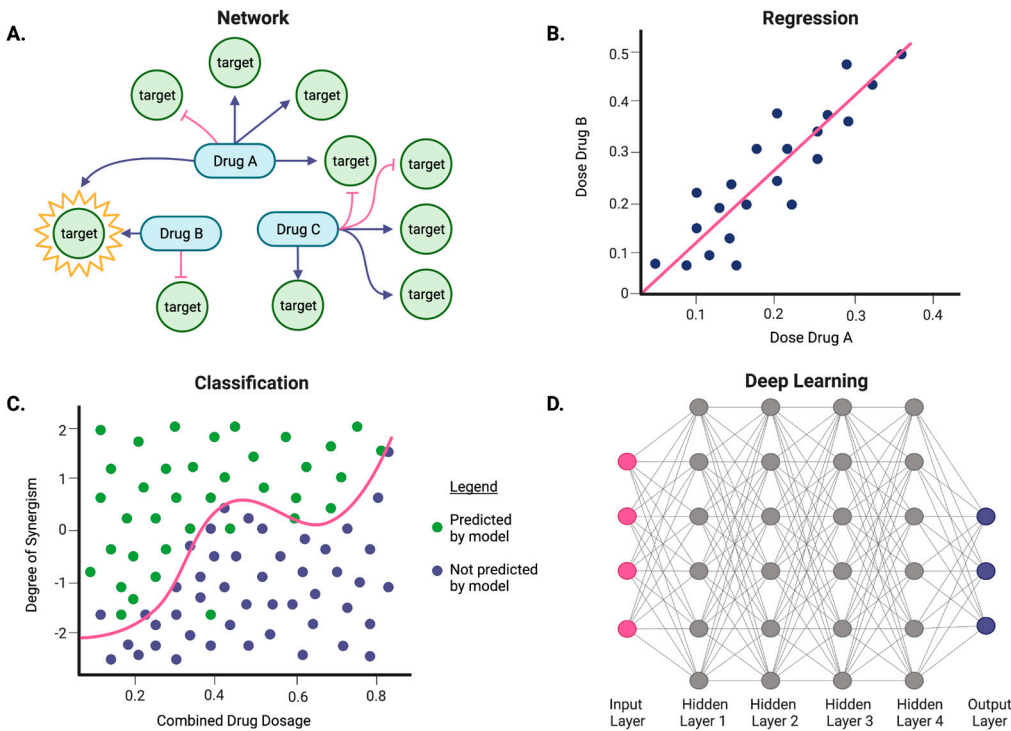


Figure 1. Visual representation of method classes used for cancer combination therapy prediction. A. Example of a drug-target network, in which drugs and targets are represented as nodes and the relations between them depicted as edges. Drugs A and B are shown to be synergistic through their activation of the same target. B. Depiction of a linear regression model, where data points are fitted to a linear relationship. C. General representation of classification models, which sort data points into categories, as opposed to regression models that fit data to a specific relationship (1B). D. Schema for an artificial neural network, the basic unit of deep learning approaches.

Table 1. Summary of machine learning method classes for combination therapy prediction.

Method Class	Definition	Notes
Network	Graphical representation of biological entities, such as genes, proteins, transcription factors, phenotypes, and drugs and how they relate to one another	<ul style="list-style-type: none">• Provide a visual representation of the relationships between biological entities, which may reveal disease physiology or drug mechanism of action.• Allow for integration of several data types.• Interpretability may be difficult to those inexperienced in network biology.
Regression	Subclass of machine learning that determines whether the relationship between two variables fits a known mathematical pattern (i.e., linear, logarithmic, polynomial, etc.)	<ul style="list-style-type: none">• Often used to predict unknown drug responses or to predict responses to novel drug combinations• Higher interpretability compared to other machine learning models due to comparatively lower model complexity• Capable of fitting data to multiple types of mathematical patterns• Requires knowing which specific type of mathematical pattern exists between two variables to make accurate predictions

Classification	Subclass of machine learning that places observations in the data into specific categories	<ul style="list-style-type: none">• Highly versatile, as several methods fall into this subclass (logistic regression, support vector machines, random forest, gradient boosting, XGBoost)• Trade-off often exists for model accuracy and interpretability
Deep Learning	Subclass of machine learning that uses multi-layered artificial neural networks composed of several processing layers to make predictions from large and complex data types	<ul style="list-style-type: none">• Can handle large, multi-faceted data types that may overwhelm other methods• Can discern significant biological relationships often overlooked by other methods• High accuracy compared to other machine learning methods• Tendency to overfit, reducing the generalizability of these models• Lower interpretability (“black boxes”)• Current lack of adequate data for modeling, as most approaches are trained on a limited number of cell lines

Networks-based Models

Given the degree of intercellular heterogeneity in cancer, it is helpful to understand how various molecular entities within cancer cells and their microenvironment interact with one another. Networks are graphical models that show associations between molecular entities within complex systems. These networks consist of nodes, which are units that symbolize molecules like proteins or larger concepts like phenotypes, and edges, which are lines that depict relationships between the nodes they connect. There are different categories of networks that represent various types of relationships between different entities. For example, drug-target networks are prominent network models for combination therapy prediction that represent both drugs and their protein targets as nodes and the relationships between drugs and their corresponding targets as edges (Figure 1A). Edges can also be weighted to quantify relationships between nodes in a network. Several network analysis algorithms, such as information propagation procedures, have been developed to take advantage of these weighted edges and provide further information on systems of interest. The types of networks that are most commonly used to identify novel drug repurposing candidates include drug-disease, drug-target, drug-drug, and protein-protein interaction networks (Azuaje, 2013). These networks can be integrated into heterogeneous, multi-omic graphs to reveal how therapeutic agents interact with biological systems through their direct targets. Interaction data for these multi-omic graphs come from publicly available databases, such as STRING (Szklarczyk *et al.*, 2023) for protein-protein interaction data, CollecTRI for transcription factor-gene regulatory relationships (Müller-Dott *et al.*, 2023), as well as the Therapeutic Target Database (Zhou *et al.*, 2022), Open Targets (Koscielny *et al.*, 2017), and Drug Repurposing Hub (Corsello *et al.*, 2017) for drug and drug-target data. Additionally, these systems-level interactions can predict therapeutic mechanisms of action, adverse events, and alternative applications of approved drugs and potentially synergistic drug combinations. We summarize major findings and examples of network-based drug combination therapy prediction methods in Table 2.

Table 2. Networks-based combination therapy prediction methods and findings.

Model	Algorithm or Finding	Impact
(Menche et al., 2015)	Disease modules - subgraphs within a biological network enriched in genes associated with disease etiology and progression	Novel drug combinations can be derived by combining drugs whose targets lie within the same disease module.
(Cheng, Kovács and Barabási, 2019)	Complementary exposure - condition in which two drugs' targets overlap with the disease module, but not each other	Of the six drug-drug-disease relationships tested, only complementary exposure correlates with therapeutic effect.
(Federico et al., 2022)	Combined protein-protein interactions, gene co-expression, and gene regulation into a multi-omic network to prioritize novel candidates across five cancer types. Introduced the druggability map.	The druggability map enables prioritization of drug repurposing candidates through consideration of both drug and disease characteristics.
(Jafari et al., 2022)	Generated bipartite networks for both patient and drug similarity. Two distinct clusters of small molecules were found in the drug similarity network.	Drug pairs derived by combining a drug from each cluster in the drug similarity network were synergistic across multiple synergy metrics.
(Sarmah et al., 2023)	Developed a temporal cell state network of cancer cells across stages of the cell cycle.	Knowledge of cell state transition dynamics may inform response to combination therapies.

Disease modules, which are subnetworks within a biological graph enriched in genes that are associated with disease etiology and progression, may be utilized in network-based drug combination repurposing (Menche *et al.*, 2015). Some network-based methods for drug combination prediction use proximity (i.e., nearest distance between nodes of interest) of drug targets to disease modules within a network to predict candidate therapies. This is based on the premise that drug combinations with targets contained within the same disease module will have increased efficacy. Cheng et al. created a separation metric to determine the distance between the drug-drug-disease modules (Cheng, Kovács and Barabási, 2019). After comparing six classes of drug-drug-disease relationship combinations, they found that the only class that correlated with therapeutic effect was complementary exposure, or the situation in which two drugs' targets overlap a disease module within the interactome, but not with each other. In a subsequent work, Federico et al. generated integrated disease networks by combining protein-protein interactions, gene co-expression, and gene regulation (Federico *et al.*, 2022). They then prioritized drug combinations for five cancers (breast cancer, hepatocellular carcinoma, prostate adenocarcinoma, stomach adenocarcinoma, colon adenocarcinoma, and lung adenocarcinoma) by accounting for the drugs' mechanisms of action, secondary structures, and the drug targets' shortest paths within the network. Drugs whose targets were directly connected were deprioritized due to their overlapping neighborhood area based on the work by Cheng et al (Cheng, Kovács and Barabási, 2019). This study also introduces the druggability map, a unique graphical instrument to prioritize drug repositioning candidates through the incorporation of both drug and disease characteristics.

Although the groups discussed above used network-based approaches to model general drug and disease entities to predict candidate combination therapies, another group leveraged patient-derived drug response to prioritize drug repurposing candidates. Jafari et al. used the Beat AML

dataset, a cohort of 672 acute myeloid leukemia samples screened for sensitivity to 122 drugs, to generate two bipartite networks: a patient similarity network and a drug similarity network (Jafari *et al.*, 2022). Analysis of clusters within the patient similarity network found characteristics and relationships that were used to account for patient heterogeneity in downstream analyses. The drug similarity network contained two distinct clusters of small molecules. They reasoned that designing combination therapies by combining the top candidates from each cluster of the drug similarity network into drug pairs may prevent drug resistance and cancer recurrence. Synergy analysis of these inter-cluster drug combinations in 135 drug-drug-cell line triplicates validated their model’s predicted regimens as highly synergistic across multiple synergy metrics, including Loewe additivity, Bliss independence, highest single agent, and zero interaction potency (Yadav *et al.*, 2015; Pemovska, Bigenzahn and Superti-Furga, 2018).

Given recent revelations on the nature of cancer cell plasticity from single-cell RNA sequencing studies, a recent publication from Sarmah *et al.* aimed to predict drug combination responses using a temporal cell state network model (Sarmah *et al.*, 2023). The authors explored the possibility that the types of cancer cells within a tumor (i.e. the different cell states across cancer cell populations), the speed by which cell state transitions occur, and how drugs affect those transitions may provide valuable information on drug combination response to therapy. They explore this hypothesis by testing three kinase inhibitors that each target a different cell cycle transition *in vitro*. They used a Markov model to gauge cell growth and single-agent drug sensitivities and then used this model to predict combinatorial drug responses. They calculated synergy using an excess over Bliss analysis, where drug synergy is defined by the observed drug response greater than that found by totaling individual drug sensitivities. Their results suggest that cell state transition dynamics and prior drug response knowledge may inform the response to drug combination therapies.

Regression-based Machine Learning Models

Regression-based machine learning models are often used in combination with prior knowledge of known drug sensitivities to predict unknown drug responses or to predict responses to drug combinations. These models predict outcomes based on whether a mathematical relationship exists between an independent and a dependent variable, with the most basic of these models fitting to a linear relationship (Figure 1B). Linear models have previously been used to reduce technical noise during the production of more robust models, to create full dose-response matrices (matrices that include all dose pairs for a drug combination pair over a desired concentration range) by predicting missing dosages, and to predict synergistic interactions (Amzallag, Ramaswamy and Benes, 2019; Ianevski *et al.*, 2019). Although these matrices are required for many synergy calculations, they are difficult to acquire, as manual drug testing becomes costly and impractical with numerous combinations and their replicates across various dosages (Amzallag, Ramaswamy and Benes, 2019; Ianevski *et al.*, 2019; Li, Xu and McIndoe, 2022). Other notes and example implementations of regression-based models for cancer combination therapy prediction are provided in Table 3.

Table 3. Regression-based combination therapy prediction methods.

Model	Algorithm	Impact
(Amzallag, Ramaswamy and Benes, 2019)	Applied a linear regression model based on Bliss independence to a dataset of drug response data points of all pairwise combinations of 108 drugs across 40 melanoma cell lines	Differentiates synergistic interactions from instances of potentiation

(Zimmer et al., 2017b)	Pairs model - integration of Bliss independence with logarithmic regression to predict the effect of higher-order drug combinations	Enables prediction of drug combinations consisting of upward to six drugs
(C. Li et al., 2022)	Keyboard - R programming package that combines three previously developed Bayesian regression-based drug prediction methods, allowing for incorporation of uncertainty and fitting to non-linear relationships	Allows for derivation of maximum tolerated drug doses, calculation of changes needed to reach optimal dosages, and determination of optimal doses for either single or multiple drug regimens

One linear model by Amzallag et al. aimed to reduce the noise produced in drug synergy prediction algorithms when the single agent data used for these calculations was captured incorrectly or incompletely (Amzallag, Ramaswamy and Benes, 2019). The authors generated a dataset of 439,000 drug response data points from testing all pairwise combinations of 108 drugs across 40 melanoma cell lines. They then applied a linear model based on the Bliss independence synergy metric (which assumes that the effect of a combination of drugs is equal to the product of the individual drugs) to all cell lines in their dataset. They found that both single agent sensitivity values and synergy values showed significantly high correlations from their linear model, and they were able to differentiate true synergistic interactions from instances of potentiation, where the addition of one drug enhances the effect of another while not directly contributing to the effect itself, using a specificity score.

Alternatively, Zimmer et al. integrated Bliss independence with a regression model to create the pairs model, which requires relatively few experiments to estimate the effect of multi-drug cocktails (Zimmer *et al.*, 2017a). They expanded upon the Bliss formula by employing drug response data of drug pairs to predict the effects of higher-order combinations that contain more than two drugs, as they had found that the interactions between pairs of drugs often predicted the overall effect of the regimens in which those pairs were included. Briefly, the formula for the pairs model smoothly converts between Bliss independence and logarithmic regression based on a parameter that defaults to only calculating by Bliss independence when equal to 0 and to only by the logarithmic-linear regression model when equal to 1. Any parameter value between 0 and 1 would interpolate between the Bliss and regression algorithms. This parameter is then adjusted based on the number of drugs in the desired drug combination (or cocktail), allowing for high-order drug combinations of up to 6 drugs, while only supplying drug pairs data as input.

Whereas the previously mentioned regression-based models assume that drug interactions fit a linear relationship by relying on the Bliss independence metric, Bayesian regression can be applied to optimize drug response predictions by assuming drug interactions have nonlinear relationships (Park, Nassar and Vikalo, 2013). Bayesian regression allows for the incorporation of uncertainty into models by estimating probability distributions over parameters, as opposed to using point estimates of parameters to make predictions like linear models. The R programming package Keyboard is a Bayesian regression-based approach developed to derive maximum tolerated doses, optimal dose increases and decreases, and optimal biological doses for single drug and drug combination experiments from clinical trial data (C. Li *et al.*, 2022). Keyboard combines three previously developed Bayesian-based drug prediction methods into its algorithms (Li *et al.*, 2017; Yan, Mandrekar and Yuan, 2017; Pan *et al.*, 2020). To predict candidate drug combinations, it considers the drug response profile of a patient cohort to a drug combination at two different dose combinations. It then predicts the maximum toxicity interval based on the updated data from the distribution of the second dose combination. This information allows the model to either increase or decrease doses with each new

cohort added to the calculations, which iteratively updates the maximum toxicity interval prediction based on the updated posterior distribution.

Classifier-based Machine Learning Models

Whereas machine learning regression-based models aim to predict drug combinations and their interactions by assuming these interactions fit a mathematical relationship, classifier-based approaches specify mathematical boundaries that classify observations into specific categories based on whether they fit into the categories’ specified ranges (e.g., classifying a drug interaction type as additive, synergistic, or antagonistic) (Figure 1C) (Güvenç Paltun, Kaski and Mamitsuka, 2021; James *et al.*, 2021). In the context of cancer drug combination prediction, these models have been applied to classify drug combination synergy *via* multiple modalities, including logistic regression, support vector machines, and decision trees. We summarize combination therapy prediction methods across these modalities in Table 4.

Table 4. Classification-based combination therapy prediction methods.

Model	Algorithm	Impact
(Iwata et al., 2015)	Logistic regression - probabilistic classifiers that determine the probability that an observation belongs to one of a finite number of specific categories	Predicted 142,988 drug combinations, including those for colon cancer and breast cancer
(Cüvitoğlu and Işık, 2017)	Support vector machine (SVM) - models based on kernel functions, which are used to transform data to a higher dimensionality	Implemented both single agent gene expression data and biological networks into an SVM to predict novel antineoplastics
(Celebi et al., 2019)	Decision trees: Random forest - ensemble approach that trains each model within the training data separately, then concludes the classification provided by the most models is the appropriate classification for an observation XGBoost - ensemble approach that analyzes each model sequentially, allowing for each subsequent model to improve upon the prediction of the previous model. XGBoost also applies regularization, which enhances the model’s generalizability.	Comparison of several machine learning methods for combination therapy prediction showed that overall, decision-tree based models outperform regression or support vector machine models in identifying synergistic combinations, with XGBoost outperforming random forest after parameter tuning.

Iwata et al. used a logistic regression model that incorporated target proteins and anatomical therapeutic chemical codes to predict potentially effective drug combinations for cancer (Iwata *et al.*,

2015). Logistic regression models are probabilistic classifiers that determine the probability that a new observation will fall into one of a finite number of categories (Güvenç Paltun, Kaski and Mamitsuka, 2021; James *et al.*, 2021). Iwata *et al.* used approved drug combinations from the FDA Orange Book and KEGG drug databases to train their model, which predicted 142,988 candidate drug combinations from known drug pairs, including some drug regimens for breast and colon cancer (Hare and Foster, 1990; Kanehisa *et al.*, 2008; Iwata *et al.*, 2015). While the limited complexity of logistic regression classifiers reduces the accuracy of these models, it also enhances their interpretability (Güvenç Paltun, Kaski and Mamitsuka, 2021; James *et al.*, 2021).

A more complex classifier model used for drug combination prediction is the support vector machine. Support vector machines (SVMs) are based on kernel functions, which include a variety of mathematical functions used to transform data from a lower to higher dimensionality (Rafique, Islam and Kazi, 2021). Cüvitoğlu and Işık used this classifier method to identify potentially effective antineoplastic drug pairs using single agent gene expression and biological network data (Cüvitoğlu and Işık, 2017). SVMs have also been used in other cancer applications, such as in the identification of cancer methylation signatures, in the prediction of response to chemotherapy, and for analyzing the risk of treatment resistance and tumor recurrence (Hu *et al.*, 2011; Dorman *et al.*, 2016; Jiang *et al.*, 2019; Wang *et al.*, 2020). However, the accuracy of SVMs is often still less than that of complex decision tree-based models such as random forest or XGBoost (Uddin *et al.*, 2019).

Decision trees are a relatively popular classifier machine learning model that takes in data at a root node and continues by some test rule, representing a branch, until the model reaches a decision, or leaf node (Rafique, Islam and Kazi, 2021). These leaf nodes then further branch into the categories of interest by which observations in the data are classified. Approaches based on decision trees include random forest models, gradient boosting, and XGBoost (Güvenç Paltun, Kaski and Mamitsuka, 2021; James *et al.*, 2021). These models are all ensemble approaches, meaning that each model is a combination of several less complex models, where each sub-model is a decision tree. Random forests select a random subset of data from a given dataset, train each model in its ensemble independently, and then use the majority decision from each sub-model to place each observation into a classification category (Srivastava, Kumar and Kumar, 2023). While random forest models combine their sub-models in parallel, gradient boosting and XGBoost combine their decision trees in series (Chen and Guestrin, 2016; Srivastava, Kumar and Kumar, 2023). This allows each sequential sub-model to improve upon the prediction of the previous sub-model. XGBoost additionally applies regularization, expanding how applicable the algorithm is to datasets outside of those used to initially train the model, thus enhancing the generalizability of these models compared to gradient boosting (Chen and Guestrin, 2016).

Celebi *et al.* compared several machine learning methods to discern which model performed best in predicting synergistic anti-cancer drug combinations (Celebi *et al.*, 2019). Although random forest and XGBoost both performed better than linear regression or support vector machines, XGBoost outperformed random forest after the models were tuned to maximize their performance, so the authors proceeded with XGBoost for all downstream analyses. While decision tree-based methods are interpretable and perform well, their accuracy is generally lower than deep learning approaches.

Deep Learning Models

Deep learning refers to a subclass of machine learning methods capable of handling large amounts of multi-dimensional data that often overwhelms other machine learning methods (Baptista, Ferreira and Rocha, 2021). Deep learning models are based on units of artificial neural networks, which are multi-layered networks composed of several processing layers (Figure 1D). These layers allow the model to learn and make predictions from complex mathematical functions (LeCun, Bengio and Hinton, 2015). Not only can deep learning incorporate larger quantities and more complex data types than other machine methods, but this ability to use multi-faceted data also allows deep learning methods to discern significant biological relationships that may not be detected by other machine learning approaches (Baptista, Ferreira and Rocha, 2021). However, the disadvantage of using

numerous features in creating a deep learning model is that it may result in overfitting, an issue in machine learning where the model is fitted too close to the data set used to train it, and is thus unable to generate accurate results for new data sets (Kernbach and Staartjes, 2022). Generalizability is thus a concern when developing deep learning models. Another limitation of deep learning techniques is simply the lack of adequate data for these models, as most deep learning approaches for predicting drug response are trained on limited numbers of cell lines. This then inevitably reduces their generalizability to densely heterogeneous patient tumors (Zhang *et al.*, 2021). This is further exemplified by Prasse *et al.*'s study, which found that fine-tuning deep neural networks on patient-derived data improves the accuracy of antineoplastic drug response predictions (Prasse *et al.*, 2022).

Despite these limitations, deep learning has still been immensely useful for advancing precision oncology. Deep learning has not only been used to predict several pharmacodynamic properties for drug discovery purposes, such as drug activity and toxicity, but it has also been shown to out-perform other machine learning methods for these tasks as well (Ma *et al.*, 2015; Mayr *et al.*, 2016; Korotcov *et al.*, 2017; Koutsoukas *et al.*, 2017; Lenselink *et al.*, 2017). In the context of cancer drug combination therapy, there have been several tools developed in recent years to predict potentially efficacious drug combination therapies for cancer, using already known antineoplastics or repurposing other approved medications for the disease (Table 5).

Table 5. Deep learning-based combination therapy prediction methods.

Model	Algorithm	Impact
DeepSynergy (Preuer et al., 2018)	Feed-forward neural network that calculates synergy scores of drug combinations for cell lines from drug chemical descriptors and cell line genomic information	Regarded as the first deep learning approach created for combination therapy prediction
CCSynergy (Hosseini and Zhou, 2023)	Uses drug bioactivity profiles from Chemical Checker to predict drug combination synergy	Enables exploration of the untested combinatorial space across hundreds of cell lines
MARSY (El Khili, Memon and Emad, 2023)	Multi-task model that combines the gene expression profiles of cancer cell lines with drug perturbation profiles to predict drug combination synergy scores	Incorporation of drug perturbation profiles allows for consideration of how pharmacologic agents alter gene expression in cell lines

DeepSynergy is regarded as the first deep learning approach developed for the prediction of drug combination synergies. DeepSynergy is a feed-forward neural network. It takes the chemical descriptors of each drug and the cell line genomic information as inputs to calculate synergy scores of drug combinations for cancer cell lines (Preuer *et al.*, 2018). Another example, CCSynergy is a deep-learning approach that uses drug bioactivity profiles from Chemical Checker for drug synergy prediction, and the use of CCSynergy to explore the untested combinatorial space revealed a compendium of potentially synergistic drug combinations across hundreds of cancer cell lines (Hosseini and Zhou, 2023). More recently, MARSY, a deep learning multi-task model that incorporates the gene expression profiles of cancer cell lines with drug perturbation profiles (i.e., the changes in gene expression of a cell line after drug treatment) was developed to predict synergy scores (El Khili, Memon and Emad, 2023). While these are currently the most recent deep learning approaches that have been developed for cancer drug combination prediction, more methods are in

active development that aim to incorporate multi-omic features to identify patient-specific anti-cancer drug regimens to further improve the specificity of predictions (Sharma *et al.*, 2023).

Discussion

We have described several computational methods developed to predict synergistic drug combinations to further precision oncology, including networks and machine-learning methods, such as regression models, classifier models, and deep learning frameworks. Each of these methods uses mathematical principles to complete various tasks in drug combination therapy prediction. Networks-based models allow for the visualization of patterns between drug and disease entities to identify candidate targets and therapies. Regression-based approaches can predict missing values in dose-response matrices to improve drug synergy calculations. Classifier methods and neural networks can predict potential anti-cancer therapies by sorting drug and disease data into categories. As these models are intended to perform specific tasks, the purpose of the study must be carefully considered when determining which of these methods to implement in one's own research. As noted by the DREAM Challenges, which compared several drug combination prediction tools for precision oncology against one another, the specific prediction algorithm matters far less than the principles it is based on and how it can be applied (Saez-Rodriguez *et al.*, 2016; Cichońska *et al.*, 2021).

Although the vast majority of drug combination prediction methods are still in the preclinical testing phase, they may soon transition to testing in randomized clinical trials. Recent clinical trials have shown promising results for the future of precision oncology as a whole. For example, the I-PREDICT and ongoing NCI-comboMATCH trials utilize next-generation sequencing (NGS)-guided matching protocols to pair patients to drug combination therapies. The results from the I-PREDICT study showed that a higher degree of matching correlated to improved patient outcomes, thus supporting the efficacy of precision combination therapy in clinical settings (Sicklick *et al.*, 2019; Meric-Bernstam *et al.*, 2023). The NCI-MATCH study utilized similar NGS methods on patient tumors to identify actionable genomic mutations across several cancer types. Although the patients treated by NGS-guided monotherapies showed improved progression-free survival compared to unmatched patients, only 3% of patients with refractory malignancies carried actionable mutations, demonstrating a need to broaden the scope of signature matching for candidate therapies via multi-omics integration (Flaherty *et al.*, 2020). The WINTHER trial was the first to match patients to drug combination therapies using a matching score based on both genomics and transcriptomics data. Not only was a higher matching score correlated with improved progression-free survival, but a significantly larger percentage of the patient cohort was able to be matched to targeted therapy regimens compared to the previously discussed NGS-guided trials, thus supporting the utility of multi-omics integration in guiding drug therapy prediction for cancer (Rodon *et al.*, 2019). More recent studies have attempted to expand the scope of multi-omics in combination therapy prediction even further. REFLECT is a machine learning method that incorporates mutational, copy number, transcriptomic, and phosphoproteomics data to generate detailed co-alteration signatures for therapy prediction (X. Li *et al.*, 2022). Given the dynamic nature of cancer across mutational signatures, copy number aberrations, and changes in gene expression due to gene regulatory alterations over time, genomics, transcriptomics, and epigenomics are the most promising -omics data to integrate to target the ever-changing landscape of cancer cell states (Barkley and Yanai, 2019). Proteomics integration also may be invaluable in the discovery of regimens including targeted therapies, such as receptor tyrosine kinase inhibitors (Bhullar *et al.*, 2018).

Aside from multi-omic integration, several clinical considerations are needed to further enhance the utility of current combination therapy prediction methods. Despite their clinical importance, toxicity and tolerability have been largely overlooked in the design of these prediction methods (Kong *et al.*, 2022). While maximizing drug synergy may enhance efficacy, studies often do not test for additional or compounding toxicities of these drug combinations, nor do they test for the specificity of these therapies for the disease over normal tissue (He *et al.*, 2018, 2021; Ianevski *et al.*, 2021). Current prediction methods also tend to not consider that drugs may have varying effects at different dosages, either alone or when interacting with another compound, and as such, drug

combination prediction algorithms should strive to predict the different interactions that can occur between drug combinations across drug dosages (Zimmer *et al.*, 2016, 2017b; Julkunen *et al.*, 2020). Finally, another needed advancement for cancer combination therapy prediction is methodology that can be used to monitor disease progression and response to treatment over time, such as Eduati *et al.*'s approach, which utilizes microfluidics and logic-based models to predict treatments for different stages of pancreatic cancer (Eduati *et al.*, 2020).

As computational methods improve to better incorporate patient-derived multi-omics data, disease-specific context, and pharmacodynamic considerations, more comprehensive models can be generated to predict effective drug regimens for complex diseases like cancer, reducing drug development time and cost and improving patient outcomes.

Summary:

- Computational drug repurposing is a time- and cost-effective alternative complementary to *de novo* drug discovery.
- Combination therapies have numerous advantages over monotherapies, including increased effect from synergistic interactions, reduced toxicity from lowered drug doses, and a reduced risk of resistance due to multiple non-overlapping mechanisms of action.
- Computational methods used for drug combination therapy prediction in cancer research include networks, regression-based machine learning, classifier machine learning models, and deep learning approaches.
- Advancements in technologies that incorporate disease mechanisms, drug characteristics, multi-omics data, and clinical considerations such as minimal toxicity and tolerability are needed to generate effective patient-specific drug combinations.

CRedit Statement: VLF: Conceptualization, Writing - Original Draft, Writing - Review & Editing, Visualization; JLF: Writing - Original Draft, Writing - Review & Editing; EJW: Writing - Original Draft, Writing - Review & Editing; TCH: Writing - Original Draft, Writing - Review & Editing; BNL: Conceptualization, Writing - Review & Editing, Supervision, Project Administration, Funding Acquisition.

Funding Statement: This work was supported by R03OD030604 (to BNL; supported JLF and BNL), UAB Lasseigne Lab funds (to BNL; supported JLF, EJW, TCH, and BNL), and 5T32GM008361-31 (supported VLF). The funders had no role in the conceptualization or writing of the manuscript.

References

- Alfarouk, K.O. *et al.* (2015) 'Resistance to cancer chemotherapy: failure in drug response from ADME to P-gp', *Cancer cell international*, 15, p. 71.
- Al-Lazikani, B., Banerji, U. and Workman, P. (2012) 'Combinatorial drug therapy for cancer in the post-genomic era', *Nature biotechnology*, 30(7), pp. 679–692.
- Amzallag, A., Ramaswamy, S. and Benes, C.H. (2019) 'Statistical assessment and visualization of synergies for large-scale sparse drug combination datasets', *BMC bioinformatics*, 20(1), p. 83.
- Ayoub, N.M. (2021) 'Editorial: Novel Combination Therapies for the Treatment of Solid Cancers', *Frontiers in oncology*, 11, p. 708943.
- Azuaje, F. (2013) 'Drug interaction networks: an introduction to translational and clinical applications', *Cardiovascular research*, 97(4), pp. 631–641.
- Baptista, D., Ferreira, P.G. and Rocha, M. (2021) 'Deep learning for drug response prediction in cancer', *Briefings in bioinformatics*, 22(1), pp. 360–379.
- Barkley, D. and Yanai, I. (2019) 'Plasticity and Clonality of Cancer Cell States', *Trends in cancer research*, pp. 655–656.
- Bhullar, K.S. *et al.* (2018) 'Kinase-targeted cancer therapies: progress, challenges and future directions', *Molecular cancer*, 17(1), p. 48.
- Booth, B. and Zimmel, R. (2004) 'Prospects for productivity', *Nature reviews. Drug discovery*, 3(5), pp. 451–456.
- Borst, P. (2012) 'Cancer drug pan-resistance: pumps, cancer stem cells, quiescence, epithelial to mesenchymal transition, blocked cell death pathways, persists or what?', *Open biology*, 2(5), p. 120066.

- Celebi, R. *et al.* (2019) 'In-silico Prediction of Synergistic Anti-Cancer Drug Combinations Using Multi-omics Data', *Scientific reports*, 9(1), pp. 1–10.
- Cheng, F., Kovács, I.A. and Barabási, A.-L. (2019) 'Network-based prediction of drug combinations', *Nature communications*, 10(1), p. 1197.
- Chen, T. and Guestrin, C. (2016) 'XGBoost: A Scalable Tree Boosting System', *arXiv [cs.LG]*. Available at: <http://arxiv.org/abs/1603.02754>.
- Cichońska, A. *et al.* (2021) 'Crowdsourced mapping of unexplored target space of kinase inhibitors', *Nature communications*, 12(1), p. 3307.
- Corsello, S.M. *et al.* (2017) 'The Drug Repurposing Hub: a next-generation drug library and information resource', *Nature medicine*, 23(4), pp. 405–408.
- Cüvitoğlu, A. and Işık, Z. (2017) 'Classification of effects of drug combinations with support vector machines', in *2017 25th Signal Processing and Communications Applications Conference (SIU)*, pp. 1–4.
- Dorman, S.N. *et al.* (2016) 'Genomic signatures for paclitaxel and gemcitabine resistance in breast cancer derived by machine learning', *Molecular oncology*, 10(1), pp. 85–100.
- Eduati, F. *et al.* (2020) 'Patient-specific logic models of signaling pathways from screenings on cancer biopsies to prioritize personalized combination therapies', *Molecular systems biology*, 16(2), p. e8664.
- El Khili, M.R., Memon, S.A. and Emad, A. (2023) 'MARSY: a multitask deep-learning framework for prediction of drug combination synergy scores', *Bioinformatics*, 39(4). Available at: <https://doi.org/10.1093/bioinformatics/btad177>.
- Federico, A. *et al.* (2022) 'Integrated Network Pharmacology Approach for Drug Combination Discovery: A Multi-Cancer Case Study', *Cancers*, 14(8). Available at: <https://doi.org/10.3390/cancers14082043>.
- Firoozbakht, F., Yousefi, B. and Schwikowski, B. (2022) 'An overview of machine learning methods for monotherapy drug response prediction', *Briefings in bioinformatics*, 23(1). Available at: <https://doi.org/10.1093/bib/bbab408>.
- Flaherty, K.T. *et al.* (2020) 'The Molecular Analysis for Therapy Choice (NCI-MATCH) Trial: Lessons for Genomic Trial Design', *Journal of the National Cancer Institute*, 112(10), pp. 1021–1029.
- Güvenç Paltun, B., Kaski, S. and Mamitsuka, H. (2021) 'Machine learning approaches for drug combination therapies', *Briefings in bioinformatics*, 22(6). Available at: <https://doi.org/10.1093/bib/bbab293>.
- Hare, D. and Foster, T. (1990) 'The Orange Book: the Food and Drug Administration's advice on therapeutic equivalence', *American pharmacy*, NS30(7), pp. 35–37.
- He, H. *et al.* (2023) 'Computational drug repurposing by exploiting large-scale gene expression data: Strategy, methods and applications', *Computers in biology and medicine*, 155, p. 106671.
- He, L. *et al.* (2018) 'Patient-Customized Drug Combination Prediction and Testing for T-cell Prolymphocytic Leukemia Patients', *Cancer research*, 78(9), pp. 2407–2418.
- He, L. *et al.* (2021) 'Network-guided identification of cancer-selective combinatorial therapies in ovarian cancer', *Briefings in bioinformatics*, 22(6). Available at: <https://doi.org/10.1093/bib/bbab272>.
- Holohan, C. *et al.* (2013) 'Cancer drug resistance: an evolving paradigm', *Nature reviews. Cancer*, 13(10), pp. 714–726.
- Hosseini, S.-R. and Zhou, X. (2023) 'CCSynergy: an integrative deep-learning framework enabling context-aware prediction of anti-cancer drug synergy', *Briefings in bioinformatics*, 24(1). Available at: <https://doi.org/10.1093/bib/bbac588>.
- Housman, G. *et al.* (2014) 'Drug resistance in cancer: an overview', *Cancers*, 6(3), pp. 1769–1792.
- Hu, X. *et al.* (2011) 'Support vector machine multiparametric MRI identification of pseudoprogression from tumor recurrence in patients with resected glioblastoma', *Journal of magnetic resonance imaging: JMRI*, 33(2), pp. 296–305.
- Ianevski, A. *et al.* (2019) 'Prediction of drug combination effects with a minimal set of experiments', *Nature machine intelligence*, 1(12), pp. 568–577.
- Ianevski, A. *et al.* (2021) 'Patient-tailored design for selective co-inhibition of leukemic cell subpopulations', *Science advances*, 7(8). Available at: <https://doi.org/10.1126/sciadv.abe4038>.
- Iwata, H. *et al.* (2015) 'Large-Scale Prediction of Beneficial Drug Combinations Using Drug Efficacy and Target Profiles', *Journal of chemical information and modeling*, 55(12), pp. 2705–2716.
- Jafari, M. *et al.* (2022) 'Bipartite network models to design combination therapies in acute myeloid leukaemia', *Nature communications*, 13(1), pp. 1–12.
- James, G. *et al.* (2021) *An Introduction to Statistical Learning with Applications in R Second Edition*. Springer Nature.

- Jarada, T.N., Rokne, J.G. and Alhaji, R. (2020) 'A review of computational drug repositioning: strategies, approaches, opportunities, challenges, and directions', *Journal of cheminformatics*, 12(1), p. 46.
- Jiang, Y. *et al.* (2019) 'Tumor Immune Microenvironment and Chemosensitivity Signature for Predicting Response to Chemotherapy in Gastric Cancer', *Cancer immunology research*, 7(12), pp. 2065–2073.
- Julkunen, H. *et al.* (2020) 'Leveraging multi-way interactions for systematic prediction of pre-clinical drug combination effects', *Nature communications*, 11(1), p. 6136.
- Kanehisa, M. *et al.* (2008) 'KEGG for linking genomes to life and the environment', *Nucleic acids research*, 36(Database issue), pp. D480–4.
- Kernbach, J.M. and Staartjes, V.E. (2022) 'Foundations of Machine Learning-Based Clinical Prediction Modeling: Part II-Generalization and Overfitting', *Acta neurochirurgica. Supplement*, 134, pp. 15–21.
- Kong, W. *et al.* (2022) 'Systematic review of computational methods for drug combination prediction', *Computational and structural biotechnology journal*, 20, pp. 2807–2814.
- Korotcov, A. *et al.* (2017) 'Comparison of Deep Learning With Multiple Machine Learning Methods and Metrics Using Diverse Drug Discovery Data Sets', *Molecular pharmaceuticals*, 14(12), pp. 4462–4475.
- Koscielny, G. *et al.* (2017) 'Open Targets: a platform for therapeutic target identification and validation', *Nucleic acids research*, 45(D1), pp. D985–D994.
- Koutsoukas, A. *et al.* (2017) 'Deep-learning: investigating deep neural networks hyper-parameters and comparison of performance to shallow methods for modeling bioactivity data', *Journal of cheminformatics*, 9(1), p. 42.
- LeCun, Y., Bengio, Y. and Hinton, G. (2015) 'Deep learning', *Nature*, 521(7553), pp. 436–444.
- Lenselink, E.B. *et al.* (2017) 'Beyond the hype: deep neural networks outperform established methods using a ChEMBL bioactivity benchmark set', *Journal of cheminformatics*, 9(1), p. 45.
- Li, C. *et al.* (2022) 'A software tool for both the maximum tolerated dose and the optimal biological dose finding trials in early phase designs', *Contemporary clinical trials communications*, 30, p. 100990.
- Li, D.H. *et al.* (2017) 'Toxicity and Efficacy Probability Interval Design for Phase I Adoptive Cell Therapy Dose-Finding Clinical Trials', *Clinical cancer research: an official journal of the American Association for Cancer Research*, 23(1), pp. 13–20.
- Li, J., Xu, H. and McIndoe, R.A. (2022) 'A novel network based linear model for prioritization of synergistic drug combinations', *PloS one*, 17(4), p. e0266382.
- Li, X. *et al.* (2022) 'Precision Combination Therapies Based on Recurrent Oncogenic Coalterations', *Cancer discovery*, 12(6), pp. 1542–1559.
- Ma, J. *et al.* (2015) 'Deep neural nets as a method for quantitative structure-activity relationships', *Journal of chemical information and modeling*, 55(2), pp. 263–274.
- Mayr, A. *et al.* (2016) 'DeepTox: Toxicity Prediction using Deep Learning', *Frontiers of Environmental Science & Engineering in China*, 3. Available at: <https://doi.org/10.3389/fenvs.2015.00080>.
- Menche, J. *et al.* (2015) 'Disease networks. Uncovering disease-disease relationships through the incomplete interactome', *Science*, 347(6224), p. 1257601.
- Meric-Bernstam, F. *et al.* (2023) 'National Cancer Institute Combination Therapy Platform Trial with Molecular Analysis for Therapy Choice (ComboMATCH)', *Clinical cancer research: an official journal of the American Association for Cancer Research*, 29(8), pp. 1412–1422.
- Müller-Dott, S. *et al.* (2023) 'Expanding the coverage of regulons from high-confidence prior knowledge for accurate estimation of transcription factor activities', *bioRxiv*. Available at: <https://doi.org/10.1101/2023.03.30.534849>.
- Ning, Y.-M. *et al.* (2010) 'Phase II trial of bevacizumab, thalidomide, docetaxel, and prednisone in patients with metastatic castration-resistant prostate cancer', *Journal of clinical oncology: official journal of the American Society of Clinical Oncology*, 28(12), pp. 2070–2076.
- Pan, H. *et al.* (2020) 'Keyboard design for phase I drug-combination trials', *Contemporary clinical trials*, 92, p. 105972.
- Park, M., Nassar, M. and Vikalo, H. (2013) 'Bayesian active learning for drug combinations', *IEEE transactions on bio-medical engineering*, 60(11), pp. 3248–3255.
- Pemovska, T., Bigenzahn, J.W. and Superti-Furga, G. (2018) 'Recent advances in combinatorial drug screening and synergy scoring', *Current opinion in pharmacology*, 42, pp. 102–110.

- Prasse, P. *et al.* (2022) 'Pre-Training on In Vitro and Fine-Tuning on Patient-Derived Data Improves Deep Neural Networks for Anti-Cancer Drug-Sensitivity Prediction', *Cancers*, 14(16). Available at: <https://doi.org/10.3390/cancers14163950>.
- Preuer, K. *et al.* (2018) 'DeepSynergy: predicting anti-cancer drug synergy with Deep Learning', *Bioinformatics*, 34(9), pp. 1538–1546.
- Rafique, R., Islam, S.M.R. and Kazi, J.U. (2021) 'Machine learning in the prediction of cancer therapy', *Computational and structural biotechnology journal*, 19, pp. 4003–4017.
- Rodon, J. *et al.* (2019) 'Genomic and transcriptomic profiling expands precision cancer medicine: the WINTHER trial', *Nature medicine*, 25(5), pp. 751–758.
- Rueff, J. and Rodrigues, A.S. (2016) 'Cancer Drug Resistance: A Brief Overview from a Genetic Viewpoint', in J. Rueff and A.S. Rodrigues (eds) *Cancer Drug Resistance: Overviews and Methods*. New York, NY: Springer New York, pp. 1–18.
- Saez-Rodriguez, J. *et al.* (2016) 'Crowdsourcing biomedical research: leveraging communities as innovation engines', *Nature reviews. Genetics*, 17(8), pp. 470–486.
- Sarmah, D. *et al.* (2023) 'Predicting anti-cancer drug combination responses with a temporal cell state network model', *PLoS computational biology*, 19(5), p. e1011082.
- Scannell, J.W. *et al.* (2012) 'Diagnosing the decline in pharmaceutical R&D efficiency', *Nature reviews. Drug discovery*, 11(3), pp. 191–200.
- Sharma, A. *et al.* (2023) 'DeepInsight-3D architecture for anti-cancer drug response prediction with deep-learning on multi-omics', *Scientific reports*, 13(1), p. 2483.
- Shim, J.S. and Liu, J.O. (2014) 'Recent advances in drug repositioning for the discovery of new anticancer drugs', *International journal of biological sciences*, 10(7), pp. 654–663.
- Sicklick, J.K. *et al.* (2019) 'Molecular profiling of cancer patients enables personalized combination therapy: the I-PREDICT study', *Nature medicine*, 25(5), pp. 744–750.
- Singhal, S. *et al.* (1999) 'Antitumor activity of thalidomide in refractory multiple myeloma', *The New England journal of medicine*, 341(21), pp. 1565–1571.
- Srivastava, R., Kumar, S. and Kumar, B. (2023) '7 - Classification model of machine learning for medical data analysis', in T. Goswami and G.R. Sinha (eds) *Statistical Modeling in Machine Learning*. Academic Press, pp. 111–132.
- Sun, D. *et al.* (2022) 'Why 90% of clinical drug development fails and how to improve it?', *Acta pharmaceutica Sinica. B*, 12(7), pp. 3049–3062.
- Szklarczyk, D. *et al.* (2023) 'The STRING database in 2023: protein-protein association networks and functional enrichment analyses for any sequenced genome of interest', *Nucleic acids research*, 51(D1), pp. D638–D646.
- Uddin, S. *et al.* (2019) 'Comparing different supervised machine learning algorithms for disease prediction', *BMC medical informatics and decision making*, 19(1), p. 281.
- Wang, X. *et al.* (2020) 'Methylation signature genes identification of cancers occurrence and pattern recognition', *Computational biology and chemistry*, 85, p. 107198.
- Wilkinson, G.F. and Pritchard, K. (2015) 'In vitro screening for drug repositioning', *Journal of biomolecular screening*, 20(2), pp. 167–179.
- Yadav, B. *et al.* (2015) 'Searching for Drug Synergy in Complex Dose-Response Landscapes Using an Interaction Potency Model', *Computational and structural biotechnology journal*, 13, pp. 504–513.
- Yan, F., Mandrekar, S.J. and Yuan, Y. (2017) 'Keyboard: A Novel Bayesian Toxicity Probability Interval Design for Phase I Clinical Trials', *Clinical cancer research: an official journal of the American Association for Cancer Research*, 23(15), pp. 3994–4003.
- Zhang, T. *et al.* (2021) 'Synergistic Drug Combination Prediction by Integrating Multiomics Data in Deep Learning Models', *Methods in molecular biology*, 2194, pp. 223–238.
- Zhou, Y. *et al.* (2022) 'Therapeutic target database update 2022: facilitating drug discovery with enriched comparative data of targeted agents', *Nucleic acids research*, 50(D1), pp. D1398–D1407.
- Zimmer, A. *et al.* (2016) 'Prediction of multidimensional drug dose responses based on measurements of drug pairs', *Proceedings of the National Academy of Sciences of the United States of America*, 113(37), pp. 10442–10447.

Zimmer, A. *et al.* (2017a) 'Prediction of drug cocktail effects when the number of measurements is limited', *PLoS biology*, 15(10), p. e2002518.

Zimmer, A. *et al.* (2017b) 'Prediction of drug cocktail effects when the number of measurements is limited', *PLoS biology*, 15(10), p. e2002518.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.