# SynDRep: a synergistic partner prediction tool based on knowledge graph for drug repurposing

Karim S. Shalaby<sup>1,2®</sup>, Sathvik Guru Rao<sup>1</sup>, Bruce Schultz<sup>3®</sup>, Martin Hofmann-Apitius<sup>1,4®</sup>, Alpha Tom Kodamullil<sup>1,5,\*,†®</sup>, Vinay Srinivas Bharadhwaj<sup>1,4,†</sup>

<sup>1</sup>Department of Bioinformatics, Fraunhofer Institute for Algorithms and Scientific Computing (SCAI), Sankt Augustin 53757, Germany <sup>2</sup>Department of Pharmaceutics and Industrial Pharmacy, Faculty of Pharmacy, Ain Shams University, Cairo 11566, Egypt <sup>3</sup>Institute of Biomedical Informatics, University Hospital Cologne, Cologne 50937, Germany

<sup>4</sup>Bonn-Aachen International Center for Information Technology (B-IT), University of Bonn, Bonn 53115, Germany <sup>5</sup>Causality Biomodels, Cochin 683503, India

\*Corresponding author. Department of Bioinformatics, Fraunhofer Institute for Algorithms and Scientific Computing (SCAI), Schloss Birlinghoven 1, Sankt Augustin 53757, Germany. E-mail: alpha.tom.kodamullil@scai.fraunhofer.de.

 $^{\dagger}$  = equal contribution.

Associate Editor: Yoshihiro Yamanishi

#### Abstract

**Motivation:** Drug repurposing is gaining interest due to its high cost-effectiveness, low risks, and improved patient outcomes. However, most drug repurposing methods depend on drug-disease-target semantic connections of a single drug rather than insights from drug combination data. In this study, we propose SynDRep, a novel drug repurposing tool based on enriching knowledge graphs (KG) with drug combination effects. It predicts the synergistic drug partner with a commonly prescribed drug for the target disease, leveraging graph embedding and machine learning (ML) techniques. This partner drug is then repurposed as a single agent for this disease by exploring pathways between them in the KG.

**Results:** HolE was the best-performing embedding model (with 84.58% of true predictions for all relations), and random forest emerged as the best ML model with an area under the receiver operating characteristic curve (ROC-AUC) value of 0.796. Some of our selected candidates, such as miconazole and albendazole for Alzheimer's disease, have been validated through literature, while others lack either a clear pathway or literature evidence for their use for the disease of interest. Therefore, complementing SynDRep with more specialized KGs, and additional training data, would enhance its efficacy and offer cost-effective and timely solutions for patients.

Availability and implementation: SynDRep is available as an open-source Python package at https://github.com/SynDRep/SynDRep under the Apache 2.0 License.

# **1** Introduction

Despite tremendous technological, regulatory, and scientific advances that increase the efficiency of drug research and development, the resulting therapeutic outcomes need to catch up with the corresponding spending on these advances (Ashburn and Thor 2004, Scannell *et al.* 2012). Additionally, the rising cost and time required to develop new drugs have resulted in lower profits for the pharmaceutical sector and a longer response time to disease outbreaks (Pushpakom *et al.* 2018). Conversely, drug repurposing, i.e. finding novel indications for current drugs, has advantages over *de novo* drug development, including shorter development time and lower cost risks (Choudhury *et al.* 2022, Hua *et al.* 2022), since compounds already investigated and approved by regulatory bodies, incorporating safety and efficacy profiles, can be reassessed critically in a new therapeutic context (Lage-Rupprecht *et al.* 2022).

In recent years, drug repurposing research has greatly benefited from the exploding growth of biomedical databases. Therefore, plenty of computational techniques have been devised to analyze different biomedical data systematically to hypothesize new indications for a drug or to find new drugs for a specific disease (Jarada et al. 2020, Luo et al. 2021, Pan et al. 2022). Computational drug repurposing approaches are mostly data-driven; they encompass the systematic analysis of data from various modalities, e.g. chemical structure, proteomic data, gene expression, genotype, or electronic health records, which can then drive the repurposing hypotheses (Hurle et al. 2013, Zong et al. 2022). For practical analysis of such vast data types, measures for appropriately aggregating them in an informative manner need to be taken. One of these measures is the organization and representation of data into a knowledge graph (KG), which aids in identifying semantic connections between multiple resources and allows for knowledge reasoning (Chen et al. 2020, Gao et al. 2022). Extending these mechanistic KGs with drug-related data to form drug-target-mechanism-oriented data models results in so-called PHARMACOMES (Lage-Rupprecht et al. 2022).

Received: October 23, 2024; Revised: March 24, 2025; Editorial Decision: March 29, 2025; Accepted: June 4, 2025

© The Author(s) 2025. Published by Oxford University Press. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

Pharmacomes with their integration of pathophysiology mechanisms, drug targets, and drugs/compounds offer the possibility to look at dual targeting strategies and combinatorial targeting of different pathophysiology mechanisms through combinations of drug repurposing candidates. Drug combinations offer excellent efficacy in treating multifactorial diseases involving more than one genetic pathway, such as cancer (Zhou et al. 2023), diabetes (Dubourg et al. 2022), Alzheimer's disease (AD) (Knorz and Quante 2022), and cardiovascular diseases (Lombardi et al. 2020). In principle, they also offer the option to specifically target comorbidity pathways. Therefore, incorporating new links among drugs into pharmacomes, indicating drug combinations, paves the way for developing new synergistic drug combinations. It warns of potential drug-drug interactions in a more comprehensive way that depends on direct as well as indirect links between drugs. Applying some link prediction algorithms afterward will predict new drug relationships, gain more insights into drug mechanisms, and eventually repurpose drug candidates for various diseases.

For a long time, drug synergy studies depended on trial and error, which suffered from high labor and time costs and exposes patients to ineffective treatment or undesirable side effects (Pang et al. 2014, Day and Siu 2016). This was then replaced by high-throughput screening (HTS), where many measurements can be produced reasonably quickly and at a lower cost (He et al. 2018). During HTS, different concentrations of two or more drugs are applied to a cell line. However, the high genomic correlation between the original tissues and the derived cell lines remains imperfect (Ferreira et al. 2013). Moreover, HTS cannot cover the whole combination space for drugs (Goswami et al. 2015). Computational methods such as deep learning and machine learning (DL/ML) models can efficiently explore the vast synergistic space using the available HTS synergy data. Recent methods range from systems biology (Feala et al. 2010), kinetic models (Sun et al. 2016), mixed integer linear programming (Pang et al. 2014), computational methods based on drug-induced gene expression profile and dose-response curves (Goswami et al. 2015), to ML approaches including random forests and Naive Baves methods (Li et al. 2015, Wildenhain et al. 2015), and DL approaches such as deep neural networks, graph autoencoder, and convolutional neural network (Preuer et al. 2018, Kuenzi et al. 2020, Sun et al. 2020, Kim et al. 2021, Liu and Xie 2021, Li et al. 2023). However, these methods are restricted to predicting synergistic combinations and do not consider drug synergy prediction as an intermediate step in the drug repurposing process. In our approach, we leverage the synergistic prediction as a foundation for the repurposing process.

We propose a new drug repurposing tool (SynDRep), which depends on enriching knowledge graphs with drug combination effects. Our approach selects repurposing candidates, by predicting synergistic drug partners of a commonly prescribed drug for the target disease. This is followed by the selection of "safe drug partners" as a single-agent therapy for the disease. The drug's candidacy for repurposing is confirmed by exploring the pathway within the KG between the drug and the target disease. Additionally, experimental evidence about the beneficial effect of the candidate on target disease supports the repurposing profile. Therefore, this approach combines the speed and cost reduction of the computational approach with the accuracy and certainty of manual curation and expands the current drug repurposing landscape with a new concept relying not only on drug-disease-target semantic connections but also on the drug-drug synergy effect.

#### 2 Methods

# 2.1 Overall workflow

The primary objective entailed the integration of drug-drug relationships into an established KG to serve as a foundational framework for subsequent drug synergy prediction and repurposing (Fig. 1). The process begins with gathering and refining data from drug combination databases, where these combinations are categorized into synergism or antagonism and then input into a neo4j instance of our chosen KG. Next, ML methods were performed based on drug physicochemical properties and enriched KG topological features. Due to the inability to validate ML model predictions, we used embedding models to predict the missing inter-drug relations in the enriched KG. However, the selected KG included certain ambiguous relationships, such as "association" and "in complex with," as well as hub nodes that have numerous connecting relationships, particularly the disease nodes. Therefore, we repeated embedding on another version of the same KG after the removal of noncausal relations and hub nodes. To overcome the underperformance of embedding models experienced in this case, we applied a combination of embedding followed by ML modeling of resultant embedding vectors. Ultimately, we identify repurposing candidates by predicting synergistic drug partners for commonly prescribed medications related to the target disease and considering safe partners as a single agent for this condition. Candidate profiles are validated by investigating the pathways present within the KG connecting the candidate drug to the target disease. Moreover, experimental evidence from the literature supporting the candidate's positive impact on the target disease further reinforces the repurposing profile.

#### 2.2 Data collection

The initial phase involved the careful selection of a comprehensive KG, such as the Human Brain Pharmacome (HBP) and appropriate sources for drug–drug combinations such as drug synergy databases, e.g. DrugcombDB, DrugcombPortal, and SYNERGxDB.

## 2.2.1 Human brain pharmacome

The KG selected for this study was HBP (pharmacome database at https://graphstore.scai.fraunhofer.de, downloaded on 25 September 2023). HBP combines knowledge from various sources with a focused drug-target-mechanism-oriented data model. It contains information curated from bibliographic databases such as PubMed, pathway databases such as Reactome, KEGG, and Pathway Commons, protein-protein interaction databases such as IntAct, BioGRID, and StringDB, and drug databases such as DrugBank, Clinical Trials, Sider, and ChEMBL. (Lage-Rupprecht et al. 2022). HBP is a comprehensive KG consisting of 136838 nodes, which represent 27 node types, including biological and molecular entities as well as biological processes, such as "Gene," "Protein," "Drug," "SNP," "BiologicalProcess," "Pathology," "Complex," "Degradation," and "Protein Modification." These nodes are interconnected through 731974 edges, representing the interactions or relations



**Figure 1.** The overall workflow of the study, including the Python packages and data sources used. The work starts with data collection and refinement from combination databases using Pandas and Requests, then the synergy data is fed into a neo4j instance of the KG. Third, ML using scikit-learn, embedding using PyKEEN, or embedding followed by ML was used to model and predict the synergies. Finally, the identification of repurposing candidates by predicting synergistic drug partners for commonly prescribed drugs for the target disease and repurposing safe partners as a single agent for this disease. Candidate profiles are confirmed by examining the existence of pathways within the knowledge graph between the candidate drug and the target disease using Pandas and NetworkX. Experimental evidence from the literature supporting the candidate's beneficial effect on the target disease further validates the repurposing profile.

between them. The edges have 72 types, including causal relations like "increases," "decreases," or "causes\_no\_change" and noncausal relations such as "association," "has\_variant," or "equivalent\_to."The data of this pharmacome has been extracted from the online source, stored locally using neo4j (version 5.22.0), and added to our GitHub page (https://github.com/SynDRep/SynDRep/blob/main/ Data/Human\_Brain\_Pharmacome/kg.tsv). It formed the base for the next step of HBP enrichment.

#### 2.2.2 Drug synergy databases

To accommodate our expansive HBP, we selected databases that contain the highest number of drugs. Drug combination effects have been gathered from DrugcombDB (Liu *et al.* 2020), DrugcombPortal (Zagidullin *et al.* 2019), and SYNERGxDB (Seo *et al.* 2020). The scores for synergism models, such as the highest single agent (HSA) model (Berenbaum 1989), Bliss model (Bliss 1939), Loewe model (Loewe 1953), and the zero interaction potency (ZIP) model (Yadav *et al.* 2015), were used to supplement the new edges created in the next step. These models consider in their calculation the different effects of drug combinations at different drug concentrations.

# 2.3 KG enrichment

The KG enrichment was done over several steps. Data extracted from drug combination databases were deduplicated. The synergism scores from the same combinations with different scores were averaged. Drug combinations were tested across different cell lines (359 cell lines in total), such as CBRC058, NCI-H322M, and KBM-7 cell lines. Some combinations of drugs produce synergism in one cell line and antagonism in another. Therefore, to remove the effect of different cell types, we selected only combinations that produced either synergism or antagonism across different cell types. Moreover, to go with a standardized approach, we selected only one synergism score (ZIP score), as the ZIP model encompasses the Loewe additivity and the Bliss independence. In addition, it is more accurate at detecting potency changes in drug combinations compared to HSA and Bliss independence models (Yadav *et al.* 2015).

## 2.4 Classical machine learning

In order to predict new synergistic relations between drugs, we started with the classical ML approaches, to assess their ability and efficiency for link prediction compared to KG embedding. Five ML models, namely: logistic regression, elastic net, gradient boosting, random forest, and support vector machine, were selected along with the features of each pair of drugs to classify their combination either into synergism or antagonism. These features encompassed aspects related to the KG, as well as physicochemical attributes of the drugs as depicted in Table 1. KG features, which depend on the network structure and topological features were extracted from HBP using NetworkX (version 7.1.3) (Hagberg et al. 2008), a Python package for the creation and study of the structure of complex networks. Physicochemical attributes of the drugs were extracted from PubChem or computed using the RDKit Python package (version 2024.3.2) (Landrum 2023). The

Table 1. The features and labels used to train the different machine learning models.

Features and metrics		Labels	Models
KG-related	Physicochemical		
Drug node degree	Molecular weight	Synergism	Logistic regression
Drug node clustering coefficient	Log P	Antagonism	Elastic net
Drug node page rank	Total polar surface area	0	Gradient boosting
Shortest path length	Number of hydrogen bond donors		Random forest
Cosine similarity	Number of hydrogen bond acceptors		Support vector machine
	Rotatable bond count		
	Tanimoto coefficient		
	Morgan fingerprint		

features, labels, and models used are listed in Table 1. The classification was performed using a 10-fold nested cross-validation approach with a Grid Search optimizer, and the model performance was assessed based on the area under the receiver operating characteristic curve (ROC-AUC) values for all models. The data was initially split into a training set (90% of the dataset, corresponding to 20 840 drug pairs) and a hold-out test set (10% of the dataset, corresponding to 2316 drug pairs). During the inner cross-validation loops, the training data was further split into 90% training (18 756 drug pairs) and 10% validation (2084 drug pairs) for hyper-parameter optimization (HPO). After completing HPO, 10 model instances were trained with the best parameters obtained from the inner loops and were subsequently evaluated on the held-out test set using the ROC-AUC score.

Following a comprehensive assessment of all models and the calculation of ROC-AUC. Since the elastic net model exhibited the highest ROC-AUC, we utilized it to predict the synergistic interactions between each pair of drugs in HBP. Subsequently, combinations predicted as synergistic were chosen to constitute the predicted synergism set. A literature check was conducted to validate the top five synergistic combinations based on their predicted probabilities during the initial prediction process. This validation step aimed to ensure the credibility and accuracy of the model's predictions by cross-referencing them with existing scientific literature.

#### 2.5 KG embedding

To embed the enriched HBP, we used PyKEEN (Python KnowlEdge EmbeddiNgs) (Ali *et al.* 2021), a Python package designed for training and evaluation of KG embedding models. We worked under stochastic local closed world assumption (SLCWA), where a randomized subset is drawn from the combination of head and tail generation strategies, initially defined in local closed world assumption, and these selected triples are treated as negatives. This approach offers several advantages, including the lower load of computation and the flexibility to include new negative sampling strategies.

## 2.5.1 Data splitting

To prevent overfitting, the set of triples (source–relation– target) that make up the network structure of the enriched HBP was then stratified using the PyKEEN into a training set (80%) and a test set (20%). To prevent the dissemination of the test set into the training set during the HPO or training of the model, we isolated the test set, and the training set was further split into training (80%) and validation (20%) sets (Supplementary Fig. S1). We checked that each split contained the corresponding percentage of triples and that the training set contained all the relation types in HBP to ensure that the test and validation sets did not contain any relation type new to the model after training. To further assess the model's efficiency in predicting drug–drug relations, one more test set was formed from the original test set, the drug–drug test set, which contained only the drug–drug relations from the test set. These two sets were used to evaluate model performance.

## 2.5.2 Model selection

We selected five models for embedding HBP: TransE (Bordes *et al.* 2013), TransR, RotatE (Lin *et al.* 2015), ComplEx (Trouillon *et al.* 2016), and HolE (Nickel *et al.* 2015). We afterward used the best-performing model to predict the new synergistic or antagonistic relations. Predictions were made using the two entities as head and tail, and the model predicts the relation type between them. The output of the prediction model is a ranking of the possible relations between these two entities according to a score produced by the model. Therefore, we selected the first three predictions to assess the model's performance by calculating the percentage of true prediction in each rank compared to all predictions in this rank (Equation 1):

$$Percentage of true predictions = \frac{Number of true predictions in a rank}{Total number of prediction in the same rank} X100$$
(1)

We assessed the percentage of true predictions at the first rank to determine its suitability for further analysis. Furthermore, the second and third ranks were also selected to check whether false predictions at rank 1 had corresponding true predictions at ranks 2 or 3, demonstrating the ability of the embedding model to enrich the true predictions in the highest ranks. This led us to exclude ranks beyond three, as most of the true predictions were within these three ranks. Additionally, we calculated the multi-class ROC-AUC using the highest-ranked prediction for each pair of drugs in the test set. This involved converting the actual and predicted relation types into binary form and then averaging the ROC-AUC values for all relation types, as shown in the following Equation 2:

$$Multiclass \ ROC - AUC = \frac{\Sigma \ ROC - AUC \ for \ each \ relation \ type}{Total \ number \ of \ relation \ types}$$
(2)

Based on the values of the percentage of true predictions and multi-class ROC-AUC, we selected the best-performing

model for further prediction of the drug-drug relations that are not in HBP.

## 2.6 Synergy prediction

After assessing all models and calculating the percentage of true prediction, we selected RotatE to predict all drug-drug relations further. Utilizing the trained RotatE model, we predicted the relations between each drug pair within HBP. Both the forward case (with drug A as a head and drug B as a tail) and the reverse case (with drug B as a head and drug A as a tail) were predicted. The predicted relationship which ranked as the first was then extracted for each drug pair. To refine our dataset, we eliminated cases where nonmutual synergism or antagonism was observed. Then, we segregated the predicted dataset into synergism and antagonism categories, focusing on selecting data that showed synergistic interactions for subsequent in-depth analysis. A comprehensive literature review was conducted to validate the model's predictive power, particularly for the predictions with the highest scores.

#### 2.7 Drug repurposing

The synergy effect frequently stems from different drugs having influences on the same, parallel or even different pathways essential for an observed phenotype, and synergy is induced by targets aggregating at specific pathways that control the state of the disease (Cokol et al. 2011, Chen et al. 2015). Consequently, we propose that if predicted synergy partners have a shared pathway in HBP related to the target disease, it is likely that the partner, to a commonly prescribed drug to the target disease, could be repurposed as a standalone treatment. While the existence of a pathway alone is insufficient to support this hypothesis, it does narrow the pool of potential repurposing candidates by focusing on drugs with both synergy and a shared pathway. As such, the selected candidates must undergo validation through in vitro or in vivo studies. Since the primary focus of HBP is on molecular interactions and patho-mechanisms within the brain and their relation to neurodegenerative diseases (NDDs), we chose AD and schizophrenia as model diseases to align with the scope of HBP. Therefore, we evaluated predicted synergistic combinations involving drugs used for AD, schizophrenia, and bipolar disorder to assess the feasibility of repurposing their synergistic partners for these conditions. Based on our predictions, we selected a list of drugs that exhibited the highest-scoring synergistic combinations for each drug. It is noteworthy that our selection criteria excluded drugs with cytotoxic or severe side effects, such as anticancer or carcinogenic drugs, ensuring that the chosen repurposing candidates prioritize safety considerations. This approach was initially followed by a meticulous search for a possible common pathway in HBP between the two drugs in the combination and the disease. Subsequently, we reviewed the literature for possible studies about using these repurposed candidates as single agents for the disease of interest.

#### 2.8 Causal-only pharmacome

Due to unclear relationships between repurposing candidates and target diseases in the pathways explored in HBP, we repeated the trial using a causal-only version of HBP. In this version, we eliminated noncausal relationships present in the HBP and retained only direct causal relationships, such as increases, decreases, or no effect between HBP entities (e.g. genes, proteins, and drugs). Entities that had no connecting relationships after the removal of noncausal relations were also omitted. Moreover, we removed hub nodes, primarily the disease nodes to generate causal-only pharmacome (COP). A separate version of COP, with disease nodes retained, was used for pathway confirmations.

Since the embedding approach outperformed ML in the HBP case, we applied it to COP as well. However, because the embedding-only approach was less efficient for COP, we introduced a third approach that combined embedding with ML. We extracted vector embeddings of the KG and used them to train and test ML models (Table 2). Utilizing the best-performing ML model, we generated final predictions for drug–drug relationships, which were then applied to drug repurposing for AD, schizophrenia, and bipolar disorder. We also tested our repurposing method on COVID-19, which lacks a node in COP, to evaluate the model's effectiveness as a rapid tool for drug repurposing during disease outbreaks.

# **3 Results**

### 3.1 KG enrichment

The KG enrichment was done over several steps as described in Section 2. The drug combination dataset contained 23 171 pairs from 882 unique drugs, which were used to enrich HBP. These combinations were further converted into edges to be added to the neo4j instance of HBP by converting the values of ZIP scores into synergistic, antagonistic, or additive effects. However, to avoid class imbalance and model overfitting due to the extremely low number of additive effect relations, additive edges were not added to HBP. The enriched HBP was then completely extracted as triples of source, relation, and target and was used to train, test, and validate the KG embedding models.

## 3.2 Classical machine learning

The cross-validation and synergy prediction results from the four selected ML models are detailed in Supplementary Results Section S1. Due to the lack of virtually validating studies and because this classical ML approach does not take relation type into consideration, we conducted a graphembedding-based approach, as explained in the next section.

## 3.3 KG embedding

Lacking a virtual validation by confirming literature, we turned our focus to novel combination prediction using graph embeddings. To do so we performed HBP embedding using different algorithms to model and predict novel drug-drug links in KG. When tested on the test set, RotatE model consistently outperformed other models in producing true predictions at the first rank as demonstrated by Fig. 2 and explained in Supplementary Results Section S2. More

**Table 2.** Embedding models trained on triplets from the causal-only pharmacome, and the machine learning models trained on the embedding vectors and used for synergism prediction.

Embedding models	Machine learning models	
TransE	Logistic regression	
TransR	Elastic net	
RotatE	Gradient boosting	
ComplEx	Random forest	
HolE	Support vector machine	



Figure 2. Percentage of true drug–drug relation predictions at different ranks for selected models. Optimum models were used to predict the drug–drug relations in the test set from the human brain pharmacome. Then, the predicted relations were compared to the actual relations to calculate the percentage of true predictions.

explanation of these results is elaborated in Supplementary Results Section S2. Based on these results, the RotatE model was selected as the model of choice for predicting relationships between drugs that do not have a direct link within HBP.

#### 3.4 Synergy prediction

Synergy predictions were generated leveraging the trained RotatE model. The prediction set was further processed as outlined in Supplementary Results Section S3. The five highest-scoring synergy combinations were subjected to a thorough literature review to validate the reliability of the model's predictions, as detailed in Supplementary Table S2. The findings revealed that most of these combinations either exhibited documented synergy or were being utilized in combination for the treatment of the specific diseases for which they were intended.

## 3.5 Drug repurposing

The assessment of the plausibility of repurposing candidates for the target disease was done by revising the literature for scientific data about their use in the selected disease and determining the common pathways between the synergistic drugs and the disease in HBP using the Python package NetworkX. This comprehensive approach enhances our understanding of potential therapeutic applications and facilitates informed decision-making regarding drug repurposing candidates. Detailed explanations of these candidates are elaborated in Supplementary Results Section S4.

Leveraging a broad and highly connected KG such as HBP, with both causal and noncausal relations like association and complexity, can lead to suboptimal model training and prediction, as well as to less explainable pathways between drugs and diseases. In addition, the presence of nodes with high degrees in KG will lead to inadequate training and inaccurate predictions as well. We believe these factors contributed to the discrepancy between our model's predictions for schizophrenia repurposing candidates and existing literature evidence (Supplementary Results Section S4a). To address this, we conducted a subsequent trial using COP, where we removed noncausal relationships and hub nodes. The hub nodes were removed before embedding, but we retained them in another copy of COP used for pathway confirmations.

## 3.6 Causal-only pharmacome

After the removal of noncausal relations and disease nodes that form hubs in the HBP, the same KG embedding models were used to model the causal-only version of the HBP. HolE was the best model to produce true predictions at the lowest rank (74.90% for all relations), as shown in Supplementary Fig. S2. Although it was also the best-performing model for drug-drug relations, the percentage of true predictions at the lowest rank was low (54.65%), indicating nearly random predictions (Fig. 3). Therefore, we changed the design of the experiment to incorporate KG embeddings, followed by ML model training and prediction using the extracted embedding vectors as input features. In this approach, we utilized the vector embeddings of COP without enrichment with drugdrug relations. The synergistic data was then used as labels for training and testing the ML models. In this run, HolE again proved to be the best-performing embedding model (84.58% for all relations), and random forest emerged as the best ML model with an ROC-AUC value of 0.796 (Figs 4 and 5). We have made COP, enriched with drug-drug relations (predicted and from databases), available at: https://doi. org/10.5281/zenodo.12806409. Consistent with prior methods, we validated the top-scoring synergistic combinations through a literature review (detailed in Table 3). This analysis found that most combinations (three out of five) either



Figure 3. Percentage of true drug-drug relation predictions at different ranks for selected models used to embed causal-only pharmacome. Optimum models were used to predict drug-drug relations in the test set. Then, the predicted relations were compared to the actual relations to calculate the percentage of true predictions.



Percentage of true predictions of different ranks for selected Models

Figure 4. Percentage of true all relations predictions at different ranks for selected models used to embed causal-only pharmacome before machine learning. Optimum models were used to predict the test set, and then the predicted relations were compared to the actual relations to calculate the percentage of true predictions. Embedding vectors of HoIE were extracted and used as input features for the training and testing of ML models.

demonstrated synergy between drug pairs or one of the drugs enhanced the effect of the other.

#### Subsequently, we selected safe drugs predicted to be synergistic partners with the previously chosen drugs for AD and schizophrenia, or bipolar disorder.

## 3.6.1 Alzheimer's candidates

Based on our research, miconazole and albendazole have emerged as promising candidates for repurposing to treat AD. They were predicted to act synergistically with three and two of the selected AD drugs, respectively. By tracing their

#### **Boxplot for different ML Models**



Figure 5. Benchmarking of machine learning models trained to classify between synergism and antagonism using the embedding vectors from causalonly pharmacome. Each boxplot shows the distribution of the ROC-AUC values over 10 repeats of the 10-fold nested cross-validation procedure.

Table 3. Validation of top scorer predictions, based on causal-only pharmacome, from published studies.

Drug A	Drug B	References	Remarks	Hit ratio
Tamsulosin	Ruxolitinib	(Wishart <i>et al</i> . 2018)	Tamsulosin decreases Ruxolitinib excretion rate, which could result in a higher serum level.	1 (one supporting study of one retrieved study)
Ruxolitinib	Zolpidem	(Wishart <i>et al.</i> 2018)	Zolpidem decreases the metabo- lism of Ruxolitinib increasing its effect.	1 (one supporting study of one retrieved study)
Ruxolitinib	Prednisolone	(Cortés <i>et al.</i> 2019)	The combination of Ruxolitinib with Prednisolone showed synergistic effects.	1 (four supporting studies of four retrieved studies)
Ruxolitinib	Cisapride	-	No study was found on their combination.	-
Deslanoside	Ruxolitinib	-	No study was found on their combination.	-

pathways to AD in the copy of COP, where we retained disease nodes, we found that they share pathways with their synergistic partners to AD.

Miconazole is a broad-spectrum antifungal with some antibacterial activity (Wishart et al. 2018). On the other hand, it offers a potential therapeutic approach for early intervention in AD by promoting myelination of the medial prefrontal cortex and ameliorating neuroinflammation-mediated AD progression in different mice models (Yeo et al. 2020, Wang et al. 2022). A prominent common pathway of miconazole with donepezil, rivastigmine, and galantamine to AD was found in COP, as shown in Fig. 6. On the other hand, Albendazole is primarily employed as an anthelmintic to treat helminth infections (Sungkar et al. 2019). However, research on H4 neuroglioma cells has shown that albendazole can reduce Tau levels, suggesting a beneficial effect on AD (Dickey et al. 2006). In

COP, it has a shared pathway with donepezil and rivastigmine, in which it intersects with them in decreasing the levels of phosphorylated microtubule-associated protein tau (MAPT), which is a hallmark of AD (Fig. 6). Comprehensive description of miconazole and albendazole pathways is elaborated in Supplementary Results Section S5.

Other candidates, such as disulfiram, auranofin, and finafloxacin, were also predicted as synergistic partners with donepezil, with donepezil and rivastigmine, and with donepezil and galantamine, respectively. Studies, in cell and animal models, showed their beneficial effects for the management of AD (Madeira et al. 2012, Roder and Thomson 2015, Reinhardt et al. 2018, Upīte et al. 2020, Guo et al. 2022, Jun and Fang 2021). However, no supporting pathway in COP could be detected for these drugs. Additionally, prochlorperazine was consistently predicted as a synergistic partner with



Figure 6. The shared pathways of miconazole, albendazole, donepezil, rivastigmine, and galantamine to Alzheimer's disease. Disease nodes were retained in a copy of the causal-only pharmacome used for pathway confirmations (ACHE: acetylcholinesterase; NOS2: nitric oxide synthase 2; APP: amyloid-beta precursor protein; MAPT: microtubule-associated protein tau; TUBA1A: tubulin alpha-1A protein; CDC42: cell division control protein 42 homolog; MAPK1 and MAPK3: mitogen-activated protein kinase 1 and 3).

the three selected AD drugs. However, it has anticholinergic properties that in higher doses might worsen AD-associated dementia (Obara *et al.* 2019). Moreover, the explored pathways between prochlorperazine and AD were controversial, with some suggesting it might have a beneficial effect for the management of AD while others suggest it might exacerbate the condition (Supplementary Fig. S3).

#### 3.6.2 Schizophrenia and bipolar disorder candidates

Upon detecting repurposing candidates for schizophrenia or bipolar disorder, we found no connecting pathways between any drug in the COP and these diseases, even for the drugs that are typically prescribed for these conditions. Therefore, we couldn't identify any repurposing candidates for schizophrenia or bipolar disorder based on COP.

## 3.6.3 COVID-19 candidates

We wanted to challenge our model further by checking its ability to predict synergistic drug combinations for diseases, which the KG was not built for, to test its ability to be used as a fast tool for repurposing drugs to new disease outbreaks. We selected a drug, baricitinib, used for COVID-19 and extracted the synergistic combinations from those predicted by our model. From these combinations, Prochlorperazine and disulfiram were selected as repurposing candidates based on their predicted synergy with baricitinib (they ranked as 8th and 10th out of 142 synergistic drug combinations, where the first seven ranks and 9th rank was for anticancer or unsafe drugs) and their beneficial effects in COVID-19 management. These effects include the inhibition of SARS-CoV-2 entry by targeting the spike protein and ACE2, which were confirmed computationally by molecular docking and experimentally in VeroE6 and HEK293T-hACE2 cell cultures (Chen *et al.* 2022, Liang *et al.* 2023).

# **4** Discussion

Pursuing a new drug candidate for a disease has been exhaustively overwhelming. Therefore, drug repurposing has recently gained significant interest. Here, we present this study of enriching existing KGs with drug synergy data to achieve a primary goal: repurposing predicted drug synergy partners as single agents for the disease of interest. Although the concept of KG-based drug repurposing has been previously studied (Al-Saleem et al. 2021, Ratajczak et al. 2022, Feng et al. 2023), we further enhanced KG with drug-drug relations, fostering more knowledge about synergistic drugs as a basis for the selection of repurposing candidates. This approach will have a significant influence on diseases with limited available therapeutic options, such as pandemics and neurodegenerative diseases. The alignment between the model's predictions and the real-world literature highlights the model's effectiveness in identifying clinically relevant and potentially impactful drug candidates. Although synergy predictions followed by safe drug selection, and pathway in pharmacome tracing, have reduced the number of selected drug candidates, they form a strong foundation for further research on these candidates.

As demonstrated from the results of drug repurposing candidates based on COP (Section 3.6) and HBP (Supplementary Results Section S4), some of our selected candidates such as albendazole and miconazole, mefloquine, ciprofloxacin, and moxifloxacin for AD, have a robust profile of clear pathways in COP with the disease of interest as well as experimental (in

cell and animal models) and clinical studies that support their willingness to be repurposed for that disease. On the other hand, another portion of selected drugs, including disulfiram, auranofin, finafloxacin, taribavirin, and ivermectin for AD, has strong literature evidence but unclear pathways in COP, which require further analysis and understanding of the relation encompassed in their pathway in COP. Finally, drugs that have no literature evidence or clear pathways were marked as the least suitable for repurposing including atovaquone for AD and pyrimethamine for Schizophrenia. In addition to these groups exists a controversial group in which literature supports their harmful effect on the selected disease; however, they appear many times in our predictions as a valuable agent for controlling that disease. This group includes mefloquine, chloroquine, and albendazole for schizophrenia. We highly recommend further clinical and experimental investigation of these drugs for that disease before the commencement of their repurposing procedures.

To challenge our model's applicability, we selected a drug for COVID-19 even knowing that there is no disease node for COVID-19 in COP. The results showed its ability to predict synergy and repurposing candidates, which a strong literature profile confirmed (Section 3.6.3). This vast ability underscores that the model does not rely on a single node or relation but on the overall interaction within the network. The potential of this approach to repurposing drugs for diseases that are out of the scope of the used pharmacome gives insights into comorbidity pathways that exist between these diseases. Specifically, we refer to the possible comorbidity between COVID-19 and neurodegenerative diseases (NDD). The ability of SynDRep to find repurposing candidates for COVID-19 may be attributed to these underlying comorbidity pathways. Therefore, this work paves the way for further research already being conducted for detecting such comorbidities (COMMUTE. Comorbidity Mechanisms Utilized in Healthcare2024).

We first took a broad approach that relied on graph topology metrics as well as the physicochemical properties of the drugs. ML models were then used to classify combinations into synergistic or antagonistic categories. However, this approach primarily neglected the "relationship type" factor within KG and relied solely on the data associated with the drug nodes. In contrast, KG embedding models take these relations into account when embedding all the nodes of the KG into vectors. This consideration improves the performance and predictions of KG embedding compared to ML modeling. Therefore, this approach emphasizes the beneficial effect of organizing data into KGs and the further extraction of this data using graph embedding techniques over the classical ML approaches. Moreover, analyzing biomedical data using network structures requires a thorough understanding of network topology. Therefore, we used the topological features along with the physicochemical features of the drugs for the training and prediction in the classical ML approach. However, these methods often demand high computational and space costs (Su et al. 2020) and result in lower performance than the graph embedding method as evidenced by the lack of literature studies for predicted top scorer partners (Supplementary Results Section S1). On the contrary, organizing the data into a graph that can describe the complex structure of data and enables the characterization of highorder geometric patterns for the networks, improves the performance of various data analysis tasks (Xu 2021). Graph

embedding techniques are able to convert sparse highdimensional graphs into continuous low-dimensional vectors that maximally preserve the graph structure properties (Cai et al. 2018). The generated highly informative and nonlinear embeddings can be subsequently used for different downstream analytic tasks such as node classification and link prediction. We applied these graph embedding techniques for the prediction of the link between pairs of drugs. RotatE excels at modeling symmetry, antisymmetry, inversion, and composition (Sun et al. 2019), which explains its superior performance in embedding HBP (Section 3.3). This was further evidenced when noncausal relations were removed to create COP, simplifying the structure. HolE and TransR performed even better than RotatE in COP due to its reduced complexity (Section 3.6). Unlike the ML approach, RotatE predictions were confirmed by published scientific studies as exploited in Supplementary Table S2. Consequently, utilizing data represented as graphs and incorporating their embeddings represents the future direction for pharmacome data mining.

In the context of drug repurposing, maintaining a clear chain of causality from drugs to disease targets is critical. Using a broad and highly connected KG such as HBP, which contains both cause-and-effect relations and less explicit relations such as association and complexity, can lead to suboptimal model training and prediction. Although selecting a cause-and-effect subgraph is optimal, the extensive relations pool in the pharmacome captures complex protein interactions that may not be strictly cause-and-effect. Additionally, the graph needs to be large enough for effective link prediction; otherwise, performance may be compromised, which was observed with the COP trial, hence pathway reviews and post-prediction literature validation were essential steps to compensate for the lack of a pure cause-and-effect subgraph by focusing on promising candidates. Another important consideration in the use of HBP is the presence of so-called "super-hubs", which are nodes with extremely high node degrees, whose presence in the KG dilutes information and hinders learning (Sardina et al. 2024). The topological imbalance in KGs has negative effects on learning using KG embedding models, where low-degree nodes embed at a much lower quality relative to high-degree nodes (Bonner et al. 2022). Moreover, high-degree nodes are mostly predicted as answers simply due to their higher degree, not their domain relevance (Bonner et al. 2022, Ratajczak et al. 2022). Based on these considerations, we performed a pruning of the HBP to remove noncausal relations between entities and super-hub nodes, which were mainly disease nodes. The disease nodes were removed before embedding, but we retained them in another copy of COP used for pathway confirmations. The results showed more promising repurposing candidates for AD. However, some candidates lacked pathways in COP. Additionally, we couldn't find any connections between schizophrenia or bipolar disorder and any drugs in COP, including those usually prescribed for these diseases. Their connecting relations might have been removed during the causal relation selection step (Section 3.6). This indicates the incompleteness of COP, which significantly impacts the repurposing approach we undertook in this study. Consequently, we recommend more manual curation of certain relation types in HBP to enhance and update their causality comprehension. This could help maintain the relation pool present in HBP, which is crucial for effective embedding and repurposing.

In contrast to our approach's advantages, it exhibits some challenges. Firstly, several sources of bias exist, including synergy score difference, cell type difference, and class imbalance. The data on drug-drug combinations were extracted from three databases: DrugcombPortal, DrugcombDB, and SYNERGxDB, which provide information on drug combinations along with their HSA, Bliss, Loewe, and ZIP scores. The subtle differences between synergy model scores may introduce bias in the evaluation of drug combinations. Therefore, we opted for the ZIP score as it is more reliable than the HSA and Bliss scores in detecting potency changes, and it combines the principles of Loewe additivity and Bliss independence. In addition, since some combinations may have a synergistic effect in one cell type and additive or antagonistic effects in another, we selected combinations that demonstrated pure antagonism or synergism across all cell types. Furthermore, most databases and studies focus on synergistic or antagonistic combinations, while scarce data about additive effect combinations are available. For instance, our work had only three additive combinations compared to tens of thousands of antagonistic or synergistic ones. Therefore, after many trials, we decided to omit these additive relations to avoid the class imbalance problem. Secondly, unlike the ML models which can classify drug combinations into synergism or antagonism, the embedding model's prediction for drug combinations is not, in all cases, a drug-drug relation but may predict any other relation available in the pharmacome. Consequently, some of the input data might get neither synergistic nor antagonistic prediction, resulting in the loss of some combinations. Thirdly, the controversy between some predictions and the published data about these drugs and diseases necessitates thorough investigations. Lastly, the diversity of drugs in the combination databases is limited; most are cancer-related and measure only cytotoxicity. Therefore, our approach must be extended to more specialized and highly curated KGs, such as cause-and-effect subgraphs, along with more balanced and versatile drug combination data.

This methodology would hold monumental potential as a robust tool for the pharmaceutical sector by broadening our search landscape and the production of more guided synergistic predictions. It leverages the certainty of manual curation from scientific literature by using a KG curated from various resources, including published scientific studies and databases. This KG-based approach is enhanced with a new concept that not only incorporates traditional drug repurposing methods-relying on semantic connections between drugs, diseases, and targets-but also utilizes reasoning derived from drug-drug synergy effects. The tool has been made available as an open-source Python package at https://github. com/SynDRep/SynDRep. We developed SynDRep with a focus on user-friendliness and provided a command-line interface to facilitate its use by scientists with biological or medical backgrounds who possess moderate knowledge of command-line prompts or python experience. The tool requires only four input files (KG.tsv, KG\_labels.tsv, Drugs. csv, Drug\_combinations.csv) and generates outputs that are saved with a single command. Clear documentation, including the required file types, is available in the README file of SynDRep. Therefore, this tool highlights the hugely beneficial effect of computational methods not only in reducing the chemical, energy, and resource waste required to conduct thousands of wet-lab investigations but also by helping in sustainability through re-using the same drugs for more diseases and the reduction of the capital required to set up new production plans. Therefore, tons of hours, labor, and costs have been spared, which can foster further projects and speed up the pace by which treatment plans can be exploited.

# Acknowledgements

K.S.S. gratefully acknowledge Andrea Zaliani for his insights into the manuscript.

## Author contributions

Karim S. Shalaby and Vinay Srinivas Bharadhwaj conceived and designed the study. Karim S. Shalaby analyzed the datasets, interpreted the results, and wrote the manuscript. Bruce Schultz collected data and proofread the manuscript. Alpha Tom Kodamullil and Martin Hofmann-Apitius acquired the funding. Martin Hofmann-Apitius, Sathvik Guru Rao, and Alpha Tom Kodamullil reviewed the manuscript. All authors have read and approved the final manuscript.

#### Supplementary data

Supplementary data are available at *Bioinformatics Advances* online.

#### **Conflict of interest**

None declared.

## Funding

This work was supported by the German Federal Ministry of Education and Research BMBF [01ZX1904C]; and the EU project eBRAIN-Health [GAP-101058516]. This work was developed in the Fraunhofer Cluster of Excellence "Cognitive Internet Technologies".

## **Data availability**

The data underlying this article are available as an opensource Python package at https://github.com/SynDRep/ SynDRep under the Apache 2.0 License.

# References

- Ali M, Berrendorf M, Hoyt CT et al. PyKEEN 1.0: a python library for training and evaluating knowledge graph embeddings. J Mach Learn Res 2021;22:1–6.
- Al-Saleem J, Granet R, Ramakrishnan S et al. Knowledge graph-based approaches to drug repurposing for COVID-19. J Chem Inf Model 2021;61:4058–67.
- Ashburn TT, Thor KB. Drug repositioning: identifying and developing new uses for existing drugs. *Nat Rev Drug Discov* 2004;3:673–83.
- Berenbaum MC. What is synergy. *Pharmacol Rev* 1989;41:93–141.
- Bliss CI. The toxicity of poisons applied jointly. Ann Appl Biol 1939; 26:585-615.
- Bonner S, Kirik U, Engkvist O *et al*. Implications of topological imbalance for representation learning on biomedical knowledge graphs. *Brief Bioinform* 2022; 23:bbac279.
- Bordes A, Usunier N, Garcia-Duran A et al. Translating Embeddings for Modeling Multi-relational Data. In: Advances in Neural Information Processing Systems. vol. 26. Nevada, USA: Curran Associates, Inc., 2013, 2787–95.

- Cai H, Zheng VW, Chang KC-C. A comprehensive survey of graph embedding: problems, techniques and applications. *IEEE Trans Knowl Data Eng* 2018;30:1616–37.
- Chen D, Liu X, Yang Y *et al.* Systematic synergy modeling: understanding drug synergy from a systems biology perspective. *BMC Syst Biol* 2015;9:56.
- Chen H, Cheng F, Li J. IDrug: integration of drug repositioning and drug-target prediction via cross-network embedding. *PLoS Comput Biol* 2020; 16:e1008040.
- Chen H-F, Hsueh P-R, Liu Y-Y *et al.* Disulfiram blocked cell entry of SARS-CoV-2 via inhibiting the interaction of spike protein and ACE2. *Am J Cancer Res* 2022;12:3333–46.
- Choudhury C, Arul Murugan N, Priyakumar UD. Structure-based drug repurposing: traditional and advanced AI/ML-aided methods. *Drug Discov Today* 2022;27:1847–61.
- Cokol M, Chua HN, Tasan M et al. Systematic exploration of synergistic drug pairs. Mol Syst Biol 2011;7:544.
- COMMUTE. Comorbidity Mechanisms Utilized in Healthcare. https:// www.commute-project.eu/en/about. (11 July 2024, date last accessed).
- Cortés AA, Diaz RA, Hernández-Campo P et al. Ruxolitinib in combination with prednisone and nilotinib exhibit synergistic effects in human cells lines and primary cells from myeloproliferative neoplasms. *Haematologica* 2019;104:937–46.
- Day D, Siu LL. Approaches to modernize the combination drug development paradigm. *Genome Med* 2016;8:115.
- Dickey CA, Ash P, Klosak N *et al.* Pharmacologic reductions of total tau levels; implications for the role of microtubule dynamics in regulating tau expression. *Mol Neurodegener* 2006;1:6.
- Dubourg J, Fouqueray P, Quinslot D *et al.* Long-term safety and efficacy of imeglimin as monotherapy or in combination with existing antidiabetic agents in Japanese patients with type 2 diabetes (TIMES 2): a 52-week, open-label, multicentre phase 3 trial. *Diabetes Obes Metab* 2022;24:609–19.
- Feala JD, Cortes J, Duxbury PM *et al.* Systems approaches and algorithms for discovery of combinatorial therapies. *Wiley Interdiscip Rev Syst Biol Med* 2010;2:181–93.
- Feng X, Ma Z, Xin Y et al. DREG: a large-scale biological knowledge graph for drug repurposing. 2023. https://doi.org/10.21203/RS.3. RS-2734221/V1
- Ferreira D, Adega F, Chaves R et al. The importance of cancer cell lines as in vitro models in cancer methylome analysis and anticancer drugs testing. In: López-Camarillo C, Aréchaga-Ocampo E (eds), Oncogenomics and Cancer Proteomics—Novel Approaches in Biomarkers Discovery and Therapeutic Targets in Cancer. Rijeka: IntechOpen, 2013, 139–66.
- Gao Z, Ding P, Xu R. KG-Predict: a knowledge graph computational framework for drug repurposing. J Biomed Inform 2022; 132:104133.
- Goswami CP, Cheng L, Alexander PS *et al*. A new drug combinatory effect prediction algorithm on the cancer cell based on gene expression and dose-response curve. *CPT Pharmacometrics Syst Pharmacol* 2015; 4:e9.
- Guo W, Chen S, Li C et al. Application of disulfiram and its metabolites in treatment of inflammatory disorders. Front Pharmacol 2022;12: 795078. https://doi.org/10.3389/fphar.2021.795078
- Hagberg AA, Schult DA, Swart PJ. Exploring network structure, dynamics, and function using NetworkX. In: Varoquaux G, Vaught T, Millman J (eds.), Proceedings of the 7th Python in Science Conference (SciPy2008). Pasadena, CA, 2008. 11–15.
- He L, Kulesskiy E, Saarela J *et al.* Methods for high-throughput drug combination screening and synergy scoring. *Methods Mol Biol Clifton NJ* 2018;1711:351–98.
- Hua Y, Dai X, Xu Y et al. Drug repositioning: progress and challenges in drug discovery for various diseases. Eur J Med Chem 2022; 234:114239.
- Hurle MR, Yang L, Xie Q et al. Computational drug repositioning: from data to therapeutics. *Clin Pharmacol Ther* 2013;93:335–41.

- Jarada TN, Rokne JG, Alhajj R. A review of computational drug repositioning: strategies, approaches, opportunities, challenges, and directions. J Cheminformatics 2020;12:46.
- Jun C, Fang B. Current progress of fluoroquinolones-increased risk of aortic aneurysm and dissection. BMC Cardiovasc Disord 2021; 21:470.
- Kim Y, Zheng S, Tang J et al. Anticancer drug synergy prediction in understudied tissues using transfer learning. J Am Med Inform Assoc 2021;28:42–51.
- Knorz AL, Quante A. Alzheimer's disease: efficacy of mono- and combination therapy. A systematic review. J Geriatr Psychiatry Neurol 2022;35:475–86.
- Kuenzi BM, Park J, Fong SH *et al.* Predicting drug response and synergy using a deep learning model of human cancer cells. *Cancer Cell* 2020;38:672–84.e6.
- Lage-Rupprecht V, Schultz B, Dick J et al. A hybrid approach unveils drug repurposing candidates targeting an Alzheimer pathophysiology mechanism. Patterns 2022;3:100433.
- Landrum G. RDKit: open-source cheminformatics. 2023. https://www. rdkit.org (27 September 2024, date last accessed).
- Li H, Zou L, Kowah JAH *et al.* Predicting drug synergy and discovering new drug combinations based on a graph autoencoder and convolutional neural network. *Interdiscip Sci Comput Life Sci* 2023; 15:316–30.
- Li P, Huang C, Fu Y *et al.* Large-scale exploration and analysis of drug combinations. *Bioinformatics* 2015;**31**:2007–16.
- Liang T, Xiao S, Wu Z *et al.* Phenothiazines inhibit SARS-CoV-2 entry through targeting spike protein. *Viruses* 2023;15:1666.
- Lin Y, Liu Z, Sun M et al. Learning entity and relation embeddings for knowledge graph completion. Proc AAAI Conf Artif Intell 2015;29: 2181–7. https://doi.org/10.1609/aaai.v29i1.9491
- Liu H, Zhang W, Zou B et al. DrugCombDB: a comprehensive database of drug combinations toward the discovery of combinatorial therapy. Nucleic Acids Res 2020; 48:D871–D881.
- Liu Q, Xie L. TranSynergy: mechanism-driven interpretable deep neural network for the synergistic prediction and pathway deconvolution of drug combinations. *PLOS Comput Biol* 2021; 17:e1008653.
- Loewe S. The problem of synergism and antagonism of combined drugs. Arzneimittelforschung 1953;3:285–90.
- Lombardi N, Crescioli G, Simonetti M et al. Adherence to triple-freedrug combination therapies among patients with cardiovascular disease. Am J Cardiol 2020;125:1429–35.
- Luo H, Li M, Yang M *et al*. Biomedical data and computational models for drug repositioning: a comprehensive review. *Brief Bioinform* 2021;22:1604–19.
- Madeira J, Renschler C, Gill N *et al.* P2-373: the anti-rheumatic gold drug auranofin could be beneficial in Alzheimer's disease by inhibiting astrocyte-mediated neuroinflammation. *Alzheimers Dement* 2012;8:P394.
- Nickel M, Rosasco L, Poggio T. Holographic Embeddings of Knowledge Graphs. 2015. https://doi.org/10.48550/arXiv. 1510.04935
- Obara K, Horiguchi S, Shimada T *et al.* Characterization of binding of antipsychotics to muscarinic receptors using mouse cerebral cortex. *J Pharmacol Sci* 2019;**140**:197–200.
- Pan X, Lin X, Cao D et al. Deep learning for drug repurposing: methods, databases, and applications. Wiley Interdiscip Rev Comput Mol Sci 2022; 12:e1597.
- Pang K, Wan Y-W, Choi WT *et al.* Combinatorial therapy discovery using mixed integer linear programming. *Bioinformatics* 2014; 30:1456–63.
- Preuer K, Lewis RPI, Hochreiter S et al. DeepSynergy: predicting anticancer drug synergy with Deep Learning. *Bioinformatics* 2018; 34:1538–1546.
- Pushpakom S, Iorio F, Eyers PA *et al.* Drug repurposing: progress, challenges and recommendations. *Nat Rev Drug Discov* 2018; 18:41–58.
- Ratajczak F, Joblin M, Ringsquandl M et al. Task-driven knowledge graph filtering improves prioritizing drugs for repurposing. BMC Bioinformatics 2022;23:84.

- Reinhardt S, Stoye N, Luderer M *et al*. Identification of disulfiram as a secretase-modulating compound with beneficial effects on Alzheimer's disease hallmarks. *Sci Rep* 2018;8:1329.
- Roder C, Thomson MJ. Auranofin: repurposing an old drug for a golden new age. *Drugs R D* 2015;15:13–20.
- Sardina J, Costabello L, Guéret C. Veni, Vidi, Vici: solving the myriad of challenges before knowledge graph learning. In: 2024 IEEE 18th International Conference on Semantic Computing (ICSC), Laguna Hills, CA, USA. 2024, 197–203.
- Scannell JW, Blanckley A, Boldon H et al. Diagnosing the decline in pharmaceutical R&D efficiency. Nat Rev Drug Discov 2012; 11:191–200.
- Seo H, Tkachuk D, Ho C et al. SYNERGxDB: an integrative pharmacogenomic portal to identify synergistic drug combinations for precision oncology. Nucleic Acids Res 2020; 48:W494–W501.
- Su C, Tong J, Zhu Y et al. Network embedding in biomedical data science. Brief Bioinform 2020;21:182–97.
- Sun X, Bao J, You Z et al. Modeling of signaling crosstalk-mediated drug resistance and its implications on drug combination. Oncotarget 2016;7:63995–4006.
- Sun Z, Deng Z-H, Nie J-Y et al. RotatE: knowledge graph embedding by relational rotation in complex space. In: 2019 International Conference on Learning Representations (ICLR 2019), New Orleans, Louisiana, USA. 2019.
- Sun Z, Huang S, Jiang P et al. DTF: deep tensor factorization for predicting anticancer drug synergy. *Bioinformatics* 2020;36:4483–9.
- Sungkar S, Putri KQ, Taufik MIS et al. The effectiveness of triple dose albendazole in treating soil transmitted helminths infection. J Parasitol Res 2019;2019:6438497.
- Trouillon T, Welbl J, Riedel S et al. Complex embeddings for simple link prediction. In: Balcan MF, Weinberger KQ (eds), Proceedings

of The 33rd International Conference on Machine Learning. New York, New York, USA: PMLR; 2016, 2071–80.

- Upīte J, Kadish I, van Groen T *et al.* Subchronic administration of auranofin reduced amyloid-β plaque pathology in a transgenic APPNL-G-F/NL-G-F mouse model. *Brain Res* 2020;**1746**:147022.
- Wang Z, Zhang Y, Feng W et al. Miconazole promotes cooperative ability of a mouse model of Alzheimer disease. Int J Neuropsychopharmacol 2022;25:951–67.
- Wildenhain J, Spitzer M, Dolma S et al. Prediction of synergism from chemical-genetic interactions by machine learning. Cell Syst 2015; 1:383–95.
- Wishart DS, Feunang YD, Guo AC et al. DrugBank 5.0: a major update to the DrugBank database for 2018. Nucleic Acids Res 2018; 46:D1074–D1082.
- Xu M. Understanding graph embedding methods and their applications. SIAM Rev 2021;63:825–53.
- Yadav B, Wennerberg K, Aittokallio T et al. Searching for drug synergy in complex dose-response landscapes using an interaction potency model. Comput Struct Biotechnol J 2015;13:504–13.
- Yeo IJ, Yun J, Son DJ et al. Antifungal drug miconazole ameliorated memory deficits in a mouse model of LPS-induced memory loss through targeting iNOS. Cell Death Dis 2020;11:623.
- Zagidullin B, Aldahdooh J, Zheng S et al. DrugComb: an integrative cancer drug combination data portal. Nucleic Acids Res 2019; 47:W43–W51.
- Zhou Z, Edil BH, Li M. Combination therapies for cancer: challenges and opportunities. BMC Med 2023;21:171. https://doi.org/10. 1186/S12916-023-02852-4
- Zong N, Wen A, Moon S *et al.* Computational drug repurposing based on electronic health records: a scoping review. *NPJ Digit Med* 2022; 5:77–8.

© The Author(s) 2025. Published by Oxford University Press.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited. Bioinformatics Advances, 2025, 00, 1–13 https://doi.org/10.1093/bioadv/vbaf092