

RESEARCH ARTICLE

# DMAPLM: A multimodal pretrained framework for computational drug repositioning

Hailin Chen<sup>1</sup>\*, Zhongling Li

School of Information and Software Engineering, East China Jiaotong University, Nanchang, China

\* [chenhailin@ecjtu.edu.cn](mailto:chenhailin@ecjtu.edu.cn)



## Abstract

Drug repositioning offers an efficient route to discover new therapeutic indications for existing drugs. However, current computational drug repositioning models often face challenges related to data scarcity, heterogeneity, and therefore limited generalizability. To address these limitations, this study introduces DMAPLM, a multimodal pretrained framework for predicting drug-disease associations for further drug repositioning screening. DMAPLM leverages a lightweight dual-encoder architecture, utilizing ChemBERTa-2 for molecular encoding of drug SMILES strings and BioBERT for semantic encoding of multi-field disease texts. The framework explicitly aligns drug and disease representations through contrastive learning and employs attention-weighted pooling to emphasize informative molecular substructures. A Random Forest classifier is finally used for association prediction based on the enhanced multimodal features. We compile a new and comprehensive benchmark dataset for performance evaluation. Extensive experiments demonstrate that DMAPLM significantly outperforms six state-of-the-art baseline models, achieving an AUROC of 0.8919 and AUPR of 0.9116 under five-fold cross-validation, representing an improvement of up to 9%. Furthermore, DMAPLM exhibits robust performance in challenging cold-start scenarios, highlighting its practical utility for identifying novel drug-disease relationships. Case studies along with molecular docking analysis confirm the interpretability and biological meaningfulness of our predictions. Our study provides a powerful and interpretable approach for computational drug repositioning.

## OPEN ACCESS

**Citation:** Chen H, Li Z (2026) DMAPLM: A multimodal pretrained framework for computational drug repositioning. *PLoS Comput Biol* 22(4): e1014192. <https://doi.org/10.1371/journal.pcbi.1014192>

**Editor:** Fuhai Li, Washington University in Saint Louis, UNITED STATES OF AMERICA

**Received:** November 21, 2025

**Accepted:** April 1, 2026

**Published:** April 22, 2026

**Copyright:** © 2026 Chen, Li. 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 use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data availability statement:** The data and source codes used in our study are available at <https://github.com/LiNiuNiuYa/DMAPLM>.

**Funding:** This work was supported by National Natural Science Foundation of China (62562031 to H.C) and Jiangxi Provincial Natural Science Foundation, China (20242BAB25083 to H.C). The funders had no role in study design, data

## Author summary

Identifying potential associations between drugs and diseases is essential for accelerating drug discovery and repurposing, yet current computational approaches often struggle when data are sparse or when new drugs or diseases lack sufficient annotations. To address these challenges, we propose a cross-modal learning framework that jointly models molecular structures and disease text descriptions. By explicitly aligning these two information sources through contrastive learning

collection and analysis, decision to publish, or preparation of the manuscript.

**Competing interests:** The authors have declared that no competing interests exist.

and attention mechanisms, our method captures complementary patterns that conventional models tend to overlook. As a result, the framework provides robust predictions even in cold-start scenarios and offers biologically interpretable insights that can support real-world translational applications. Our study demonstrates the promise of integrating chemical and biomedical language representations to better understand drug–disease relationships and guide future drug development.

## Introduction

The pharmaceutical industry is continually confronted with significant challenges in the process of drug development, including escalating research and development costs, prolonged timelines, and high failure rates [1]. In response, drug repositioning [2]—also known as drug repurposing—has garnered considerable attention as an efficient strategy to identify novel therapeutic indications for existing drugs. By leveraging approved drugs or those in clinical development for new usage, this approach offers a markedly reduced risk profile relative to *de novo* drug discovery, largely owing to the established safety, pharmacokinetic, and pharmacodynamic profiles of the existing drugs [3]. Consequently, drug repositioning holds immense potential to expedite the availability of treatments for unmet medical needs, particularly for rare diseases and emergent health threats, yielding notable clinical and economic benefits.

Although traditional experimental drug repurposing, initiated by hypotheses on a drug's mechanism of action or *in vitro* screening of cell lines and animal models against specific disease targets, has yielded successes, it is frequently hampered by low efficiency, limited scope, and a high degree of unpredictability in identifying a broad spectrum of clinically viable repurposed drug candidates [4,5]. Moreover, existing experimental approaches are often biased by prior knowledge, potentially overlooking unconventional yet effective drug-disease associations.

With the accumulation of various types of biomedical data, such as chemical structures, gene expression profiles and adverse event profiles, researchers have developed a series of computational methods to predict novel drug-disease associations for further drug repositioning screening. These methods can be roughly divided into three categories, namely network propagation-based [6–8], traditional machine learning-based [9–11] and deep learning-based [12–14]. More recently, with the emergence of large language models (LLMs) such as ChatGPT, biomedical AI has entered a new era. Pretrained models, such as BioBERT [15], ChemBERTa [16], and ESM-2 [17], have shown strong feature representation capabilities by learning from large-scale corpora, and successfully applied to tasks such as molecular property prediction [18] and protein design [19]. Some recent studies have applied LLMs to drug repurposing [20–23].

Despite these impressive progress, current computational drug repurposing methods face the following challenges and limitations. Firstly, current gold-standard datasets, such as the well-known Gottlieb dataset [24], for training and evaluating drug repositioning prediction models are small and outdated, and are therefore missing many associations. Secondly, data scarcity and heterogeneity remain a persistent

problem. While many drug and disease databases exist, they are often incomplete, biased, and lack standardization, hindering the robustness and generalizability of prediction models. Thirdly, the interpretability and explainability of complex computational models, particularly deep learning architectures, are often lacking. This “black box” nature can limit trust and hinder the translation of predictions into actionable insights for experimental validation. Finally, the integration of multi-modal data (e.g., genomics, proteomics, transcriptomics, electronic health records) to create comprehensive prediction models is still in its initial phases, and effectively analyzing such disparate data sources presents considerable technical challenges.

To address these problems, we compile a new benchmark dataset by searching molecular atlas and pharma-information from DrugMAP 2.0 (released in 2024) [25], and propose DMAPLM, a lightweight dual-encoder framework based on pretrained language models (PLMs), for drug-disease association prediction in this study. Unlike existing methods, DMAPLM directly integrates embeddings of drug molecules and disease texts learned from pretrained language models through contrastive learning and attention mechanisms, without requiring graph construction or task-specific fine-tuning, making it especially suitable for data-sparse and cold-start scenarios.

We comprehensively evaluate the performance of DMAPLM on the collected benchmark dataset. Experimental results show that under five-fold cross-validation and cold-start settings, DMAPLM improves AUROC and AUPR by up to 9% over state-of-the-art methods. Case studies along with molecular docking analyses further confirm that predicted associations align with literature-supported biological mechanisms, highlighting DMAPLM’s potential for real-world drug repurposing applications. The main contributions of this study are as follows:

1. **Lightweight cross-modal architecture:** We propose a dual-encoder framework that leverages pretrained language models—specifically ChemBERTa and BioBERT—to extract molecular and textual representations without constructing complex graph structures.
2. **Explicit cross-modal alignment:** We apply contrastive learning to explicitly align molecular structure and disease text semantic spaces, while capturing complementary features through attention mechanisms to enhance cross-modal discriminability.
3. **Robustness and interpretability:** The framework achieves stable performance under data-sparse and cold-start scenarios, with biologically interpretable predictions suitable for real-world clinical translational applications.

This paper is organized as follows. Section II presents the datasets and our proposed method, section III reports experimental results, section IV discusses the findings, implications, and limitations, and section V concludes the study.

## Materials and methods

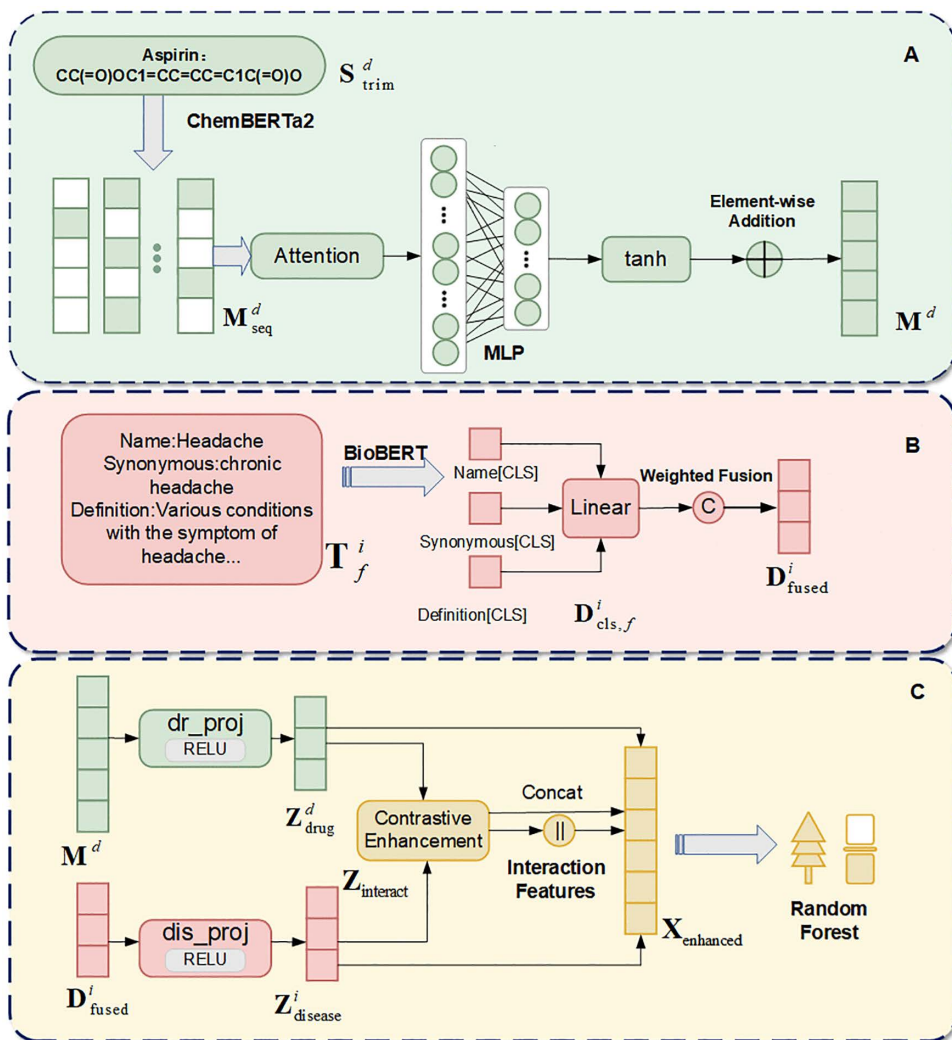
### A. Datasets

To obtain the latest information for performance evaluation, we search the DrugMAP 2.0 (<https://idrblab.org/drugmap>) [25], a manually curated pharmaceutical database, for experimentally confirmed drug-disease associations. We retain only associations with “Approved” status to focus on clinically validated drugs. Meanwhile, we only keep drugs with valid SMILES (Simplified Molecular-Input Line-Entry System) representations. For diseases, we extract three of their textual fields from the DrugMAP 2.0 database: Disease Name, Synonymous, and Definition.

After data preprocessing, we finally construct a benchmark dataset comprising 1,455 diseases, 2,622 drugs, and 5,993 drug-disease associations, with an association density of 0.0016 (0.16%).

### B. Method architecture

The model architecture of DMAPLM for drug-disease association prediction is illustrated in [Fig 1](#), which mainly consists of three components: (A) attention-based molecular encoding using ChemBERTa-2 [26], (B) multi-field disease text encoding using BioBERT [15], and (C) contrastive learning enhancement for feature extraction with Random Forest-based classification.



**Fig 1. Workflow of DMAPLM for predicting drug-disease associations.** Module A: ChemBERTa2 encodes SMILES sequences into 384-dimensional embeddings via attention pooling. Module B: BioBERT encodes multi-field disease texts with weighted fusion. Module C: Embeddings enhanced by contrastive learning for Random Forest prediction.

<https://doi.org/10.1371/journal.pcbi.1014192.g001>

**Pre-trained language module for molecular and biomedical text encoding.** For drug molecules, we employ ChemBERTa-2 (DeepChem/ChemBERTa-77M-MTR) as the molecular encoder to transform drug SMILES strings into contextualized embeddings. ChemBERTa-2 is a RoBERTa-like pre-trained language model specifically designed for molecular representation learning based on SMILES notation. The model was trained on 77 million drug molecules from PubChem through self-supervised learning, enabling it to capture the intrinsic chemical semantics and structural patterns of molecular compounds. For each drug molecule, the SMILES string is first tokenized using the ChemBERTa-2 tokenizer, with sequences truncated to a maximum length of 512 tokens to ensure computational efficiency. The molecular encoding process is formulated as:

$$S_{trim}^d = \text{Tokenizer}(\text{SMILES}^d), \quad (1)$$

$$\mathbf{M}_{\text{seq}}^d = \text{ChemBERTa2}(\mathbf{S}_{\text{trim}}^d), \quad (2)$$

where  $\mathbf{S}_{\text{trim}}^d \in \mathbb{R}^{L_d}$  represents the tokenized and truncated SMILES sequence for drug  $d$  with length  $L_d \leq 512$ , and  $\mathbf{M}_{\text{seq}}^d \in \mathbb{R}^{L_d \times 384}$  denotes the contextualized molecular embeddings extracted from the final hidden layer of ChemBERTa-2. Each token is encoded as a 384-dimensional vector, capturing the sequential dependencies and local chemical environments within the molecular structure. These token-level representations preserve fine-grained structural information that will be subsequently aggregated to obtain drug-level representations for downstream prediction tasks.

For disease entities, we apply BioBERT (dmis-lab/biobert-base-cased-v1.1) as the disease text encoder to transform disease textual descriptions into semantic embeddings. BioBERT is a domain-specific pre-trained language model adapted from BERT (Bidirectional Encoder Representations from Transformers) for biomedical text mining. The model was pre-trained on large-scale biomedical corpora including PubMed abstracts and PMC full-text articles, enabling it to effectively capture complex biomedical terminologies and semantic relationships in disease descriptions. To construct comprehensive disease representations, we extract three complementary textual fields for each disease: Disease Name, Synonymous terms, and detailed Definitions. These multi-field descriptions are separately encoded and fused through weighted aggregation. For each textual field  $f$ , the encoding process is formulated as:

$$\mathbf{T}_f^i = \text{Tokenizer}(\text{Text}_f^i), \quad (3)$$

$$\mathbf{D}_{\text{cls},f}^i = \text{BioBERT}(\mathbf{T}_f^i)_{[\text{CLS}]}, \quad (4)$$

where  $\mathbf{T}_f^i$  represents the tokenized text sequence for disease  $i$  and field  $f$  truncated to a maximum length of 512 tokens, and  $\mathbf{D}_{\text{cls},f}^i \in \mathbb{R}^{768}$  denotes the disease representation extracted from the [CLS] token [27,28] of BioBERT's final hidden layer. The [CLS] token serves as an aggregate representation of the entire input sequence, capturing the global semantic meaning of the disease description. To leverage information from multiple textual fields, we perform weighted fusion across different fields as follows:

$$\mathbf{D}_{\text{fused}}^i = \frac{\sum_{f \in F} w_f \cdot \mathbf{D}_{\text{cls},f}^i}{\sum_{f \in F} w_f}, \quad (5)$$

where  $F = \{\text{Name, Synonymous, Definitions}\}$  represents the set of textual fields, and  $w_f$  denotes the weight assigned to the field  $f$  to reflect the relative importance of different information sources. The fused representation  $\mathbf{D}_{\text{fused}}^i \in \mathbb{R}^{768}$  is further projected to a 128-dimensional space through a linear transformation to obtain the final disease embedding  $\mathbf{D}^i \in \mathbb{R}^{128}$ , providing a compact and informative representation for downstream drug-disease association prediction.

**Attention-based pooling for molecular embeddings.** After obtaining token-level molecular representations  $\mathbf{M}_{\text{seq}}^d \in \mathbb{R}^{L_d \times 384}$ , we use an attention-weighted pooling mechanism [29] to aggregate the sequence into a fixed-size drug-level representation. The attention pooling process is calculated as:

$$e_j = \text{MLP}(\mathbf{m}_j) = W_2 \cdot \tanh(W_1 \cdot \mathbf{m}_j + b_1) + b_2, \quad (6)$$

$$\alpha_j = \frac{\exp(e_j - \max_k e_k)}{\sum_{k=1}^{L_d} \exp(e_k - \max_k e_k)}, \quad (7)$$

$$\mathbf{M}^d = \sum_{j=1}^{L_d} \alpha_j \cdot \mathbf{m}_j, \quad (8)$$

where  $\mathbf{m}_j \in \mathbb{R}^{384}$  represents the embedding of the  $j$ -th token from  $\mathbf{M}_{\text{seq}}^d$ ,  $e_j$  denotes the attention score received by a two-layer feedforward network with tanh activation, and  $\alpha_j$  is the normalized attention weight. The subtraction of  $\max_k e_k$  in the softmax operation ensures numerical stability. The final pooled representation  $\mathbf{M}^d \in \mathbb{R}^{384}$  aggregates information from all tokens, with each token weighted by its learned attention coefficient. This attention mechanism allows the model to automatically identify and emphasize the most informative molecular substructures for drug-disease association prediction.

**Contrastive learning for feature enhancement.** To enhance the prediction power of drug and disease representations, we incorporate a contrastive learning framework [30] that projects embeddings into a shared latent space. Two projection networks  $f_{\text{drug}}$  and  $f_{\text{disease}}$  are used to transform the original embeddings:

$$\mathbf{Z}_{\text{drug}}^d = f_{\text{drug}}(\mathbf{M}^d), \quad \mathbf{Z}_{\text{disease}}^i = f_{\text{disease}}(\mathbf{D}^i), \quad (9)$$

where  $\mathbf{Z}_{\text{drug}}^d$  and  $\mathbf{Z}_{\text{disease}}^i \in \mathbb{R}^{128}$  represent the projected embeddings. Both projection networks consist of two linear layers with ReLU activation and layer normalization. An interaction function  $f_{\text{interact}}$  operates on the concatenated projected features to capture cross-modal patterns:

$$\mathbf{Z}_{\text{interact}} = f_{\text{interact}}([\mathbf{Z}_{\text{drug}}^d; \mathbf{Z}_{\text{disease}}^i]), \quad (10)$$

The final enhanced representation combines the original and projected features:

$$\mathbf{X}_{\text{enhanced}} = [\mathbf{M}^d; \mathbf{D}^i; \mathbf{Z}_{\text{drug}}^d; \mathbf{Z}_{\text{disease}}^i; \mathbf{Z}_{\text{interact}}], \quad (11)$$

The projection networks are optimized using an InfoNCE-based contrastive loss. For a batch of drug-disease pairs, we compute the similarity matrix:

$$S_{ij} = \frac{\mathbf{Z}_{\text{drug}}^i \cdot (\mathbf{Z}_{\text{disease}}^j)^T}{\tau}, \quad (12)$$

where  $\tau$  is the temperature parameter. The contrastive loss with high similarity is used for separating positive pairs from negative ones:

$$\mathcal{L}_{\text{contrast}} = - \sum_{(d,i): y_{di}=1} \log \frac{\exp(S_{di})}{\sum_j \exp(S_{dj})}, \quad (13)$$

This strategy enables learning of discriminative representations that capture both entity-specific properties and relational patterns between drugs and diseases.

**Random forest for drug-disease association prediction.** We use Random Forest to predict potential drug-disease associations based on the PLM-derived embeddings. Random Forest is an ensemble learning algorithm that constructs multiple decision trees through bootstrap sampling and aggregates their predictions via

majority voting. For each drug-disease pair with enhanced feature representation  $\mathbf{X}_{\text{enhanced}}$ , the prediction result is computed as:

$$\hat{y} = \frac{1}{T} \sum_{t=1}^T h_t(\mathbf{X}_{\text{enhanced}}), \quad (14)$$

where  $T$  is the number of trees and  $h_t$  denotes the output of the  $t$ -th tree. Each tree is trained on a bootstrap sample with random feature subsets at each split, providing robustness against overfitting.

**Optimal threshold selection via Youden index.** To convert probability scores into binary predictions, we apply the Youden index to determine the optimal threshold. The Youden index maximizes the sum of sensitivity and specificity as:

$$J(\theta) = \text{TPR}(\theta) - \text{FPR}(\theta), \quad (15)$$

$$\theta^* = \text{argmax}_{\theta} J(\theta), \quad (16)$$

where TPR and FPR denote true positive rate and false positive rate, respectively. Binary predictions are obtained via  $\hat{y}_{\text{binary}} = \mathbf{1}(\hat{y} \geq \theta^*)$ .

## Results

### A. Experimental setting

In this section, known drug-disease associations are labeled as positive samples, while pairs without known association information serve as negative samples. Given the limited number of positive samples, we apply a balanced sampling strategy by randomly selecting an equal number of unknown associations as negative samples. This approach mitigates potential bias caused by data imbalance and ensures a fair evaluation of model performance. We use five-fold cross-validation (5-CV) for performance assessment. The dataset is evenly split into five subsets, with four used for training and one for testing in each fold, ensuring that every sample is used for both training and testing exactly once. This procedure is repeated five times. We calculate AUC, AUPR, F1-score and Accuracy for performance evaluation.

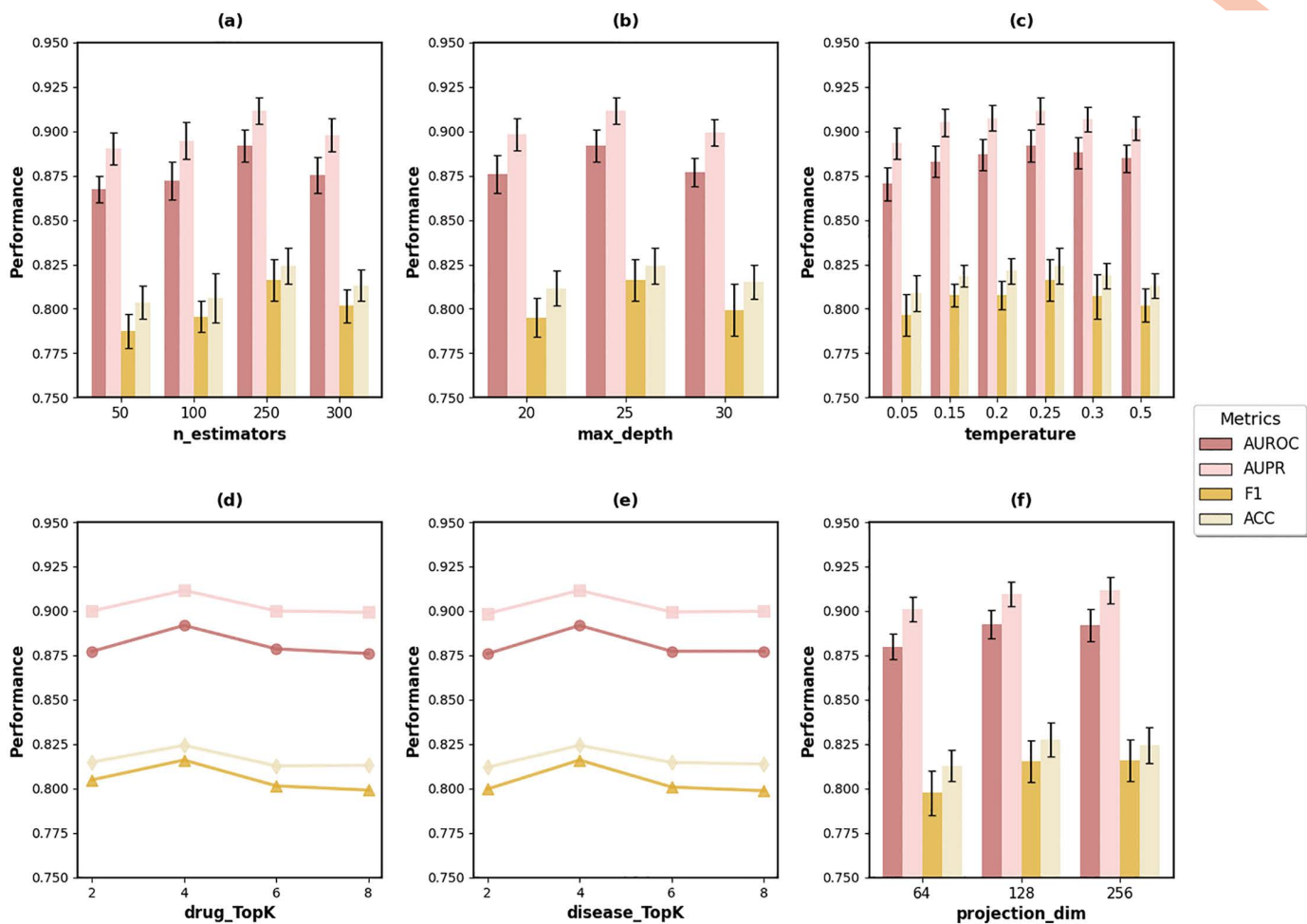
### B. Hyperparameter configuration

We apply grid search to evaluate different hyperparameter settings on prediction performance based on 5-fold cross-validation. We conduct sensitivity analysis on key hyperparameters including `n_estimators`, `max_depth`, `temperature`, `drug_TopK`, `disease_TopK`, and `projection_dim`. The results are shown in [Fig 2](#).

Based on the experimental results, the key hyperparameters are set as follows: The Random Forest uses 250 estimators with a maximum depth of 25. The minimum samples required to split an internal node is 5, and the minimum samples required at a leaf node is 2. Feature selection uses the square root of total features at each split. Class weights are balanced to address data imbalance. For similarity network construction, disease TopK is set to 4 and drug TopK is set to 4 to select the most relevant neighbors. The contrastive learning module uses a temperature parameter of 0.25 and projection dimension of 256. Training employs a learning rate of 0.0005 with early stopping to prevent overfitting.

### C. Ablation tests

We conduct ablation studies based on 5-fold cross-validation ([Table 1](#)) to assess the performance contribution of different components in DMAPLM. Removing pre-trained language model embeddings causes dramatic performance decline, demonstrating the critical importance of PLM representations. Replacing PLM embeddings with one-hot encoding shows



**Fig 2. Hyperparameter sensitivity analysis of DMAPLM.** Optimal configuration:  $n\_estimators=250$ ,  $max\_depth=25$ ,  $temperature=0.25$ ,  $TopK=4$ , and  $projection\_dim=256$ .

<https://doi.org/10.1371/journal.pcbi.1014192.g002>

moderate performance but significant degradation from PLM features. Excluding contrastive learning leads to performance reduction, confirming its importance in learning discriminative representations. Deleting attention-weighted pooling causes performance reduction, highlighting the value of adaptive feature aggregation.

#### D. Model comparison

We further compare the prediction performance of DMAPLM with six state-of-the-art (SOTA) models using the benchmark dataset based on 5-fold cross-validation. The baseline models for performance comparison include four deep learning SOTA models and two traditional machine learning models:

- LAGCN [31]: A model that incorporates layer attention mechanisms in graph convolutional layers to enhance drug-disease association prediction.
- HGTDR [32]: A heterogeneous graph transformer model that constructs knowledge graphs and utilizes transformer networks with fully connected layers for drug repurposing prediction.

**Table 1. Results of ablation tests based on 5-fold cross-validation.**

Method	AUC	AUPR	F1-score	Accuracy
<b>DMAPLM</b>	<b>0.8919</b>	<b>0.9116</b>	<b>0.8161</b>	<b>0.8244</b>
DMAPLM w/o PLM	0.5950	0.6012	0.5555	0.6220
DMAPLM w/o CL	0.8462	0.8711	0.7682	0.7877
DMAPLM w/o Att	0.8875	0.9069	0.8079	0.8144
DMAPLM w/ One-hot	0.7642	0.7644	0.7076	0.7405

DMAPLM w/o PLM is the version that removes pre-trained language model embeddings.

DMAPLM w/o CL is the version that deletes contrastive learning enhancement.

DMAPLM w/o Att is the version that removes attention-weighted pooling.

DMAPLM w/ One-hot is the version that replaces PLM embeddings with one-hot encoding.

<https://doi.org/10.1371/journal.pcbi.1014192.t001>

- DTI-LM [20]: A pretrained language model framework that integrates molecular graph and protein sequence features using graph attention networks for drug-target interaction prediction.
- HNF-DDA [21]: A transformer-style heterogeneous network embedding model that employs subgraph contrastive learning and all-pairs message passing for drug-disease association prediction.
- DrugLAMP [22]: A multi-modal framework combining pretrained language models with pocket-guided co-attention and paired multi-modal attention mechanisms for drug-target interaction prediction.
- Node2Vec [33]: A network embedding method that learns continuous feature representations for nodes through biased random walks.

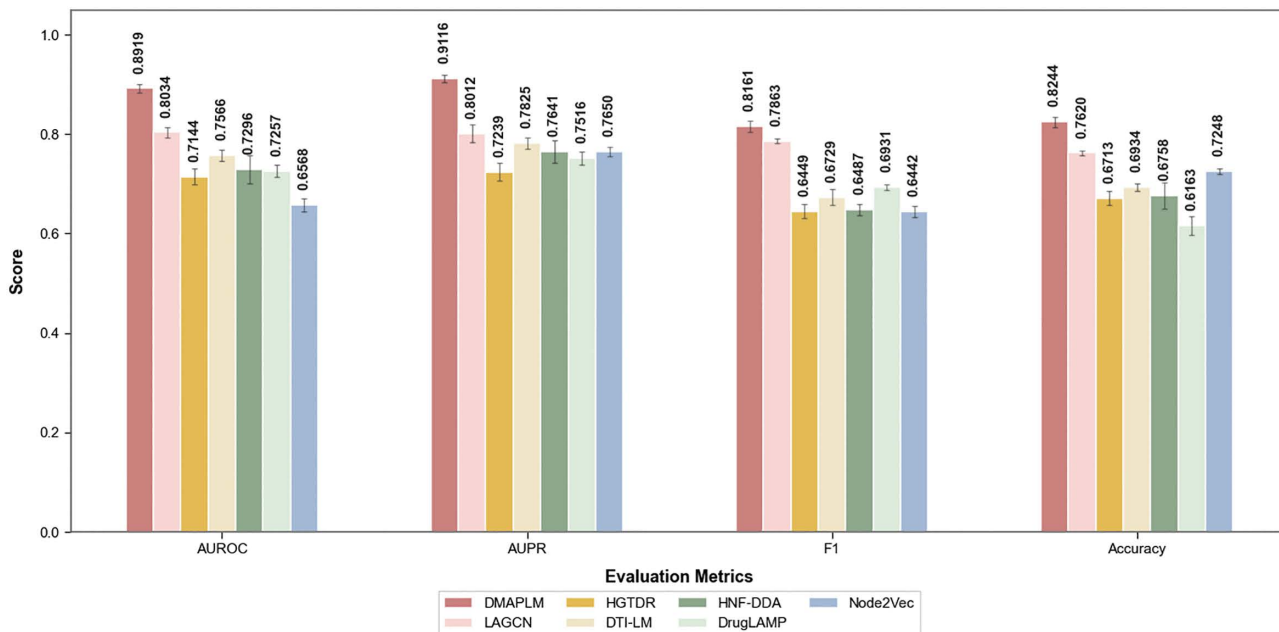
Among these SOTA models, all focus on the broad field of drug repositioning area, with HGTDR and HNF-DDA specifically for drug-disease association prediction, while DTI-LM and DrugLAMP were designed for drug-target interaction prediction. To ensure a fair comparison, all tests are conducted under identical conditions, including five-fold cross-validation, consistent random seed initialization, and the same data partitioning strategy.

As shown in Fig 3, DMAPLM outperforms all other competing models across all evaluation metrics. It achieves an AUROC of 0.8919 and AUPR of 0.9116, exceeding the second-best performing model LAGCN by 8.85% in AUROC (0.8919 vs 0.8034) and 9.22% in AUPR (0.9116 vs 0.8194). These results demonstrate that DMAPLM effectively integrates pretrained language model embeddings with contrastive learning and provides significant improvement in predicting drug-disease associations.

### E. Cold-start prediction

We conduct comprehensive cold-start prediction experiments using the standard C1, C2, and C3 protocols following established practices in computational drug discovery [34,35]. These protocols evaluate three critical scenarios that simulate real-world drug discovery challenges: C1 simulates new drug scenarios by randomly selecting 20% of drugs and their interactions as the test set; C2 evaluates new disease scenarios using 20% of diseases; C3 represents the most challenging double cold-start scenario by simultaneously selecting 20% of both drugs and diseases for testing.

As shown in Fig 4, DMAPLM achieves the best performance in all cold-start scenarios: AUROC and AUPR are 0.8150 and 0.8338 for C1 (new drugs), 0.7805 and 0.8098 for C2 (new diseases), and 0.7456 and 0.7355 for C3 (double cold-start). Even in the most challenging C3 scenario, DMAPLM maintains stable prediction performance, demonstrating its generalization advantage in practical drug repurposing applications.



**Fig 3. Performance comparison of DMAPLM with baseline models.** DMAPLM significantly outperforms all other methods, with most notable improvements in AUROC and AUPR over the second-best model LAGCN.

<https://doi.org/10.1371/journal.pcbi.1014192.g003>

## F. Top-K evaluation

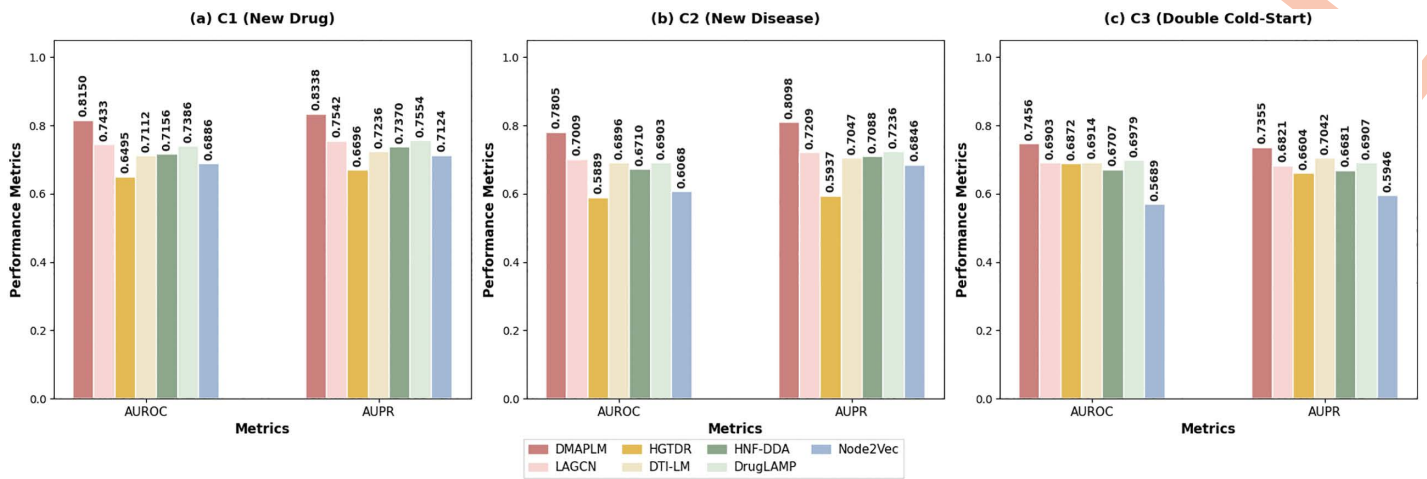
To evaluate the prioritization ability of DMAPLM in practical drug screening, we use Top-K metrics to simulate clinical decision-making scenarios, including Precision@K (P@K), Recall@K (R@K), Mean Reciprocal Rank (MRR), and Normalized Discounted Cumulative Gain (NDCG@10).

As shown in Table 2, DMAPLM achieves superior performance on all Top-K metrics. Specifically, in the most critical Top-1 prediction, DMAPLM's P@1 reaches 0.8139, a 3.73% improvement compared to the second-best model LAGCN (0.7846), indicating that our model can accurately identify the most promising candidate drugs. Meanwhile, the MRR value of 0.8990 (1.59% higher than LAGCN's 0.8849) confirms the model's advantage in early ranking, which is crucial for reducing downstream experimental validation costs. In addition, the high P@3 and R@3 scores (0.5155 and 0.9087) both surpass suboptimal methods, indicating that DMAPLM maintains high precision and recall when expanding the candidate list, making it suitable for clinical scenarios.

## G. Robustness analysis

We test the robustness of DMAPLM by systematically introducing Gaussian noise to the input features to assess the model's resilience to data quality issues commonly encountered in biomedical fields. Gaussian noise with varying intensities (0%, 10%, 20%, 30%, 40%, 50%) is systematically injected into both the PLM embeddings and the final concatenated features during training and testing phases. The noise level represents the standard deviation of the Gaussian noise relative to the feature standard deviation.

As shown in Fig 5, DMAPLM demonstrates progressive performance degradation as noise levels increase. Under moderate level (20%), the model maintains strong performance with only 4.69% AUROC degradation. Even under severe noise (50%), the model retains reasonable prediction capability with 19.21% AUROC and 18.83% AUPR degradation.



**Fig 4. Cold-start prediction performance evaluation under three different scenarios.** C1 tests new drugs, C2 tests new diseases, C3 tests both simultaneously. DMAPLM achieves best performance across all scenarios, maintaining stability in the most challenging C3.

<https://doi.org/10.1371/journal.pcbi.1014192.g004>

**Table 2. Top-k metrics comparison.**

Model	P@1	R@1	P@3	R@3	MRR	NDCG@10
DMAPLM	0.8139	0.5770	0.5155	0.9087	0.8990	0.6626
LAGCN	0.7846	0.5566	0.5108	0.9068	0.8849	0.6542

<https://doi.org/10.1371/journal.pcbi.1014192.t002>

The results indicate that DMAPLM's learned representations are robust to input perturbations and can maintain clinically relevant performance when faced with noisy biomedical data.

## H. feature quality assessment

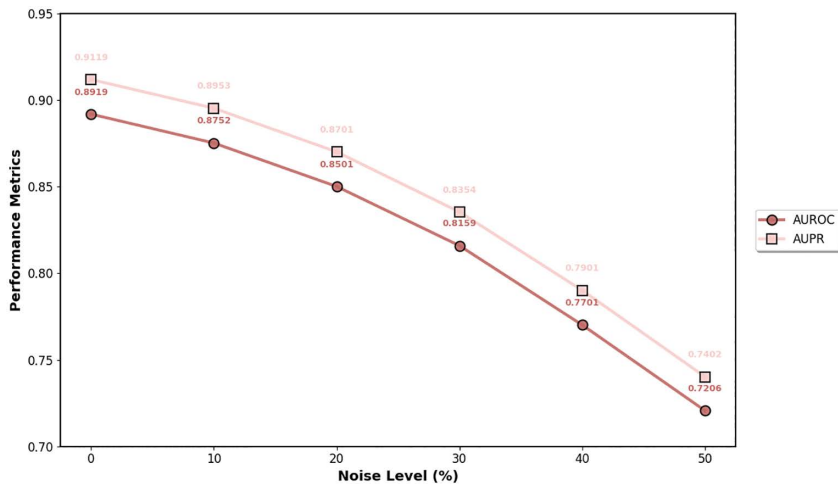
To assess the effect of contrastive learning, we analyze similarity distributions in the embedding space. From one representative training fold, we sample 1,000 positive (known drug-disease) and 1,000 negative pairs, comparing three settings: Baseline, PLM, and PLM+CL.

Cosine similarity was used to measure feature separability [30], following the principle that positive pairs should be close while negative pairs should be distant. A single-fold analysis was used to clearly show the distributional differences without averaging effects.

As shown in Fig 6, contrastive learning significantly improves the discriminability of the embedding space. Visualization shows that Baseline embeddings exhibit strong overlap between positive and negative samples; PLM introduces partial separation, while PLM+CL achieves clear boundaries. Quantitative analysis confirms this improvement: positive pair similarity for PLM+CL reaches 0.679 (PLM: 0.346), negative pair similarity decreases to 0.146 (PLM: 0.182), and Cohen's d attains 4.56, far exceeding PLM (2.18) and Baseline (0.24) [36]. This enhanced separability results in a cosine similarity-based classification AUC of 0.915, representing 21.5% and 62.2% improvements over PLM and Baseline, respectively, confirming that contrastive learning produces highly discriminative features that could support accurate association prediction.

## I. PLM-based embedding clustering analysis

We obtain MeSH Tree Numbers for diseases via the NLM REST API [https://id.nlm.nih.gov/mesh/] [37]. The first three characters of each Tree Number define the Primary\_Class: C01 (bacterial and fungal infections), C04 (neoplasms), C10



**Fig 5. Robustness analysis under varying noise levels.** Even though, results show progressive performance degradation, DMAPLM still maintains reasonable prediction ability under severe noise.

<https://doi.org/10.1371/journal.pcbi.1014192.g005>

(nervous system diseases), C14 (cardiovascular diseases), C17 (skin and connective tissue diseases), and C23 (pathological conditions and signs). Of the 873 diseases with retrieved Tree Numbers, the top 6 categories contain 600 diseases (70%): C01 (160), C04 (122), C10 (103), C23 (76), C14 (72), and C17 (67).

We perform KMeans clustering ( $k=6$ ) on disease embeddings. Table 3 shows PLM embeddings achieve  $NMI=0.342$ ,  $ARI=0.367$ , and  $Silhouette=0.174$ . TF-IDF obtains  $NMI=0.027$ ,  $ARI=-0.004$ , and random baseline achieves  $NMI=0.011$ ,  $ARI=0.001$ . PLM embeddings significantly outperform baselines in capturing disease taxonomy.

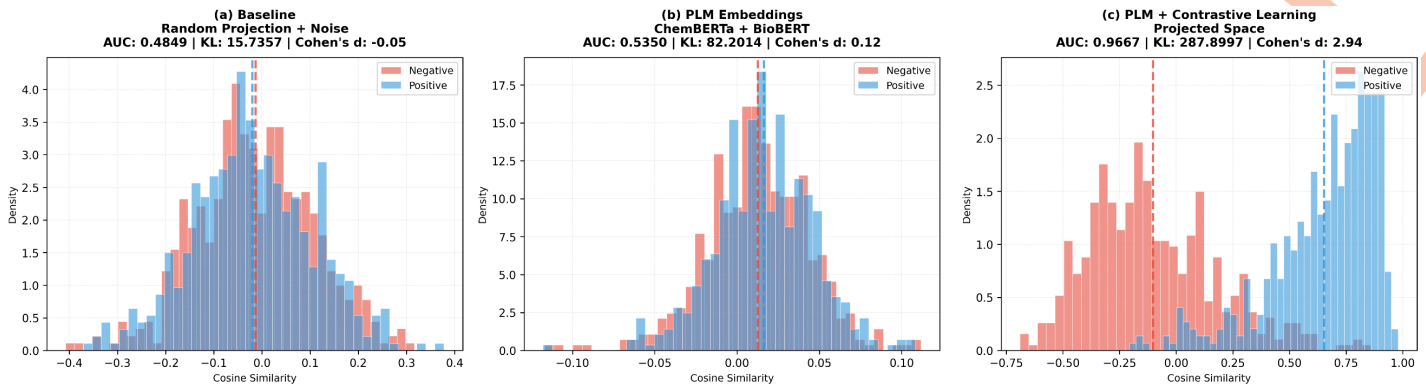
Fig 7 shows t-SNE projections of embeddings. PLM embeddings display clear cluster structures with distinct category separation. TF-IDF and random embeddings show no clear patterns. The visualization confirms that PLM effectively encodes disease features consistent with MeSH taxonomy.

## J. Attention weight visualization

We apply attention weight visualization to demonstrate the interpretability of our model. We select three drugs (Rapamycin, Leurubicin and Adenosine) as examples. Fig 8 shows that our model automatically identifies pharmacologically relevant substructures. When predicting Rapamycin's association with prostate cancer, attention concentrates on the macrolide core and epoxide group—precisely the FKBP12-binding pharmacophore confirmed by crystal structures. Meanwhile, for Leurubicin and Adenosine against bacterial infection, attention patterns clearly differentiate their distinct mechanisms (aromatic ring systems vs. phosphate chains), indicating that DMAPLM can learn mechanistic features of drugs.

## K. Case studies

To test DMAPLM's ability to predict novel drug-disease associations under realistic scenarios, we conduct case studies using the highest-scoring predictions obtained from the C1 and C2 cold-start validation protocols described in Section E. Specifically, the top-ranked associations involved Rapamycin-Prostate cancer, Leurubicin-Acute myelogenous leukaemia, and Adenosine-Bacterial infection. In each case, the original known associations are removed from the binary interaction matrix to simulate unknown conditions, and DMAPLM is then applied to predict new links. The resulting top predictions are validated through systematic literature mining on PubMed (<https://pubmed.ncbi.nlm.nih.gov/>) using PMID-indexed evidence.



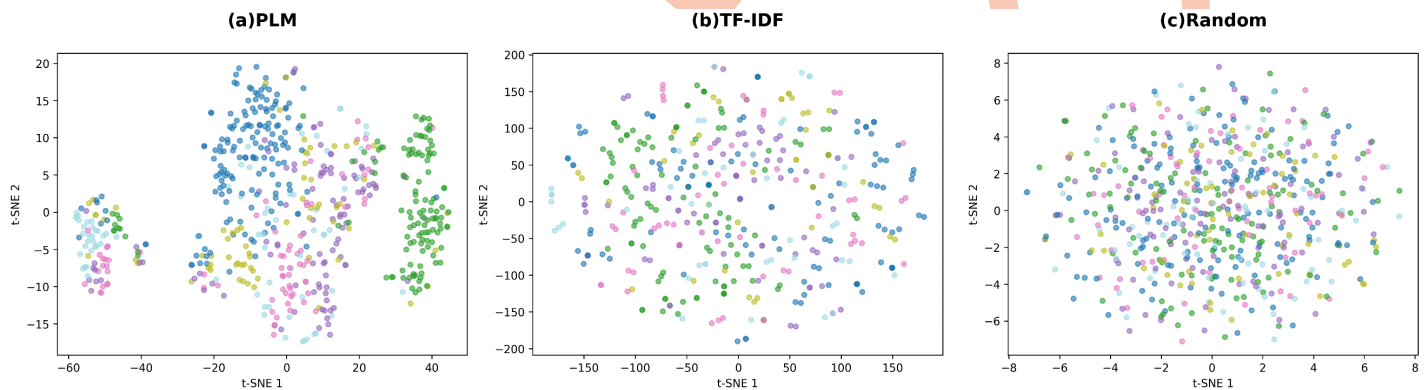
**Fig 6. Cosine similarity distributions for feature quality assessment.** Baseline shows severe overlap; PLM achieves partial separation; PLM + CL reaches clear boundaries.

<https://doi.org/10.1371/journal.pcbi.1014192.g006>

**Table 3. Disease Embedding Clustering Metrics.**

Method	NMI	ARI	Silhouette
PLM	0.342	0.367	0.174
TF-IDF	0.027	-0.004	-0.183
Random	0.011	0.001	0.006

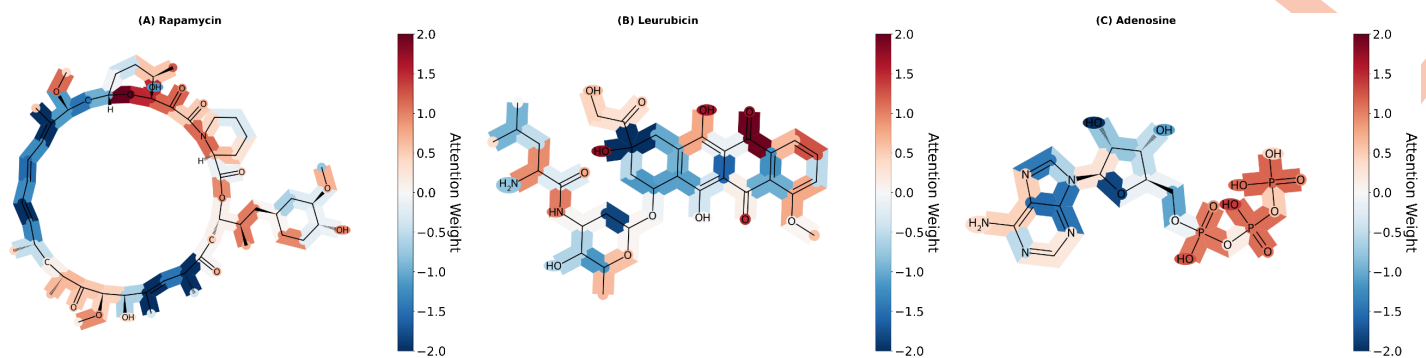
<https://doi.org/10.1371/journal.pcbi.1014192.t003>



**Fig 7. t-SNE Visualization of Disease Embeddings.** PLM displays clear clustering with distinct category boundaries; TF-IDF and Random show no patterns.

<https://doi.org/10.1371/journal.pcbi.1014192.g007>

As shown in Table 4, all the top five predictions for the three drugs have been confirmed by literature. For example, Shorning et al. [38] demonstrated rapamycin's therapeutic potential in prostate cancer by highlighting the critical role of the PI3K–AKT–mTOR signaling axis in this malignancy. Breistøl et al. [39] confirmed the antitumor efficacy of N-L-leucyl-doxorubicin (Leurubicin) in human melanoma xenograft models, demonstrating superior tumor growth inhibition compared to doxorubicin. Importantly, the study verified that Leurubicin functions as a prodrug and achieves higher intratumoral concentrations of active doxorubicin, thereby enhancing therapeutic efficacy. Leurubicin is an N-L-leucyl



**Fig 8. Attention weight visualization of three drugs.**

<https://doi.org/10.1371/journal.pcbi.1014192.g008>

prodrug of the anthracycline doxorubicin, exhibiting antineoplastic activity (PubChem CID: 68897). Xiang et al. [41] confirmed adenosine's protective role against bacterial infection through NLRP3 inflammasome activation.

We further test three drugs with mixed therapeutic and adverse profiles and validate their top predicted disease associations. As shown in Table 5, most of these predictions show strong literature support. For example, Hong et al. [42] demonstrated both atorvastatin and rosuvastatin's therapeutic efficacy in coronary heart disease, showing that high-intensity statin treatment with either atorvastatin 40 mg or rosuvastatin 20 mg daily significantly reduced the 3-year composite risk of death, myocardial infarction, stroke, or coronary revascularization in a randomized trial of 4,400 patients with established coronary artery disease. Haidar et al. [43] confirmed pramlintide's therapeutic effectiveness against type 1 diabetes, achieving increased time in glycemic range from 74% to 84% ( $P=0.0014$ ) and improved daytime control in a randomized crossover trial using a novel dual-hormone artificial pancreas system. Wald et al. [44] confirmed dexamethasone's therapeutic effectiveness against bacterial meningitis in children, achieving improved mortality and neurologic outcomes in pneumococcal meningitis as an adjuvant therapy. We note that these computational predictions primarily reflect therapeutic associations, though rare instances may represent adverse events (e.g., hydrocortisone-associated meningitis [45]). Experimental validation is required to determine the clinical nature of each prediction.

To validate DMAPLM's predictions for rapamycin's anticancer activity, we conduct molecular docking studies. Rapamycin (also known as sirolimus), a macrocyclic lactone isolated from Easter Island soil, was initially discovered as an antifungal agent but later recognized for its potent immunosuppressive and antiproliferative properties [50,51].

Rapamycin's mechanism of action is well-established: it binds the intracellular receptor FKBP12, and this complex then binds the FRB (FKBP12-Rapamycin Binding) domain of mTOR, allosterically inhibiting mTOR kinase activity [52]. In 1996, Choi et al. solved the crystal structure of the FKBP12-rapamycin-FRB ternary complex (PDB: 1FAP) [53], followed by Liang et al.'s refined 2.2 Å resolution structure (PDB: 4FAP) in 1999 [54].

The FDA approved rapamycin (1999) and its analogs temsirolimus (2007) and everolimus (2009) as mTOR inhibitors for organ transplantation and multiple cancers [55,56]. Moreover, dysregulation of the mTOR signaling pathway has been widely observed across diverse human cancers [57,58], particularly in prostate cancer, lung cancer, and plasma cell myeloma—the cancer types predicted by DMAPLM for rapamycin treatment.

We use the high-resolution FKBP12-rapamycin-FRB ternary complex structure (PDB: 4FAP) as the docking template. The protein structure is prepared using AutoDockTools 1.5.6, including addition of polar hydrogens, Gasteiger charge calculation, and conversion to PDBQT format. Rapamycin's 3D structure is obtained from PubChem (CID: 5284616) and processed with Open Babel for energy minimization.

**Table 4. The top five diseases predicted by DMAPLM for three drugs.**

Drug	Disease	Rank	Evidence
Rapamycin	Prostate cancer	1	PMID:32630372
Rapamycin	Advanced cancer	2	PMID:22149876
Rapamycin	Lung cancer	3	PMID:37057884
Rapamycin	Plasma cell myeloma	4	PMID:39875173
Rapamycin	Intracranial meningioma	5	PMID:30082628
Leurubicin	Bacterial infection	1	PMID:37308748
Leurubicin	Acute myelogenous leukaemia	2	PMID:40293351
Leurubicin	Leukemia	3	PMID:20572037
Leurubicin	Chronic obstructive pulmonary disease	4	PMID:6254538
Leurubicin	Melanoma Bacterial infection	5	PMID:10533461
Adenosine	Bacterial infection	1	PMID:23717478
Adenosine	MALT lymphoma	2	PMID:36539557
Adenosine	Plasma cell myeloma	3	PMID:32409420
Adenosine	Hepatosplenic T-cell lymphoma	4	PMID:2981904
Adenosine	Myeloproliferative neoplasm	5	PMID:33748423

<https://doi.org/10.1371/journal.pcbi.1014192.t004>

**Table 5. The top five drugs predicted by DMAPLM for three diseases.**

Disease	Drug	Rank	Evidence
Coronary Heart Disease	Atorvastatin	1	PMID:36877807
Coronary Heart Disease	Rosuvastatin	2	PMID:36877807
Coronary Heart Disease	Enalapril	3	PMID:41070303
Coronary Heart Disease	Esmolol	4	PMID:26847114
Coronary Heart Disease	Bumetanide	5	NA
Type-1 Diabetes	Pramlintide	1	PMID:31974099
Type-1 Diabetes	Dapagliflozin	2	PMID:32311204
Type-1 Diabetes	Exenatide	3	PMID:32573927
Type-1 Diabetes	Simvastatin	4	PMID:17519305
Type-1 Diabetes	Atorvastatin	5	PMID:37951886
Meningitis	Dexamethasone	1	PMID:40874187
Meningitis	Betamethasone	2	PMID:27404370
Meningitis	Ciprofloxacin XR	3	PMID:29944651
Meningitis	Fusidic Acid	4	PMID:35810737
Meningitis	Hydrocortisone	5	PMID:26358724

<https://doi.org/10.1371/journal.pcbi.1014192.t005>

Molecular docking is performed using AutoDock Vina v1.2.7. The search space includes the entire FRB domain and surrounding regions, with exhaustiveness set to 8 for thorough conformational sampling. The software generates 9 binding conformations ranked by binding affinity. Results are visualized and analyzed using PyMOL 3.1.

AutoDock Vina generates 9 rapamycin-mTOR binding conformations (Table 6). Binding affinities ranged from -6.807 to -9.921 kcal/mol. The best conformation (Mode 1) shows a binding affinity of -9.921 kcal/mol, indicating strong binding capability. The top five conformations all exceed -7.5 kcal/mol with RMSD values <2.5 Å, demonstrating consistent binding modes. This binding energy falls within the moderate-to-strong binding range (-9.0 to -12.0 kcal/mol) [59], suggesting favorable bioactivity potential.

**Table 6. Molecular docking results of Rapamycin with mTOR (PDB: 4FAP).**

Model	Binding Affinity (kcal/mol)	RMSD l.b. (Å)	RMSD u.b. (Å)
1	-9.92	0.000	0.000
2	-8.47	2.37	6.95
3	-7.99	1.90	3.75
4	-7.50	1.80	3.58
5	-7.04	8.59	12.60

Note: RMSD l.b. and RMSD u.b. represent lower-bound and upper-bound root-mean-square deviations relative to the best conformation.

<https://doi.org/10.1371/journal.pcbi.1014192.t006>

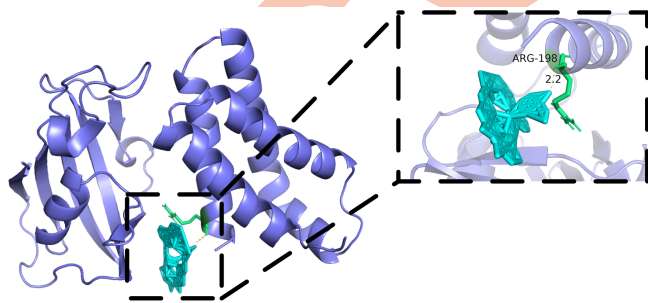
PyMOL visualization (Fig 9) reveals that rapamycin successfully docked into the FRB domain binding pocket. Analysis identifies a key hydrogen bond between rapamycin and ARG-198 (Chain B) with a distance of 2.2 Å (Fig 9), within the ideal range for strong hydrogen bonding (2.0-3.0 Å) [60]. Additionally, rapamycin forms extensive hydrophobic interactions with the  $\alpha 1$  and  $\alpha 4$  helices of the FRB domain, consistent with the interaction patterns observed in Liang et al.'s crystal structure [54].

This case study provides structural evidence supporting DMAPLM's prediction of rapamycin's anticancer potential. The docking results confirm strong rapamycin-mTOR interactions, consistent with its known inhibitory mechanism. Together with literature evidence on mTOR dysregulation in prostate, lung, and myeloma cancers, these findings demonstrate biological plausibility for the predicted associations.

## Discussion

In this study, we introduce DMAPLM, a multimodal pretrained framework designed to address key challenges in computational drug repositioning, including data sparsity, modality heterogeneity, and limited generalizability. By combining ChemBERTa-2-based molecular encoding with BioBERT-based disease text representation, DMAPLM enables a unified cross-modal embedding space optimized through contrastive alignment. This design allows the model to capture complementary semantic and structural information that are often overlooked by graph-based or matrix-completion approaches.

Across five-fold and cold-start evaluations, DMAPLM consistently outperformed state-of-the-art baselines, highlighting its robustness in realistic scenarios where novel drugs or diseases lack sufficient annotation. Notably, the model maintained strong performance even in the double cold-start setting, underscoring its ability to predict unobserved interaction patterns. This property is of practical relevance for repurposing tasks, where candidate compounds or emerging disease phenotypes frequently lie outside the training dataset.



**Fig 9. Docking pose of Rapamycin bound to the FRB domain of mTOR (PDB: 4FAP).** Rapamycin docks into FRB domain, forming hydrogen bond with ARG-198 with binding affinity of -9.921 kcal/mol, validating biological plausibility of predictions.

<https://doi.org/10.1371/journal.pcbi.1014192.g009>

The case studies further demonstrate DMAPLM's capacity to generate biologically plausible hypotheses. High-confidence associations were subsequently supported through literature mining and molecular docking. The structural consistency between predicted binding modes and known biochemical mechanisms underscore the value of integrating pretrained language models for representing both molecular substructures and disease semantics.

A key advantage of DMAPLM lies in its interpretability. Ablation studies reveal that DMAPLM's performance is primarily driven by pretrained semantic representations capturing both chemical and biomedical knowledge. Additionally, contrastive learning produces a discriminative embedding space where positive and negative drug-disease pairs are clearly separated, enabling interpretable similarity-based classification. Attention weight visualization of drugs further demonstrates the interpretability of our model.

Despite these strengths, several limitations remain. First, the model's performance is inherently constrained by the availability and coverage of curated drug-disease data; missing annotations or inconsistent disease terminology may introduce bias. Second, although contrastive pretraining improves representation alignment, the resulting embeddings still reflect static molecular and textual information and do not fully capture context-dependent biological processes. Third, the validation strategy relies on *in silico* docking and literature evidence; experimental assays are required before clinical translatability can be established.

Future work may focus on incorporating dynamic biological data such as transcriptomic perturbation profiles or patient-specific multi-omics features, which may enhance interpretability and mechanistic insight. Integrating causal reasoning or pathway-aware modeling could further improve the identification of therapeutically actionable relationships. In addition, extending DMAPLM to support few-shot or zero-shot prediction paradigms may broaden its utility for understudied diseases.

## Conclusions

DMAPLM provides an effective and interpretable multimodal framework for computational drug repositioning. By leveraging pretrained language models and contrastive representation alignment, the method achieves substantial improvements in predictive accuracy and generalization across multiple evaluation settings. Case studies supported by molecular docking and literature evidence demonstrate that DMAPLM can generate biologically meaningful hypotheses with translational potential. Overall, our study offers a scalable and robust computational tool for accelerating drug repurposing, and its integration with experimental validation pipelines holds promise for facilitating more efficient therapeutic discovery.

## Author contributions

**Conceptualization:** Hailin Chen.

**Data curation:** Hailin Chen, Zhongling Li.

**Formal analysis:** Hailin Chen.

**Funding acquisition:** Hailin Chen.

**Investigation:** Hailin Chen, Zhongling Li.

**Methodology:** Hailin Chen, Zhongling Li.

**Software:** Zhongling Li.

**Supervision:** Hailin Chen.

**Validation:** Hailin Chen.

**Writing – original draft:** Hailin Chen, Zhongling Li.

**Writing – review & editing:** Hailin Chen, Zhongling Li.

## References

1. Khanna I. Drug discovery in pharmaceutical industry: productivity challenges and trends. *Drug Discov Today*. 2012;17(19–20):1088–102. <https://doi.org/10.1016/j.drudis.2012.05.007> PMID: 22627006
2. Ashburn TT, Thor KB. Drug repositioning: identifying and developing new uses for existing drugs. *Nat Rev Drug Discov*. 2004;3(8):673–83. <https://doi.org/10.1038/nrd1468> PMID: 15286734
3. Novac N. Challenges and opportunities of drug repositioning. *Trends Pharmacol Sci*. 2013;34(5):267–72. <https://doi.org/10.1016/j.tips.2013.03.004> PMID: 23582281
4. Parvathaneni V, Kulkarni NS, Muth A, Gupta V. Drug repurposing: a promising tool to accelerate the drug discovery process. *Drug Discov Today*. 2019;24(10):2076–85. <https://doi.org/10.1016/j.drudis.2019.06.014> PMID: 31238113
5. Choudhury C, Arul Murugan N, Priyakumar UD. Structure-based drug repurposing: Traditional and advanced AI/ML-aided methods. *Drug Discov Today*. 2022;27(7):1847–61. <https://doi.org/10.1016/j.drudis.2022.03.006> PMID: 35301148
6. Chen H, Zhang H, Zhang Z, Cao Y, Tang W. Network-based inference methods for drug repositioning. *Comput Math Methods Med*. 2015;2015:130620. <https://doi.org/10.1155/2015/130620> PMID: 25969690
7. Liu H, Song Y, Guan J, Luo L, Zhuang Z. Inferring new indications for approved drugs via random walk on drug-disease heterogeneous networks. *BMC Bioinformatics*. 2016;17(Suppl 17):539. <https://doi.org/10.1186/s12859-016-1336-7> PMID: 28155639
8. Huang Y, Bin Y, Zeng P, Lan W, Zhong C. NetPro: Neighborhood Interaction-Based Drug Repositioning via Label Propagation. *IEEE/ACM Trans Comput Biol Bioinform*. 2023;20(3):2159–69. <https://doi.org/10.1109/TCBB.2023.3234331> PMID: 37018341
9. Yang M, Luo H, Li Y, Wu F-X, Wang J. Overlap matrix completion for predicting drug-associated indications. *PLoS Comput Biol*. 2019;15(12):e1007541. <https://doi.org/10.1371/journal.pcbi.1007541> PMID: 31869322
10. Zhang W, Xu H, Li X, Gao Q, Wang L. DRIMC: an improved drug repositioning approach using Bayesian inductive matrix completion. *Bioinformatics*. 2020;36(9):2839–47. <https://doi.org/10.1093/bioinformatics/btaa062> PMID: 31999326
11. Jamali AA, Tan Y, Kusalik A, Wu F-X. NTD-DR: Nonnegative tensor decomposition for drug repositioning. *PLoS One*. 2022;17(7):e0270852. <https://doi.org/10.1371/journal.pone.0270852> PMID: 35862409
12. Jiang H-J, Huang Y-A, You Z-H. SAEROF: an ensemble approach for large-scale drug-disease association prediction by incorporating rotation forest and sparse autoencoder deep neural network. *Sci Rep*. 2020;10(1):4972. <https://doi.org/10.1038/s41598-020-61616-9> PMID: 32188871
13. Zhao B-W, Hu L, You Z-H, Wang L, Su X-R. HINGRL: predicting drug-disease associations with graph representation learning on heterogeneous information networks. *Brief Bioinform*. 2022;23(1):bbab515. <https://doi.org/10.1093/bib/bbab515> PMID: 34891172
14. Yang R, Fu Y, Zhang Q, Zhang L. GCNGAT: Drug-disease association prediction based on graph convolution neural network and graph attention network. *Artif Intell Med*. 2024;150:102805. <https://doi.org/10.1016/j.artmed.2024.102805> PMID: 38553169
15. Lee J, Yoon W, Kim S, Kim D, Kim S, So CH, et al. BioBERT: a pre-trained biomedical language representation model for biomedical text mining. *Bioinformatics*. 2020;36(4):1234–40. <https://doi.org/10.1093/bioinformatics/btz682> PMID: 31501885
16. Chithrananda S, Grand G, Ramsundar B. ChemBERTa: large-scale self-supervised pretraining for molecular property prediction, arXiv preprint arXiv:2010.09885 2020.
17. Lin Z, Akin H, Rao R, Hie B, Zhu Z, Lu W, et al. Evolutionary-scale prediction of atomic-level protein structure with a language model. *Science*. 2023;379(6637):1123–30. <https://doi.org/10.1126/science.ade2574> PMID: 36927031
18. Jablonka KM, Schwaller P, Ortega-Guerrero A, Smit B. Leveraging large language models for predictive chemistry. *Nat Mach Intell*. 2024;6(2):161–9. <https://doi.org/10.1038/s42256-023-00788-1>
19. Madani A, Krause B, Greene ER, Subramanian S, Mohr BP, Holton JM, et al. Large language models generate functional protein sequences across diverse families. *Nat Biotechnol*. 2023;41(8):1099–106. <https://doi.org/10.1038/s41587-022-01618-2> PMID: 36702895
20. Ahmed KT, Ansari MI, Zhang W. DTI-LM: language model powered drug-target interaction prediction. *Bioinformatics*. 2024;40(9):btae533. <https://doi.org/10.1093/bioinformatics/btae533> PMID: 39221997
21. Shang Y, Wang Z, Chen Y, Yang X, Ren Z, Zeng X, et al. HNF-DDA: subgraph contrastive-driven transformer-style heterogeneous network embedding for drug-disease association prediction. *BMC Biol*. 2025;23(1):101. <https://doi.org/10.1186/s12915-025-02206-x> PMID: 40241152
22. Luo Z, Wu W, Sun Q. Accurate and transferable drug–target interaction prediction with DrugLAMP. *Bioinformatics*. 2024;40(btae693).
23. Fan S, Yang K, Lu K, Dong X, Li X, Zhu Q, et al. DrugRepPT: a deep pretraining and fine-tuning framework for drug repositioning based on drug's expression perturbation and treatment effectiveness. *Bioinformatics*. 2024;40(12):btae692. <https://doi.org/10.1093/bioinformatics/btae692> PMID: 39563444
24. Gottlieb A, Stein GY, Ruppin E, Sharan R. PREDICT: a method for inferring novel drug indications with application to personalized medicine. *Mol Syst Biol*. 2011;7:496. <https://doi.org/10.1038/msb.2011.26> PMID: 21654673
25. Li F, Mou M, Li X, Xu W, Yin J, Zhang Y, et al. DrugMAP 2.0: molecular atlas and pharma-information of all drugs. *Nucleic Acids Res*. 2025;53(D1):D1372–82. <https://doi.org/10.1093/nar/gkae791> PMID: 39271119
26. Ahmad W, Simon E, Chithrananda S. Chemberta-2: Towards chemical foundation models. arXiv preprint. 2022. <https://doi.org/10.48550/arXiv.2209.01712>

27. Devlin J, Chang M-W, Lee K. Bert: Pre-training of deep bidirectional transformers for language understanding. In: Proceedings of the 2019 conference of the North American chapter of the association for computational linguistics: human language technologies, 2019. 4171–86.
28. Reimers N, Gurevych I. Sentence-bert: Sentence embeddings using siamese bert-networks. arXiv preprint. 2019. <https://doi.org/10.48550/arXiv.1908.10084>
29. Yang Z, Yang D, Dyer C, He X, Smola A, Hovy E. Hierarchical Attention Networks for Document Classification. In: Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, 2016. 1480–9. <https://doi.org/10.18653/v1/n16-1174>
30. Chen T, Kornblith S, Norouzi M. A simple framework for contrastive learning of visual representations. In: International conference on machine learning, 2020. 1597–607.
31. Zhang W, Yue X, Lin W, Wu W, Liu R, Huang F, et al. Predicting drug-disease associations by using similarity constrained matrix factorization. BMC Bioinformatics. 2018;19(1):233. <https://doi.org/10.1186/s12859-018-2220-4> PMID: 29914348
32. Gharizadeh A, Abbasi K, Ghareyazi A, Mofrad MRK, Rabiee HR. HGTD: Advancing drug repurposing with heterogeneous graph transformers. Bioinformatics. 2024;40(7):btac349. <https://doi.org/10.1093/bioinformatics/btac349> PMID: 38913860
33. Grover A, Leskovec J. node2vec: Scalable Feature Learning for Networks. KDD. 2016;2016:855–64. <https://doi.org/10.1145/2939672.2939754> PMID: 27853626
34. Quan N, Ma S, Zhao K, Bi X, Zhang L. MFCADTI: improving drug-target interaction prediction by integrating multiple feature through cross attention mechanism. BMC Bioinformatics. 2025;26(1):57. <https://doi.org/10.1186/s12859-025-06075-7> PMID: 39966727
35. Luo Y, Zhao X, Zhou J, Yang J, Zhang Y, Kuang W, et al. A network integration approach for drug-target interaction prediction and computational drug repositioning from heterogeneous information. Nat Commun. 2017;8(1):573. <https://doi.org/10.1038/s41467-017-00680-8> PMID: 28924171
36. Cohen J. Statistical power analysis for the behavioral sciences. Routledge. 2013.
37. Lipscomb CE. Medical Subject Headings (MeSH). Bull Med Libr Assoc. 2000;88(3):265–6. PMID: 10928714
38. Shorning BY, Dass MS, Smalley MJ, Pearson HB. The PI3K-AKT-mTOR Pathway and Prostate Cancer: At the Crossroads of AR, MAPK, and WNT Signaling. Int J Mol Sci. 2020;21(12):4507. <https://doi.org/10.3390/ijms21124507> PMID: 32630372
39. Breistøl K, Hendriks HR, Fodstad O. Superior therapeutic efficacy of N-L-leucyl-doxorubicin versus doxorubicin in human melanoma xenografts correlates with higher tumour concentrations of free drug. Eur J Cancer. 1999;35(7):1143–9. [https://doi.org/10.1016/s0959-8049\(99\)00074-x](https://doi.org/10.1016/s0959-8049(99)00074-x) PMID: 10533461
40. Wang S, Jiang M, Wang Y, Chang L, Zhao B, Liu X, et al. Venetoclax plus Modified-Intensity Idarubicin and Cytarabine Treatment as First-Line Treatment for Newly Diagnosed Pediatric Acute Myeloid Leukemia. Clin Cancer Res. 2025;31(13):2608–16. <https://doi.org/10.1158/1078-0432.CCR-25-0479> PMID: 40293351
41. Xiang Y, Wang X, Yan C, Gao Q, Li S-A, Liu J, et al. Adenosine-5'-triphosphate (ATP) protects mice against bacterial infection by activation of the NLRP3 inflammasome. PLoS One. 2013;8(5):e63759. <https://doi.org/10.1371/journal.pone.0063759> PMID: 23717478
42. Hong S-J, Lee Y-J, Lee S-J, Hong B-K, Kang WC, Lee J-Y, et al. Treat-to-Target or High-Intensity Statin in Patients With Coronary Artery Disease: A Randomized Clinical Trial. JAMA. 2023;329(13):1078–87. <https://doi.org/10.1001/jama.2023.2487> PMID: 36877807
43. Haidar A, Tsoukas MA, Bernier-Twardy S, Yale J-F, Rutkowski J, Bossy A, et al. A Novel Dual-Hormone Insulin-and-Pramlintide Artificial Pancreas for Type 1 Diabetes: A Randomized Controlled Crossover Trial. Diabetes Care. 2020;43(3):597–606. <https://doi.org/10.2337/dc19-1922> PMID: 31974099
44. Wald ER. Dexamethasone as Adjuvant Therapy for Bacterial Meningitis in Children: What About Streptococcus pneumoniae?. Open Forum Infect Dis. 2025;12(8):ofaf456. <https://doi.org/10.1093/ofid/ofaf456> PMID: 40874187
45. Parihar V, Maguire S, Shahin A, Ahmed Z, O'Sullivan M, Kennedy M, et al. Listeria meningitis complicating a patient with ulcerative colitis on concomitant infliximab and hydrocortisone. Ir J Med Sci. 2016;185(4):965–7. <https://doi.org/10.1007/s11845-015-1355-9> PMID: 26358724
46. Pascual R, Valencia M, Bustamante C. Antenatal betamethasone produces protracted changes in anxiety-like behaviors and in the expression of microtubule-associated protein 2, brain-derived neurotrophic factor and the tyrosine kinase B receptor in the rat cerebellar cortex. Int J Dev Neurosci. 2015;43:78–85. <https://doi.org/10.1016/j.ijdevneu.2015.04.005> PMID: 25889225
47. Guzzo EFM, de Lima Rosa G, Domingues AM, Padilha RB, Coitinho AS. Reduction of seizures and inflammatory markers by betamethasone in a kindling seizure model. Steroids. 2023;193:109202. <https://doi.org/10.1016/j.steroids.2023.109202> PMID: 36828350
48. Warrington TP, Bostwick JM. Psychiatric adverse effects of corticosteroids. Mayo Clin Proc. 2006;81(10):1361–7. <https://doi.org/10.4065/81.10.1361> PMID: 17036562
49. Unterman A, Nolte JES, Boaz M, Abady M, Shoenfeld Y, Zandman-Goddard G. Neuropsychiatric syndromes in systemic lupus erythematosus: a meta-analysis. Semin Arthritis Rheum. 2011;41(1):1–11. <https://doi.org/10.1016/j.semarthrit.2010.08.001> PMID: 20965549
50. Vézina C, Kudelski A, Sehgal SN. Rapamycin (AY-22,989), a new antifungal antibiotic. I. Taxonomy of the producing streptomycete and isolation of the active principle. J Antibiot (Tokyo). 1975;28(10):721–6. <https://doi.org/10.7164/antibiotics.28.721> PMID: 1102508
51. Sehgal SN. Sirolimus: its discovery, biological properties, and mechanism of action. Transplant Proc. 2003;35(3 Suppl):7S-14S. [https://doi.org/10.1016/s0041-1345\(03\)00211-2](https://doi.org/10.1016/s0041-1345(03)00211-2) PMID: 12742462
52. Li J, Kim SG, Blenis J. Rapamycin: one drug, many effects. Cell Metabolism. 2014;19:373–9.

53. Choi J, Chen J, Schreiber SL, Clardy J. Structure of the FKBP12-Rapamycin Complex Interacting with Binding Domain of Human FRAP. *Science*. 1996;273(5272):239–42. <https://doi.org/10.1126/science.273.5272.239>
54. Liang J, Choi J, Clardy J. Refined structure of the FKBP12-rapamycin-FRB ternary complex at 2.2 Å resolution. *Acta Crystallogr D Biol Crystallogr*. 1999;55(Pt 4):736–44. <https://doi.org/10.1107/s0907444998014747> PMID: [10089303](https://pubmed.ncbi.nlm.nih.gov/10089303/)
55. Liu Q, Thoreen C, Wang J, Sabatini D, Gray NS. mTOR Mediated Anti-Cancer Drug Discovery. *Drug Discov Today Ther Strateg*. 2009;6(2):47–55. <https://doi.org/10.1016/j.ddstr.2009.12.001> PMID: [20622997](https://pubmed.ncbi.nlm.nih.gov/20622997/)
56. Guertin DA, Sabatini DM. Defining the role of mTOR in cancer. *Cancer Cell*. 2007;12(1):9–22. <https://doi.org/10.1016/j.ccr.2007.05.008> PMID: [17613433](https://pubmed.ncbi.nlm.nih.gov/17613433/)
57. Marques-Ramos A, Cervantes R. Expression of mTOR in normal and pathological conditions. *Mol Cancer*. 2023;22(1):112. <https://doi.org/10.1186/s12943-023-01820-z> PMID: [37454139](https://pubmed.ncbi.nlm.nih.gov/37454139/)
58. Panwar V, Singh A, Bhatt M, Tonk RK, Azizov S, Raza AS, et al. Multifaceted role of mTOR (mammalian target of rapamycin) signaling pathway in human health and disease. *Signal Transduct Target Ther*. 2023;8(1):375. <https://doi.org/10.1038/s41392-023-01608-z> PMID: [37779156](https://pubmed.ncbi.nlm.nih.gov/37779156/)
59. Kitchen DB, Decornez H, Furr JR, Bajorath J. Docking and scoring in virtual screening for drug discovery: methods and applications. *Nat Rev Drug Discov*. 2004;3(11):935–49. <https://doi.org/10.1038/nrd1549> PMID: [15520816](https://pubmed.ncbi.nlm.nih.gov/15520816/)
60. Jeffrey GA, Jeffrey GA. An introduction to hydrogen bonding. New York: Oxford University Press. 1997.