Article

HeTDR: Drug repositioning based on heterogeneous networks and text mining

Graphical abstract

Highlights

- We developed a novel DL-based method for drug repositioning (HeTDR)

- HeTDR succeeds in fusing networks topology information and text mining information

- HeTDR obtains high accuracy, excessing most state-of-the-art models

- HeTDR could represent an algorithm integrating multiple sources of information
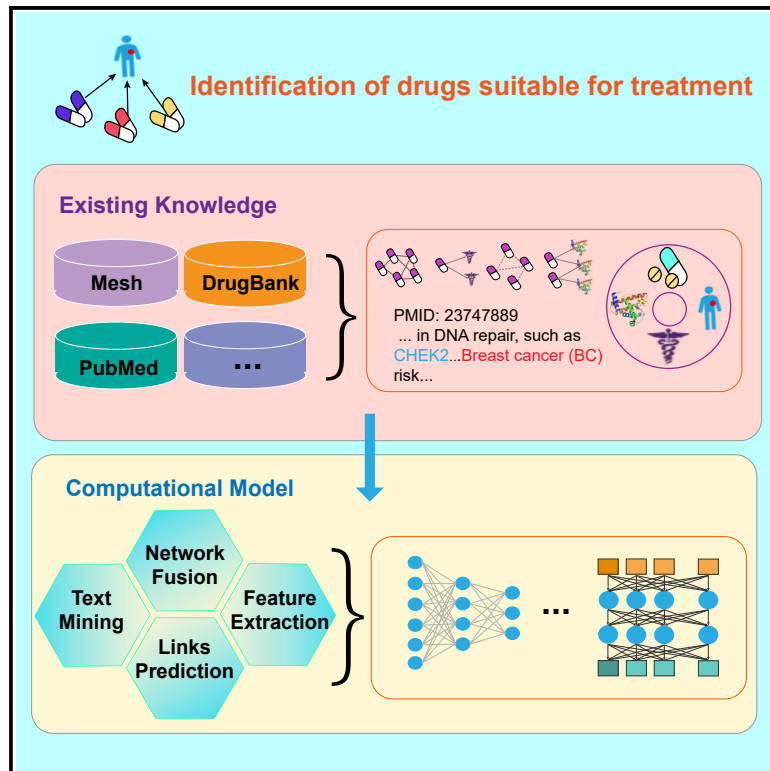
Authors

Shuting Jin, Zhangming Niu, Changzhi Jiang, ..., Xurui Jin, Xiangrong Liu, Xiangxiang Zeng

Correspondence

xrliu@xmu.edu.cn

In brief

Drug repositioning is a useful way to discover new drug candidates for curing diseases. However, integrating multiple networks and text mining information for drug repositioning is still a major challenge. We propose a drug repositioning method based on heterogeneous networks and text mining (HeTDR), which can combine drug features from multiple networks and disease features from biomedical corpora for drug repositioning. HeTDR obtains high accuracy in predicting drug-disease interactions and is capable of finding novel indications of approved drugs.

CellPress

Patterns

Article

HeTDR: Drug repositioning based on heterogeneous networks and text mining

Shuting Jin,[1,2,3] Zhangming Niu,[4] Changzhi Jiang,[1] Wei Huang,[1] Feng Xia,[1] Xurui Jin,[4] Xiangrong Liu,[1,2,6,*] and Xiangxiang Zeng[5]

[1]Department of Computer Science, Xiamen University, Xiamen 361005, China
[2]National Institute for Data Science in Health and Medicine, Xiamen University, Xiamen 361005, China
[3]Shenzhen Research Institute of Xiamen University, Shenzhen 518000, China
[4]MindRank AI Ltd., Hangzhou, Zhejiang 311113, China
[5]School of Information Science and Engineering, Hunan University, Changsha 410082, China
[6]Lead contact
*Correspondence: xrliu@xmu.edu.cn
https://doi.org/10.1016/j.patter.2021.100307

THE BIGGER PICTURE Traditional drug discovery and development are often time consuming and high risk. Drug repositioning aims to expand existing indications or discover new targets by studying the approved drug compounds, thereby reducing the time, costs, and risk of drug development. We propose a novel method in drug repositioning based on heterogeneous networks and text mining (HeTDR), which combines drugs features from multiple networks and diseases features from biomedical corpora to predict the correlation scores between drugs and diseases. This prediction model has provided a potential solution for multiple information fusion and to exhibit accurate performance leading to the discovery of new drugs for indications. This algorithm could contribute a new idea to the acceleration and development of future drug repositioning by using computational methods and provide computer-aided guidance for biologists in clinical settings.

1 2 3 4 5   **Development/Pre-production:** Data science output has been rolled out/validated across multiple domains/problems

---

SUMMARY

Using existing knowledge to carry out drug-disease associations prediction is a vital method for drug repositioning. However, effectively fusing the biomedical text and biological network information is one of the great challenges for most current drug repositioning methods. In this study, we propose a drug repositioning method based on heterogeneous networks and text mining (HeTDR). This model can combine drug features from multiple drug-related networks, disease features from biomedical corpora with the known drug-disease associations network to predict the correlation scores between drug and disease. Experiments demonstrate that HeTDR has excellent performance that is superior to that of state-of-the-art models. We present the top 10 novel HeTDR-predicted approved drugs for five diseases and prove our model is capable of discovering potential candidate drugs for disease indications.

INTRODUCTION

The time for new drug development has been gradually increasing, and the cost it takes to bring a new drug to market is becoming more expensive up to 2.6 billion dollars.[1,2] In recent years, an increasing number of studies have shown that some approved drugs can be used to treat new indications, this process is referred to as drug repositioning.[3] Drug re-positioning aims to expand the existing indications or discover new targets by studying the approved drug compounds, thereby reducing the time, costs, and risks in drug development.[4] In the beginning, most of the repositioned drugs were accidentally discovered in clinical settings. With the development of high-throughput technology, the increase in large-scale genomics and pharmacological and chemical datasets has made it possible to predict the relationship between drug

with disease by using systematic and reasonable calculation methods.

Drug repositioning is a long-standing problem and many calculation methods have been proposed to predict drug-disease associations for drug repositioning. These studies could be roughly classified into three groups: machine learning, network-based methods, and literature mining. Most of the machine learning methods formulated drug repositioning as a classification task, and some machine learning classification algorithms are widely used, such as support vector machines (SVMs),[5–7] logistic regression,[7,8] and random forest,[7] to identify the potential indications for approved drugs. However, those supervised classification methods will randomly generate negative samples, which could lead to biased decisions. The network-based methods are the most widely used methods for drug repositioning. Its mainstream algorithms include: network inference,[9–12] random walk,[13,14] and matrix factorization.[15,16] These methods usually rely heavily on the richness of interaction network data. In addition, many studies that used such methods proposed the need to construct more complete and large-scale data information to improve prediction performance. The performance of literature mining-based methods relies on biomedical entities' co-occurrence and semantic inference of some keywords of interest. These methods will be limited by the ambiguity of natural language and the limitations of text-mining technology.[17–20] Recently, deep learning has achieved tremendous development, and the models of network representation[21,22] and text mining[23] have been continuously improved. However, comprehensively considering topological properties and statistical correlation to discover potential drug-disease relationships remains a challenge.

In this study, we develop a novel method, called HeTDR, that allows the incorporation of topological structure information from the heterogeneous networks and the features information from biomedical text mining. In our model, the drug features extraction module is based on similarity network fusion (SNF)[24] and sparse autoencode (SAE),[25,26] and the disease features extraction module is based on Bidirectional Encoder Representations from Transformers for Biomedical Text Mining (BioBERT).[27] Some previous studies proved that combining the drug-disease associations network with feature information could make the prediction model obtain more robust results than using the feature information alone.[10,28] Accordingly, we adopt a model that can combine the attribute information with the topological structure information for final prediction. The contributions of HeTDR can be enumerated as follows: (1) HeTDR integrates nine drug-related networks into a low-dimensional and compact feature representation common to all networks, which can better capture the overall information of the drug. (2) HeTDR obtains the information of disease from biomedical corpora, thereby improving the accuracy of drug-disease associations prediction. (3) HeTDR makes full use of both topology and attribute information to overcome the influence of network data often being partially observed. (4) HeTDR combines a network-based method and a text mining-based method to provide a new solution for drug repositioning. In the computational experiments, HeTDR obtains high accuracy in predicting drug-disease interactions, and significantly outperforms existing state-of-the-art methods. Moreover, case studies show that HeTDR can help discover novel associations not included in known drug-disease pairs and find novel indications of approved drugs.

RESULTS

The HeTDR pipeline

The workflow of HeTDR is shown in Figure 1, and consists of three parts: (1) HeTDR uses SNF to integrate nine drug-related networks into one network with global information and then utilizes the SAE to obtain high-quality features representation of the drug. (2) HeTDR uses the BioBERT model to obtain disease features information from biomedical corpora. Specifically, we use the pre-trained parameters of the BioBERT, and select the relation extraction task for fine-tuning training. After the fine-tuning process, we extract the representation of sub-words and obtain the representations of all diseases by frequent sub-words. (3) HeTDR combines the drug-disease associations network with the features information of the drug and disease to infer the potential associations between drug and disease.

Evaluation of prediction performance by ablation analysis

We partition 6,677 reported drug-disease pairs into three subsets, 80% of the known drug-disease pairs for training, 10% for validation, and 10% for testing. For the validation set and test set, we randomly generate negative samples with a 1:1 ratio matched with the positive samples. To evaluate the performance of HeTDR, we use the following performance metrics: area under the receiver operating characteristic curve (AUROC), the area under the precision-recall curve (AUPR), and F1-measure (F1).

HeTDR showed a high performance (AUROC = 0.959, AUPR = 0.955, and F1 = 0.901) (Figure 2). We conduct ablation studies on HeTDR to explore the effect of drug features and disease features. We run experiments with the same data splits, model parameters, and evaluation protocol to keep the comparison as fair as possible. The result is presented in Figure 2. We implement three simplified variants of HeTDR.

(1) HeTDR_Di: by removing the drug features.
(2) HeTDR_Dr: by removing the disease features.
(3) HeTDR_SNF: by removing the SAE.

Minimal difference is found in the results obtained by using drug features (AUROC = 0.916, AUPR = 0.915, and F1 = 0.830) or disease features (AUROC = 0.921, AUPR = 0.919, and F1 = 0.838) alone, but combining these features information on disease and drug has the highest accuracy of drug-disease associations prediction (AUROC = 0.959, AUPR = 0.955, and F1 = 0.901). The result of using SAE when obtaining drug features is better than the output when SNF is used to obtain the drug features (AUROC = 0.934, AUPR = 0.928, and F1 = 0.862), which illustrates the importance of filtering information when fusing multiple types of networks data. HeTDR makes full use of multiple drug-related information and disease-related information, which accounts for its superior performance over other methods.

Evaluation of prediction performance on cross-validation

To evaluate HeTDR more comprehensively, we conduct 5-fold cross-validation and compare HeTDR with classic methods. We compare our results with seven methods: SVM,[29] Katz,[30] MBiRW,[14] DTINet,[31] DRRS,[32] DeepDR,[10] and HNet-DNN.[12]
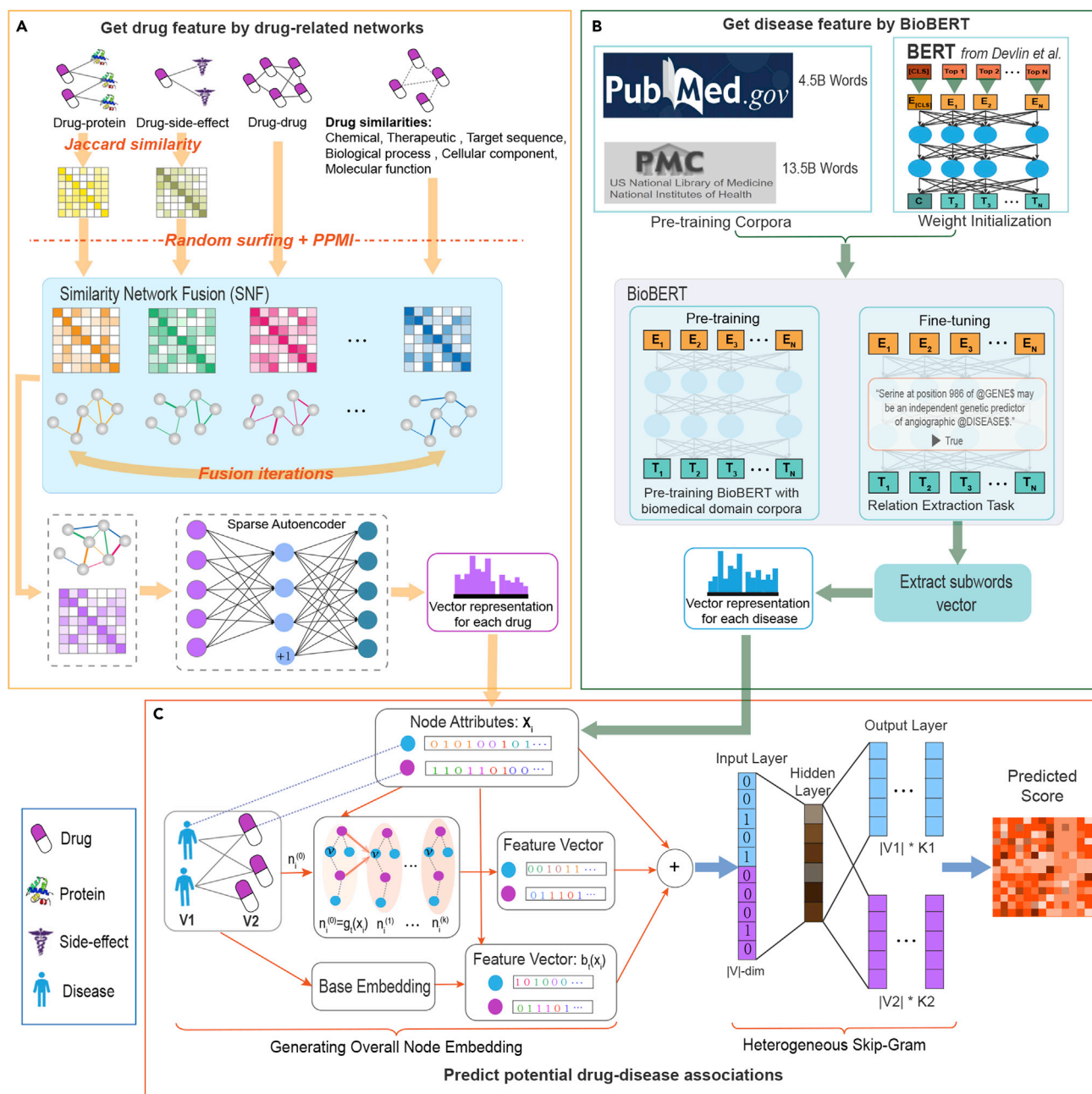
Figure 1. Flowchart of HeTDR

The model consists of three parts: (A) HeTDR integrates nine drug-related networks to obtain global information of drugs. In the heterogeneous interaction networks, we first use the Jaccard similarity coefficient to calculate the similarity network. Then, we fuse these drug-related networks into one network by SNF and apply SAE to obtain low-dimensional features of the drugs. (B) HeTDR obtains vector representation of the disease features by text mining biomedical corpora. In the pre-training stage, we directly use the model parameters pre-trained by BioBERT. Then, we select the relation extraction task for fine-tuning training. After the fine-tuning process has taken place, we extract the representations of sub-words and use the representations of these sub-words to obtain the representations of all diseases. (C) HeTDR predicts potential drug-disease associations by an embedding learning method, which can capture both the drug-disease associations network topological structural proximity and node attributes proximity.

SVM is the traditional machine learning algorithm; Katz is the path-based classic algorithm calculating similarities between nodes in a network for associations prediction; MBiRW utilizes some comprehensive similarity measures and the BiRandom Walk (BiRW) algorithm for drug repositioning; DTINet is a matrix factor-ization-based model and it can integrate diverse information from heterogeneous networks; DRRS is a computational drug reposi-tioning method using low-rank matrix appropriation and random-ized algorithm; DeepDR is a network-based deep learning method to infer potential novel drug-disease associations; HNet-DNN
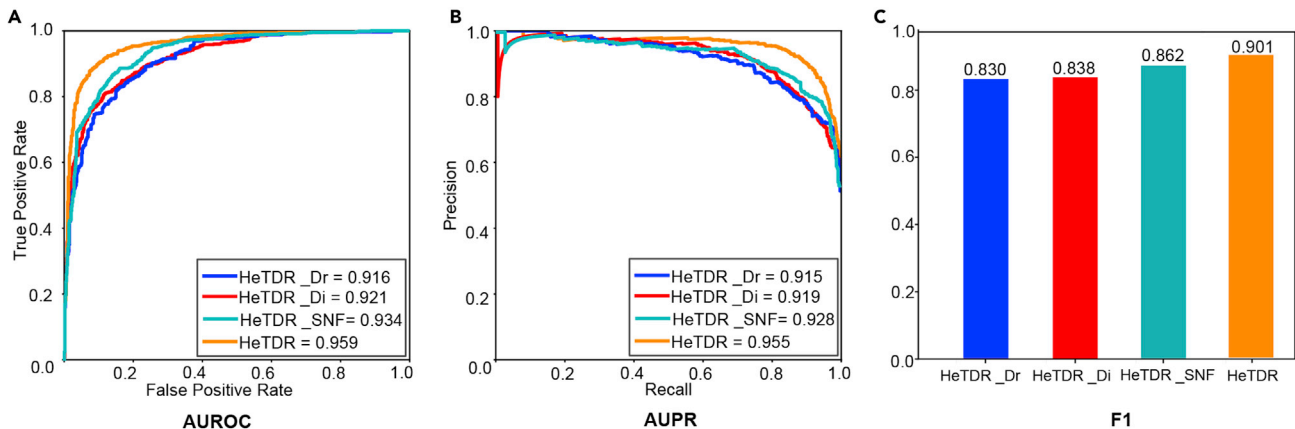
Figure 2. Performance of HeTDR comparing the different features
(A) ROC curves of prediction results by using different features.
(B) PR curves of prediction results by using different attributes.
(C) F1 scores of prediction results by using different features.

uses the deep neural network (DNN) to predict new drug-disease interactions. We use AUROC and AUPR to evaluate the performance of these methods. In particular, each method is configured to its default setting or best parameter values reported in its paper. In the 5-fold cross-validation results, HeTDR obtains high performance (AUROC = 0.979 and AUPR = 0.976), outperforming other state-of-the-art methods: SVM (AUROC = 0.602 and AUPR = 0.580), Katz (AUROC = 0.701 and AUPR = 0.722), MBiRW (AUROC = 0.812 and AUPR = 0.810), DTINet (AUROC = 0.870 and AUPR = 0.884), DRRS (AUROC = 0.912 and AUPR = 0.877), DeepDR (AUROC = 0.915 and AUPR = 0.927), and HNet-DNN (AUROC = 0.951 and AUPR = 0.908) (Figure 3).

HeTDR identifies novel drug-disease associations
We save the results of HeTDR predictions of drug-disease associations in "Evaluate prediction performance by ablation analysis." We delete 6,677 known drug-disease associations used in the prediction model, and select the novel top 150 pairs with highest similarity drug-disease pairs predicted by HeTDR, and visualize the interactions network (Figure 4). Among the list of top 150 predictions, HeTDR can capture the experimental or clinical reported drug-disease associations. For example, novel predictions show

that DB08827 (lomitapide) and DB05528 (mipomersen) can act on C0342882 (familial hypercholesterolemia-heterozygous). The novel prediction was confirmed by previous studies. Lomitapide is an inhibitor of a microsomal triglyceride transfer protein that lowers hepatic low-density lipoprotein cholesterol production, as new therapeutic modalities treating homozygous familial hypercholesterolemia.[33,34] Mipomersen decreases the levels of apolipoprotein B, low-density lipoprotein non-high-density lipoprotein cholesterol, and total cholesterol, which are used in patients with homozygous familial hypercholesterolemia as an adjunct to diet and other lipid-lowering medications.[34,35] For C0026705 (mucopolysaccharidosis II), the significant association with DB00090 (laronidase) is successfully identified by HeTDR.

To interpret the results of the HeTDR model, we analyze the novel associations predicted in the example. We use the features learned by the HeTDR model to calculate the drug-drug and disease-disease similarity relationships in the association prediction module at the model. We select the top 20 diseases predicted to be most related to C0342882 and C0026705 (see Table S1), and use known drug-disease associations to find approved drugs that can be used to treat these diseases (Figure 5). In the model, C0342881 (familial hypercholesterolemia-homozygous)
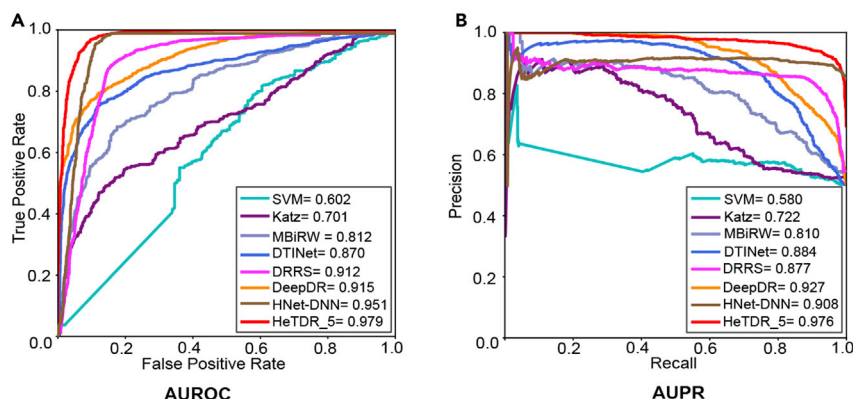

Figure 3. HeTDR outperforms other state-of-the-art methods for drug-disease associations prediction
(A) ROC curves of prediction results obtained by applying HeTDR and five previously reported methods in 5-fold cross-validation.
(B) PR curves of prediction results obtained by HeTDR and five previously reported methods in 5-fold cross-validation.
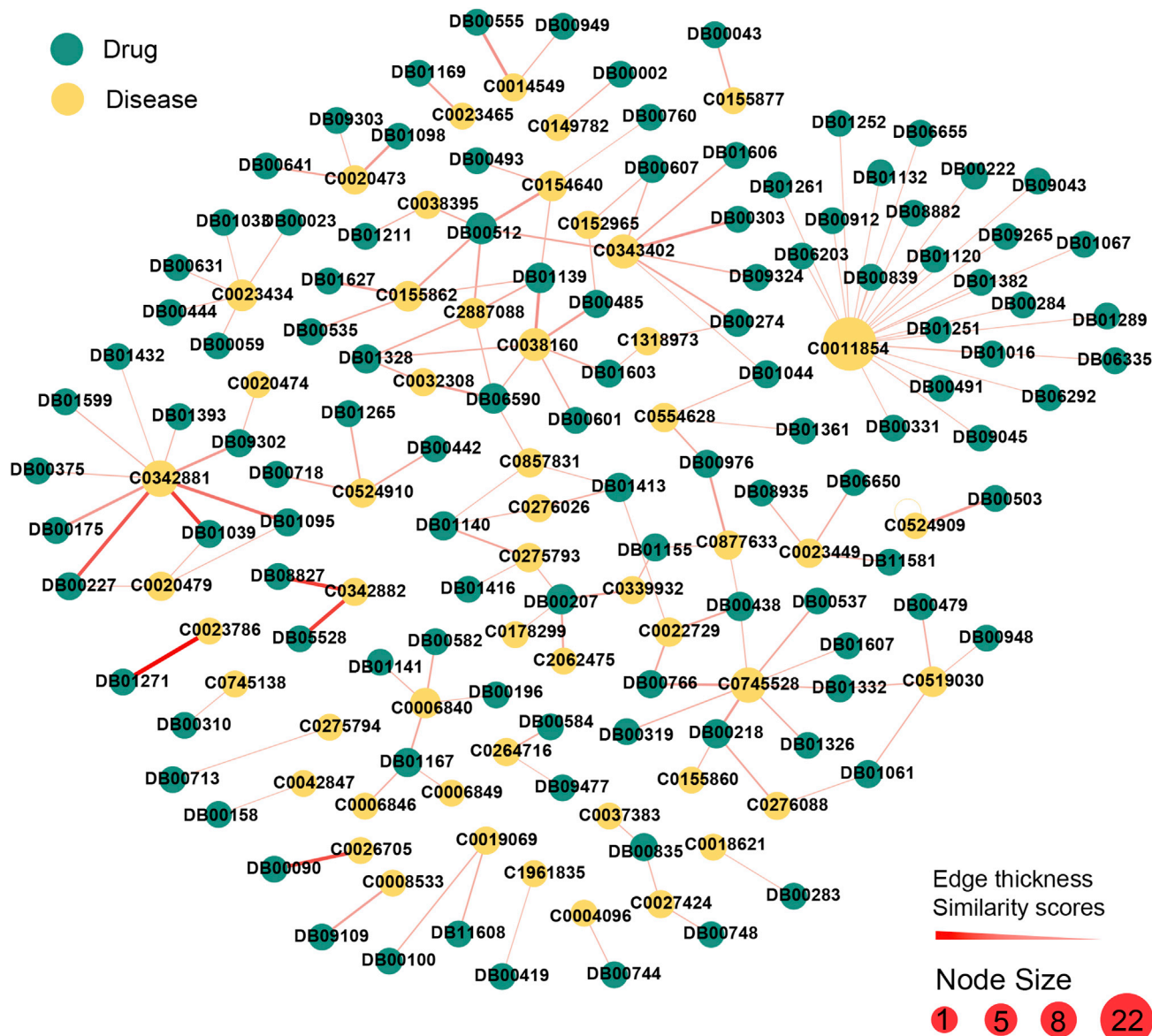
Figure 4. Network visualization of the drug-disease associations predicted by HeTDR
In this network, the predicted novel top 150 drug-disease pairs network is visualized. The label of the node represents the ID of the drugs (Drugbank_ID) or diseases (UMLS_ID). The node size denotes the degree. The weight of edges (drug-disease pairs) denotes the predicted score by HeTDR. The novel top 150 pairs of the highest similarity drug-disease associations can be found in Table S1. This image was generated by Gephi (https://gephi.org).

is most related to C0342882, and DB05528 and DB08827 are drugs that have been approved to act on C0342881. The disease most related to C0026705 captured by our model is C0023786 (mucopolysaccharidosis I), and DB00090 is an approved drug that can be used to treat C0023786. Accordingly, DB00090 has a high potential to cure C0026705. These results demonstrate that our model can make full use of features information to capture approximate relationships of drug-drug and disease-disease, and reveal potential drug-disease associations based on these approximate relationships and the known associations. The novel top 150 pairs of the highest similarity drug-disease associations can be found in Table S2.

Case studies

To further demonstrate the ability of HeTDR for discovering novel drug-disease associations, we choose five diseases for case studies, namely, Alzheimer's disease, obesity, asthma, epilepsy, and Parkinson's disease. Similarly, we delete known drug-disease associations used in the prediction model, and select the top 10 drugs according to the predicted highest association scores, which are considered drug candidates for the disease, as shown in Table 1. To provide a more reliable reference for follow-up researchers, we also provide the associated prediction scores for the top 20 drugs of all diseases, and the number of side effects associated with these drugs (see Table S3). All the
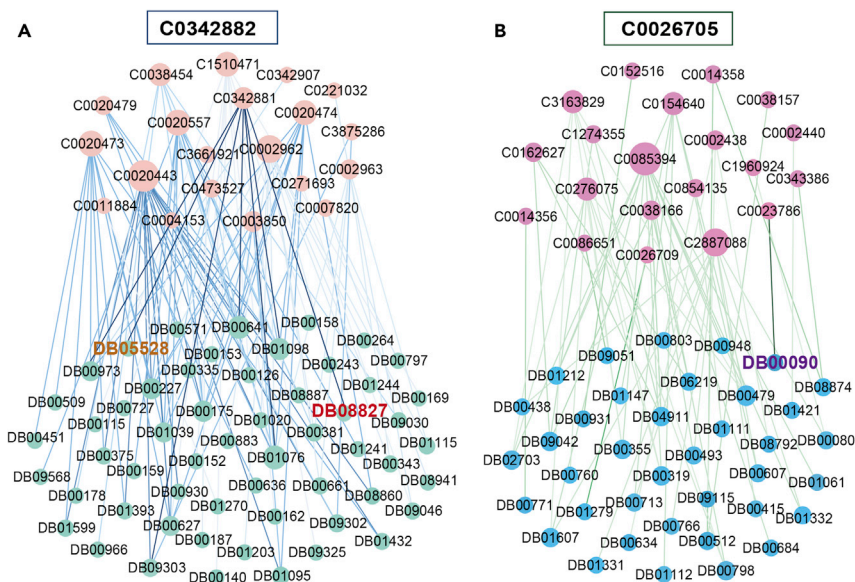
Figure 5. The interpretability of HeTDR Identifies novel associations

(A) The upper nodes are the top 20 most relevant diseases of C0342882.

(B) The upper nodes are the top 20 most relevant diseases of C0026705. The edges are the known drug-disease associations, and the heavier color of the edge represents edge linking disease rank higher in the top 20. This image was generated by Gephi (https://gephi.org).

drug-related networks by using SNF and obtained low-dimensional and compact drug features representation by using SAE. Then, HeTDR used the BioBERT to obtain disease features representation from biomedical corpora. Finally, HeTDR combined clinically reported drug-disease pairs with disease and drug features representations to predict potential drug-disease associations. HeTDR performed better than existing drug reposi-

side effects and the drug-side effect interactions can be found in the data and code availability.

We first use the Comparative Toxicogenomics Database (CTD) (http://ctdbase.org/) and the DrugBank database (https://go.drugbank.com) to validate the top 10 drugs predicted by HeTDR. The CTD (http://ctdbase.org/) is a publicly available database resource providing manually curated key information about the chemical-disease, chemical-gene, and gene-disease interactions from the literature. The DrugBank database is comprehensive molecular information that integrates bioinformatics and chemoinformatics resources and provides detailed drugs data, drugs target information, drugs mechanism, and drugs-related clinical trials information. In addition, we verify whether novel associations that could not be identified in CTD and DrugBank can be supported by previously published studies (Table 1). The results demonstrate that the HeTDR is an effective method to find the potential drug-disease associations and can help develop candidate drugs for the disease.

Then, we compute the associated prediction scores by HeTDR_Di and HeTDR_Dr, and screen out the top 20 predicted drugs that may act on the five diseases. Through analysis, it is found that some of the potential drugs with higher rankings predicted by HeTDR, such as zileuton and metixene, are also captured in the top 20 drugs predicted by HeTDR_Di and HeTDR_Dr. Some drugs, such as Prednisone for the treatment of Alzheimer's disease, can only be found in the top 20 results of HeTDR_Dr, and corticotropin also used for the treatment of Alzheimer's disease, can only be found in the results of HeTDR_Di. This also shows that HeTDR combined with network-based and text-based information has complementary effects, and can better reveal potential novel associations.

DISCUSSION

In this article, a novel HeTDR model for drug repositioning was built with both network topology attributes and text mining information. HeTDR first fused diverse information from a multitude of different

tioning methods because we fused diverse information from a multitude of different drug-related network types and integrated disease-related information from biomedical text mining. The analysis and verification also showed that combining the information from the text and the network can make better predictions compared with only using the data from a single source. In addition, HeTDR could preserve the known drug-disease associations network's topological structure and node attribute proximity to predict novel drug-disease associations. Experiments have shown that the HeTDR model achieved state-of-the-art performance in drug-disease associations prediction. Case studies of five diseases further proved the effectiveness of our model in finding novel drug-disease associations, as validated by database records or literature.

We acknowledge that HeTDR still has some room for improvement. First, we directly used BioBERT's pre-training parameters, which may have certain limitations, and we expect to obtain more effective disease features information from medical records to further improve the predictive ability of our model in the future. Second, our model obtains features through different modules, and developing an end-to-end model that uses downstream tasks to obtain better features for associations prediction is possible. Despite the shortcomings, we tried to avoid the impact of these problems by verifying the effectiveness of each module in obtaining features. The constructed HeTDR model is still the most powerful model that integrates multiple types of information.

In summary, our model could be used as an effective method to predict drug-disease associations, develop a new idea for drug repositioning calculation, and provide computer-aided guidance for biologists in clinical trials.

EXPERIMENTAL PROCEDURES

Resource availability

Lead contact

Further information and requests for resources should be directed to and will be fulfilled by the lead contact, Xiangrong Liu (xrliu@xmu.edu.cn).

Table 1. The top 10 related candidate drugs for Alzheimer's disease, obesity, asthma, epilepsy, and Parkinson's disease

Disease name	Rank	Drug name	Description	Rank	Drug name	Description
Alzheimer's disease	1	corticotropin	CTD_I	6	ergoloid	[36–40]
	2	natalizumab	N/A	7	fludrocortisone	CTD_I
	3	dexamethasone	CTD_I	8	prednisone	CTD_I
	4	teriflunomide	N/A	9	budesonide	CTD_I
	5	canakinumab	[41–45]	10	dalfampridine	CTD_I
Obesity	1	lisdexamfetamine	CTD_I	6	nicotine	CTD_M
	2	methylphenidate	CTD_I	7	guanfacine	CTD_I
	3	dextroamphetamine	CTD_I	8	naloxegol	N/A
	4	atomoxetine	[46–51]	9	methylnaltrexone	[52,53]
	5	dexmethylphenidate	N/A	10	disulfiram	CTD_I
Asthma	1	zileuton	CTD_T	6	ipratropium bromide	CTD_T
	2	salbutamol	CTD_T	7	arformoterol	CTD_T
	3	montelukast	CTD_T	8	fluticasone propionate	CTD_T
	4	formoterol	CTD_T	9	dexpanthenol	CTD_I
	5	salmeterol	CTD_T	10	ephedrine	CTD_T
Epilepsy	1	lorazepam	DrugBank_T	6	oxcarbazepine	CTD_T
	2	clonazepam	CTD_T	7	levetiracetam	CTD_T
	3	ethosuximide	CTD_T	8	tiagabine	CTD_I
	4	stiripentol	DrugBank_T	9	nitrazepam	CTD_T
	5	topiramate	CTD_T	10	ethotoin	[54,55]
Parkinson's disease	1	pergolide	CTD_T	6	octreotide	CTD_I
	2	metixene	DrugBank_T	7	cabergoline	CTD_T
	3	orphenadrine	CTD_T	8	gabapentin	CTD_T
	4	rivastigmine	CTD_T	9	lanreotide	[56]
	5	gabapentin enacarbil	[57,58]	10	pegvisomant	N/A

CTD_I, inferred; CTD_M, curated (marker/mechanism); CTD_T, curated (therapeutic); DrugBank_T, therapeutic; N/A, could not be confirmed.

Materials availability
There are no physical materials associated with this study.

Data and code availability
The code and data are available at https://github.com/stjin-XMU/HeTDR, https://doi.org/10.5281/zenodo.4915882.

Data and preparation
We use the drug-related networks constructed by previous research,[10] to obtain the drug features and predict the drug-disease associations. Specif-

Table 2. Materials of networks

Node types	Numbers	Edge types	Numbers/Types
Drugs	1,519	drug-drug	290,836
Proteins	1,025	drug-protein	6,744
Side effects	12,904	drug-side effect	382,041
Diseases	1,229	drug-disease	6,677
		drug-drug similarities	chemical similarities, therapeutic similarities, drugs' target sequence similarities, gene ontology (GO) biological process, GO cellular component, and GO molecular function

ically, we use the drug-drug associations network, drug-protein associations network, drug-side effect associations network, and six drug similarity networks obtained from different omics data to obtain drug features. The known drug-disease interactions are used for the final association prediction. The type of all drug similarity networks and the numbers of all drugs, proteins, side effects, diseases, and associations are listed in Table 2. Specific construction details can be found in the supplemental experimental procedures.

For the embedding features of disease obtained through text mining, we directly use the pre-trained parameters of the BioBERT model.[27] The text corpora used in the training process of the model includes: (1) English Wikipedia, (2) BooksCorpus, (3) PubMed Abstracts, and (4) PMC Full-text articles.

Methods

Obtain drug features based on a heterogeneous network
To better utilize nine multiple drug-related associations networks to capture the features information of the drug, we first obtain the positive pointwise mutual information (PPMI) matrix of the drugs in each association network. Then, we use the SNF to fuse these PPMI matrices obtained from multiple networks. Finally, we use SAE to obtain a high-quality representation of the drug features (Figure 1A).

We randomly sort the vertices in a given drug-related associations network. As to the ith vertex, assume that there is a transition matrix T that obtains the transition probabilities between vertices. p_r is denoted as a row of vectors, and the j-th entry represents the probability of reaching the jth vertex after transferring r steps. The p_0 is the initial one-hot vector with the ith vertex value is 1, and all other values are 0. We can learn the probability of transitions between vertices iteratively. In each iteration process, the random surfing process will continue

with probability of α, and there is a $1 - \alpha$ probability to return to the original vertex and restart this process. The recurrence relation could be described as follows:

$$p_r = \alpha p_{r-1} T + (1 - \alpha) p_0. \qquad \text{(Equation 1)}$$

By performing the above process on each node, a probabilistic co-occurrence (PCO) matrix can be obtained. After yielding the PCO matrix, we calculate a shifted PPMI matrix.[59] The PPMI matrix-specific calculation method is as follows:

$$PPMI = \max \left(\log \frac{P(i,j) \times \sum_i^{N_d} \sum_j^{N_P} P(i,j)}{\sum_i^{N_d} P(i,j) \times \sum_j^{N_P} P(i,j)}, 0 \right), \qquad \text{(Equation 2)}$$

where P represents the original PCO matrix, and N_d and N_P represent the serial numbers of rows and columns, respectively. The process occurs as a pre-processing and the random surfing-based representation is mitigating the sparsity of some individual network types.

For the PPMI matrix obtained by each network, a computational method is needed to better integrate these matrix data to establish a comprehensive view of a given drug. We denote the correlation score between drug i and drug j in the PPMI matrix as $M(i,j)$ and use the SNF[24] to obtain a comprehensive biological view of a given drug. For the fused matrix from multiple of matrices, a full and sparse kernel is defined on the vertex set V. The following formula is used to normalize the full kernel:

$$F(i,j) = \begin{cases} \dfrac{M(i,j)}{2\sum\limits_{k \neq i} M(i,j)}, j \neq i \\[4mm] \dfrac{1}{2}, j = i \end{cases} . \qquad \text{(Equation 3)}$$

The K nearest neighbors is utilized to measure local affinity in a given network, as follows:

$$S(i,j) = \begin{cases} \dfrac{M(i,j)}{\sum\limits_{k \in N_i} M(i,k)}, j \in N_i \\[4mm] 0, otherwise \end{cases} , \qquad \text{(Equation 4)}$$

where N_i is denoted as a set of i's neighbors, including i in network. Where the F contains the complete information about the similarity of each drug to all other drugs, and S only encodes the similarity to the k most similar drugs for each drug. In the SNF process, F is always used as the initial state and S is used as the kernel matrix.

The F and S of each PPMI matrix are obtained by the above two formulas. Let us first take the fusion of the previous two PPMI matrices as a case, we get the status F_1 and F_2 from two PPMI matrices by Equation 3, and obtain the kernel matrices S_1 and S_2 by Equation 4. When t = 0, we denote the initial two status matrices $F_1^{t=0} = F_1$ and $F_2^{t=0} = F_2$. Iteratively updating the similarity matrix corresponding to each network data by using the following formulas is a key step of SNF:

$$F_1^{t+1} = S_1 \times F_2^t \times (S_1)^T, \qquad \text{(Equation 5)}$$

$$F_2^{t+1} = S_2 \times F_1^t \times (S_2)^T. \qquad \text{(Equation 6)}$$

After t iterations, F_1^{t+1} is the status matrix of F_1, F_2^{t+1} is the similarity matrix of F_2. The status matrices are updated in this procedure with generating two parallel interchanging diffusion processes each time. After t steps, the two matrices fused into a matrix F_f is computed as follows:

$$F_f = \frac{F_1^t + F_2^t}{2}. \qquad \text{(Equation 7)}$$

As shown in the framework Figure 1A, we sequentially merge our multiple PPMI matrices pairwise. Based on message passing theory, a non-linear method is used in the process of network fusion to iteratively update each network to make it more similar to other networks in each iteration. Through several iterations, SNF converges multiple networks into a single network. This method of merging multiple network data can reduce the noise of the network and preserve the strongly associated edges in the network.

After fusing multiple drug-related networks, we use the SAE[26] method to obtain high-quality and low-dimensional drug features. The encoder function from the input layer to the hidden layer of the SAE is:

$$H = \sigma(W_{encode} X_i + B_{encode}). \qquad \text{(Equation 8)}$$

The decoder function from the hidden layer to the output layer is:

$$Y = \sigma(W_{decode} H + B_{decode}), \qquad \text{(Equation 9)}$$

where $\sigma(x) = 1/(1 + e^{-x})$ is the activation function, W is the connection parameter, and B is the bias vector. The sparse penalty is added to the target function of the autoencoder to capture the effective features of the drug. The activation of the jth hidden unit is denoted as $h_j(x)$, we use the following formula to get the average activation amount of the jth hidden unit:

$$\widehat{\rho_J} = \frac{1}{m} \sum_{i=1}^{m} h_j(X_i). \qquad \text{(Equation 10)}$$

The loss function of SAE with sparse penalty is as follow:

$$L_{sparse}(W,B) = L(W,B) + \gamma \sum_{t=1}^{N_2} KL(\rho || \widehat{\rho_J}), \qquad \text{(Equation 11)}$$

$$KL(\rho || \widehat{\rho_J}) = \rho \log \frac{\rho}{\widehat{\rho_J}} + (1 - \rho) \log \frac{1 - \rho}{1 - \widehat{\rho_J}}, \qquad \text{(Equation 12)}$$

where $L(W,B)$ is the loss function of the neural network, γ is the hyperparameter to control the weight of sparsity in loss function, N_2 is the number of hidden layer units and ρ is a very small value closed to 0 as a sparsity parameter. The $KL(\rho || \widehat{\rho_J})$ is called KL-divergence, which possesses the property that $KL(\rho || \widehat{\rho_J}) = 0$ if $\widehat{\rho_J} = \rho$. Otherwise, it increases monotonically as $\widehat{\rho_J}$ diverges from ρ, which acts as the sparsity constraint. We adopt the gradient descent algorithm to minimize the $L_{sparse}(W,B)$ to optimize the parameters W and B.

Obtaining disease features based on text mining

For the disease features acquisition module, we considered, at the early stage of model design, that drugs that could act on a certain disease and may have an effect on related or similar diseases. Therefore, when we obtain disease features, we hope to capture their possible associations through text mining. The proposed BERT model[60] has enabled a qualitative leap for the text mining algorithm and brought a milestone change in the natural language-processing field, which uses transformer as the main framework and is pre-trained on BooksCorpus and English Wikipedia. However, the word distribution between general corpus and biomedical corpus is different, which is why the pure BERT model cannot achieve good results in biomedical text mining tasks.

In our work, to obtain effective disease features representation through text mining, we used the BioBERT model,[27] which is a biomedical domain-based pre-trained language-representation model. For tokenization, the BioBERT utilizes the WordPiece tokenization.[61] With word tokenization, any disease name words can be represented by frequently occurring sub-words. Given the limitation of training costs, we directly use the pre-trained parameters of the BioBERT model and select the relation extraction task in the three representative biomedical text mining tasks for fine-tuning training. The BioBERT used the original BERT sentence classifier. In the fine-tuning process, diseases and genes are anonymized target entities in the datasets to prevent the supervised information used in the fine-tuning process from overlapping with drug-disease relations in the test set and to avoid the possibility that the test data may be contaminated by the text corpora. For the fine-tuning, we select a batch size of 32 and a learning rate of $2e - 5$. After the fine-tuning process done, we extract the representations of sub-words and use the representations of these sub-words to obtain the representations of all diseases (Figure 1B). After obtaining the diseases features, we verified their effectiveness of the diseases features. More details for evaluating the effectiveness of diseases features are available in the supplemental experimental procedures.

Drug-disease associations prediction

The features obtained from the first two steps can be used as the attribute information of drugs and diseases. In the process of embedding learning, we need to preserve node attribute proximity and the drug-disease associations network topological structure. We refer to the GATNE-I model[62] to predict the drug-disease interactions Figure 1C. We divide the overall embedding of a given node v_i into three parts: node attributes embedding, neighborhood aggregation embedding, and base embedding.

To better integrate the attribute information of heterogeneous network nodes, the drug and disease features obtained in the above two steps are defined as the attribute X_i of nodes. We define the base embedding $b_t(X_i)$, which is a parameterized function of v_i's attribute X_i, b_t is a transformation function and $t = \varnothing(i)$ is node v_i's corresponding node type. For neighborhood aggregation embedding, the kth level embedding $n_i^k \in \mathbb{R}^s$, ($1 \leq k \leq K$) of node v_i is aggregated from neighbor's embedding, and s is the dimension of neighborhood aggregation embedding. We compute the neighborhood aggregation embedding by the mean aggregator function[63] as:

$$n_i^{(k)} = \sigma\left(\widehat{W}^{(k)} \cdot MEAN\left(\left\{n_j^{(k-1)}, \forall v_j \in \mathcal{N}_i\right\}\right)\right), \quad \text{(Equation 13)}$$

where σ is an activation function, $\widehat{W}^{(k)}$ is the weight matrices to propagate information between different layers, and $\mathcal{N}_i$ are the neighbors of node v_i. The initial neighborhood aggregation embeddings $n_i^{(0)}$ is also a parameterized function of v_i's attribute X_i as $n_i^{(0)} = g_t(X_i)$. The subscript t is v_i's corresponding node type and g_t is a transformation function.

Finally, the overall embedding of node v_i's function is as follows:

$$v_i = b_t(X_i) + \alpha M^T n_i R_i + \beta D_t^T X_i, \quad \text{(Equation 14)}$$

$$R_i = softmax\left(w^T \tanh(W n_i)\right)^T, \quad \text{(Equation 15)}$$

where α is a hyperparameter, $M \in \mathbb{R}^{s \times d}$ is a trainable transformation matrix, d is the dimension of overall embedding, w is trainable parameter with size d_a, W is trainable parameter with size $d_a \times s$, β is a coefficient, the symbol T represents the transposition of the matrix or the vector, and D_t is a feature transformation matrix on v_i's corresponding node type t.

Next, optimize the model. Because the drug-disease associations network is a heterogeneous network, to ensure the semantic relationship between different types of nodes can be correctly merged into the skip-gram model, meta-path-based random walks are used to obtain node sequence and skip-gram are performed over the node sequence to learn embeddings.[64] Specifically, for a given network $G = (V, E, T)$ and a meta-path scheme $\mathcal{L} : V_1 \rightarrow V_2 \cdots V_t \cdots \rightarrow V_l$, the flow of the walker is conditioned on the predefined mate-path $\mathcal{L}$ and the transition probability at step t is define as follows:

$$P(v_i|v_i, \mathcal{L}) = \begin{cases} \frac{1}{|\mathcal{N}_i \cap V_{t+1}|} & (v_i, v_j) \in E, \ v_j \in V_{t+1} \\ 0 & (v_i, v_j) \in E, \ v_j \notin V_{t+1} \\ 0 & (v_i, v_j) \notin E \end{cases}, \quad \text{(Equation 16)}$$

where $v_i \in V_t$. Supposing the random walk with length l follows a path $P = (v_1, v_2, \cdots, v_l)$ such that $(v_{t-1}, v_t) \in E(t = 2, 3, \cdots l)$, v_t's context is denoted as $C = \{v_k | v_k \in P, |k - t| \leq c, t \neq k\}$, where w is the radius of the window size.

Therefore, given a node v_i with its context C of a path, we aim to minimize the following negative log likelihood:

$$-\log P_\varepsilon\left(\{v_j|v_j \varepsilon C\}|v_i\right) = \sum_{v_j \in C} -\log P_\varepsilon(v_j|v_i), \quad \text{(Equation 17)}$$

where ε denotes all the parameters, and use the heterogeneous softmax function normalized with respect to the node type of node v_j, the probability of v_j given v_i is defined as:

$$P_\varepsilon(v_j|v_i) = \frac{\exp\left(c_j^T \cdot v_i\right)}{\sum_{k \in V_t} \exp(c_k^T \cdot v_i)}, \quad \text{(Equation 18)}$$

where $v_j \in V_t$, c_k is the context embedding of node v_k and v_i is the overall embedding of node v_i.

Finally, heterogeneous negative sampling is used to approximate the objective function $-\log P_\varepsilon(v_j|v_i)$ for each node pair (v_i, v_j) as:

$$H = -\log \sigma\left(c_j^T \cdot v_i\right) - \sum_{l=1}^{L} \mathbb{E}_{v_k \sim P_t(v)}\left[\log \sigma\left(-c_k^T \cdot v_i\right)\right], \quad \text{(Equation 19)}$$

where σ is the sigmoid function, L is the number of negative samples corresponding to a positive training sample, and v_k is randomly drawn from a noise distribution $P_t(v)$ defined on node v_j's corresponding node set V_t.

AUTHOR CONTRIBUTIONS

Conceptualization, X.Z., X.L., and S.J.; methodology, X.Z., Z.N., and S.J.; investigation, X.L., X.Z., S.J., Z.N., C.J., and W.H.; writing – original draft, S.J., C.J., and F.X.; writing – review & editing, X.Z., S.J., X.J., and W.H.; funding acquisition, X.L. and X.Z.; resources, S.J., X.L., and X.Z.; supervision, X.L. and X.Z.

REFERENCES

1. Dickson, M., and Gagnon, J.P. (2004). Key factors in the rising cost of new drug discovery and development. Nat. Rev. Drug Discov. *3*, 417–429. https://doi.org/10.1038/nrd1382.

2. Pushpakom, S., Iorio, F., Eyers, P.A., Escott, K.J., Hopper, S., Wells, A., Doig, A., Guilliams, T., Latimer, J., McNamee, C., et al. (2019). Drug repurposing: progress, challenges and recommendations. Nat. Rev. Drug Discov. *18*, 41–58. https://doi.org/10.1038/nrd.2018.168.

3. Li, J., Zheng, S., Chen, B., Butte, A.J., Swamidass, S.J., and Lu, Z. (2016). A survey of current trends in computational drug repositioning. Brief. Bioinformatics *17*, 2–12. https://doi.org/10.1093/bib/bbv020.

4. Ashburn, T.T., and Thor, K.B. (2014). Drug repositioning: identifying and developing new uses for existing drugs. Nat. Rev. Drug Discov. *3*, 673–683. https://doi.org/10.1038/nrd1468.

5. Napolitano, F., Zhao, Y., Moreira, V.M., Tagliaferri, R., Kere, J., D'Amato, M., and Greco, D. (2013). Drug repositioning: a machine-learning approach through data integration. J. Cheminformatics *5*, 30. https://doi.org/10.1186/1758-2946-5-30.

6. Wang, Y., Chen, S., Deng, N., and Wang, Y. (2013). Drug repositioning by kernel-based integration of molecular structure, molecular activity, and phenotype data. PLoS One *8*, e78518. https://doi.org/10.1371/journal.pone.0078518.

7. Kim, E., Choi, A.-s., and Nam, H. (2019). Drug repositioning of herbal compounds via a machine-learning approach. BMC Bioinformatics *20*, 33–43. https://doi.org/10.1186/s12859-019-2811-8.

8. Gottlieb, A., Stein, G.Y., Ruppin, E., and Sharan, R. (2011). PREDICT: a method for inferring novel drug indications with application to personalized medicine. Mol. Syst. Biol. *7*, 496. https://doi.org/10.1038/msb.2011.26.

9. Martinez, V., Navarro, C., Cano, C., Fajardo, W., and Blanco, A. (2015). DrugNet: network-based drug–disease prioritization by integrating heterogeneous data. Artif. Intellig. Med. *63*, 41–49. https://doi.org/10.1016/j.artmed.2014.11.003.

10. Zeng, X., Zhu, S., Liu, X., Zhou, Y., Nussinov, R., and Cheng, F. (2019). deepDR: a network-based deep learning approach to in silico drug repositioning. Bioinformatics *35*, 5191–5198. https://doi.org/10.1093/bioinformatics/btz418.

11. Lu, L., and Yu, H. (2018). DR2DI: a powerful computational tool for predicting novel drug-disease associations. J. Comput. Aided Mol. Des. *32*, 633–642. https://doi.org/10.1007/s10822-018-0117-y.

12. Liu, H., Zhang, W., Song, Y., Deng, L., and Zhou, S. (2020). HNet-DNN: inferring new drug-disease associations with deep neural network based on heterogeneous network features. J. Chem. Inf. Model. *60*, 2367–2376. https://doi.org/10.1021/acs.jcim.9b01008.

13. Liu, H., Song, Y., Guan, J., Luo, L., and Zhuang, Z. (2016). Inferring new indications for approved drugs via random walk on drug-disease heterogenous networks. BMC Bioinformatics *17*, 269–277. https://doi.org/10.1186/s12859-016-1336-7.

14. Luo, H., Wang, J., Li, M., Luo, J., Peng, X., Wu, F.X., and Pan, Y. (2016). Drug repositioning based on comprehensive similarity measures and bi-random walk algorithm. Bioinformatics *32*, 2664–2671.

15. Dai, W., Liu, X., Gao, Y., Chen, L., Song, J., Chen, D., Gao, K., Jiang, Y., Yang, Y., and Chen, J. (2015). Matrix factorization-based prediction of novel drug indications by integrating genomic space. Comput. Math. Methods Med. *2015*. https://doi.org/10.1155/2015/275045.

16. Ceddia, G., Pinoli, P., Ceri, S., and Masseroli, M. (2019). Non-negative matrix tri-factorization for data integration and network-based drug repositioning. In 2019 IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB) (IEEE), pp. 1–7. https://doi.org/10.1109/CIBCB.2019.8791474.

17. Brown, A.S., and Patel, C.J. (2017). MeSHDD: literature-based drug-drug similarity for drug repositioning. J. Am. Med. Inform. Assoc. *24*, 614–618. https://doi.org/10.1093/jamia/ocw142.

18. Yang, H.-T., Ju, J.H., Wong, Y.T., Shmulevich, I., and Chiang, J.H. (2017). Literature-based discovery of new candidates for drug repurposing. Brief. Bioinformatics *18*, 488–497. https://doi.org/10.1093/bib/bbw030.

19. Li, F., Zhang, M., Fu, G., and Ji, D. (2017). A neural joint model for entity and relation extraction from biomedical text. BMC Bioinformatics *18*, 1–11. https://doi.org/10.1186/s12859-017-1609-9.

20. Wang, P., Hao, T., Yan, J., and Jin, L. (2017). Large-scale extraction of drug-disease pairs from the medical literature. J. Assoc. Inf. Sci. Technol. *68*, 2649–2661. https://doi.org/10.1002/asi.23876.

21. Yang, C., Xiao, Y., Zhang, Y., Sun, Y., and Han, J. (2020). Heterogeneous network representation learning: a unified framework with survey and benchmark. IEEE Trans. Knowledge Data Eng. https://doi.org/10.1109/TKDE.2020.3045924.

22. Hu, Z., Dong, Y., Wang, K., and Sun, Y. (2020). Heterogeneous graph transformer. In Proceedings of the Web Conference 2020. https://doi.org/10.1145/3366423.3380027.

23. Antons, D., Grünwald, E., Cichy, P., and Salge, T.O. (2020). The application of text mining methods in innovation research: current state, evolution patterns, and development priorities. R&D Manag. *50*, 329–351. https://doi.org/10.1111/radm.12408.

24. Wang, B., Mezlini, A.M., Demir, F., Fiume, M., Tu, Z., Brudno, M., Haibe-Kains, B., and Goldenberg, A. (2014). Similarity network fusion for aggregating data types on a genomic scale. Nat. Methods *11*, 333. https://doi.org/10.1038/nmeth.2810.

25. Ng, A. (2011). Sparse autoencoder. CS294A Lecture Notes *72*, 1–19.

26. Jiang, H.-J., Huang, Y.-A., and You, Z.-H. (2020). SAEROF: an ensemble approach for large-scale drug-disease association prediction by incorporating rotation forest and sparse autoencoder deep neural network. Sci. Rep. *10*, 1–11. https://doi.org/10.1038/s41598-020-61616-9.

27. Lee, J., Yoon, W., Kim, S., Kim, D., Kim, S., So, C.H., and Kang, J. (2020). BioBERT: a pre-trained biomedical language representation model for biomedical text mining. Bioinformatics *36*, 1234–1240. https://doi.org/10.1093/bioinformatics/btz682.

28. Moghadam, H., Rahgozar, M., and Gharaghani, S. (2016). Scoring multiple features to predict drug disease associations using information fusion and aggregation. SAR QSAR Environ. Res. *27*, 609–628. https://doi.org/10.1080/1062936X.2016.1209241.

29. Cortes, C., and Vapnik, V. (1995). Support-vector networks. Machine Learn. *20*, 273–297. https://doi.org/10.1007/BF00994018.

30. Chen, X., Huang, Y.A., You, Z.H., Yan, G.Y., and Wang, X.S. (2017). A novel approach based on KATZ measure to predict associations of human microbiota with non-infectious diseases. Bioinformatics *33*, 733–739. https://doi.org/10.1093/bioinformatics/btw715.

31. Luo, Y., Zhao, X., Zhou, J., Yang, J., Zhang, Y., Kuang, W., Peng, J., Chen, L., and Zeng, J. (2017). A network integration approach for drug-target interaction prediction and computational drug repositioning from heterogeneous information. Nat. Commun. *8*, 1–13. https://doi.org/10.1038/s41467-017-00680-8.

32. Luo, H., Li, M., Wang, S., Liu, Q., Li, Y., and Wang, J. (2018). Computational drug repositioning using low-rank matrix approximation and randomized algorithms. Bioinformatics *34*, 1904–1912. https://doi.org/10.1093/bioinformatics/bty013.

33. Aljenedil, S., Alothman, L., Bélanger, A.M., Brown, L., Lahijanian, Z., Bergeron, J., Couture, P., Baass, A., Ruel, I., and Brisson, D. (2020). Lomitapide for treatment of homozygous familial hypercholesterolemia: the Québec experience. Atherosclerosis *310*, 54–63. https://doi.org/10.1016/j.atherosclerosis.2020.07.028.

34. Bélanger, A.M., Akioyamen, L., Alothman, L., and Genest, J. (2020). Evidence for improved survival with treatment of homozygous familial hypercholesterolemia. Curr. Opin. Lipidol. *31*, 176–181. https://doi.org/10.1097/MOL.0000000000000686.

35. Reeskamp, L.F., Kastelein, J.J., Moriarty, P.M., Duell, P.B., Catapano, A.L., Santos, R.D., and Ballantyne, C.M. (2019). Safety and efficacy of mipomersen in patients with heterozygous familial hypercholesterolemia. Atherosclerosis *280*, 109–117. https://doi.org/10.1016/j.atherosclerosis.2018.11.017.

36. Jenike, M.A., Albert, M.S., Heller, H., LoCastro, S., and Gunther, J. (1986). Combination therapy with lecithin and ergoloid mesylates for Alzheimer's disease. J. Clin. Psychiatry *47*, 249–251.

37. Flynn, B.L., and Ranno, A.E. (1999). Pharmacologic management of Alzheimer disease part II: antioxidants, antihypertensives, and ergoloid derivatives. Ann. Pharmacother. *33*, 188–197. https://doi.org/10.1345/aph.17172.

38. Cover, C.C., Poulin, J.E., Gustafson, M.R., Wyant, T., Gamble, D.N., and Kay, M.M. (1996). Posttranslational changes in band 3 in adult and aging brain following treatment with ergoloid mesylates, comparison to changes observed in Alzheimer's disease. Life Sci. *58*, 655–664. https://doi.org/10.1016/S0024-3205(96)80004-X.

39. Singer, J.M., Hamot, H.B., and Patin, J.R. (1985). Differential patient response to ergoloid mesylates according to current etiopathic notions of dementia. Aging 2000: Our Health Care Destiny (Springer), pp. 405–420. https://doi.org/10.1007/978-1-4612-5058-6_33.

40. Gu, S., Fu, W.Y., Fu, A.K., Tong, E.P.S., Ip, F.C., Huang, X., and Ip, N.Y. (2018). Identification of new EphA4 inhibitors by virtual screening of FDA-approved drugs. Sci. Rep. *8*, 1–7. https://doi.org/10.1038/s41598-018-25790-1.

41. Nizami, S., Hall-Roberts, H., Warrier, S., Cowley, S.A., and Di Daniel, E. (2019). Microglial inflammation and phagocytosis in Alzheimer's disease: potential therapeutic targets. Br. J. Pharmacol. *176*, 3515–3532. https://doi.org/10.1111/bph.14618.

42. Dansokho, C., and Heneka, M.T. (2018). Neuroinflammatory responses in Alzheimer's disease. J. Neural Transm. *125*, 771–779. https://doi.org/10.1007/s00702-017-1831-7.

43. Sheridan, C. (2017). Novartis trial validates inflammasome as chronic disease driver. Nat. Biotechnol. *35*, 893–894. https://doi.org/10.1038/nbt1017-893.

44. Mitroulis, I., Skendros, P., and Ritis, K. (2010). Targeting IL-1β in disease; the expanding role of NLRP3 inflammasome. Eur. J. Intern. Med. *21*, 157–163. https://doi.org/10.1016/j.ejim.2010.03.005.

45. Chauhan, D., Vande Walle, L., and Lamkanfi, M. (2020). Therapeutic modulation of inflammasome pathways. Immunological Rev. *297*, 123–138. https://doi.org/10.1111/imr.12908.

46. McElroy, S.L., Guerdjikova, A., Kotwal, R., Welge, J.A., Nelson, E.B., Lake, K.A., Keck, P.E., Jr., and Hudson, J.I. (2007). Atomoxetine in the treatment of binge-eating disorder: a randomized placebo-controlled trial. J. Clin. Psychiatry *68*, 390–398.

47. Gadde, K.M., Yonish, G.M., Wagner, H.R., Foust, M.S., and Allison, D.B. (2006). Atomoxetine for weight reduction in obese women: a preliminary randomised controlled trial. Int. J. Obes. *30*, 1138–1142. https://doi.org/10.1038/sj.ijo.0803223.

48. Pott, W., Albayrak, Ö., Hinney, A., Hebebrand, J., and Pauli-Pott, U. (2013). Successful treatment with atomoxetine of an adolescent boy with attention deficit/hyperactivity disorder, extreme obesity, and reduced melanocortin 4 receptor function. Obes. Facts *6*, 109–115. https://doi.org/10.1159/000348792.

49. Mirbolooki, M.R., Constantinescu, C.C., Pan, M.L., and Mukherjee, J. (2013). Targeting presynaptic norepinephrine transporter in brown adipose tissue: a novel imaging approach and potential treatment for diabetes and obesity. Synapse *67*, 79–93. https://doi.org/10.1002/syn.21617.

50. Shibao, C., Raj, S.R., Gamboa, A., Diedrich, A., Choi, L., Black, B.K., Robertson, D., and Biaggioni, I. (2007). Norepinephrine transporter blockade with atomoxetine induces hypertension in patients with impaired autonomic function. Hypertension *50*, 47–53. https://doi.org/10.1161/HYPERTENSIONAHA.107.089961.

51. Spencer, T.J., Kratochvil, C.J., Sangal, R.B., Saylor, K.E., Bailey, C.E., Dunn, D.W., Geller, D.A., Casat, C.D., Lipetz, R.S., and Jain, R. (2007). Effects of atomoxetine on growth in children with attention-deficit/hyperactivity disorder following up to five years of treatment. J. Child Adolesc. Psychopharmacol. *17*, 689–699. https://doi.org/10.1089/cap.2006.0100.

52. Yuan, C.-S., Wang, C.Z., Attele, A., and Zhang, L. (2009). Methylnaltrexone reduced body weight gain in ob/ob mice. J. Opioid Manag. *5*, 213–218. https://doi.org/10.5055/jom.2009.0023.

53. Gatti, A., and Sabato, A.F. (2012). Management of opioid-induced constipation in cancer patients. Clin. Drug Invest. *32*, 293–301. https://doi.org/10.2165/11598000-000000000-00000.

54. Biton, V., Gates, J.R., Ritter, F.J., and Loewenson, R.B. (1990). Adjunctive therapy for intractable epilepsy with ethotoin. Epilepsia *31*, 433–437. https://doi.org/10.1111/j.1528-1157.1990.tb05499.x.

55. Ashna, A., Van Helden, D., and Laver, D. (2018). Phenytoin and ethotoin inhibit ryanodine receptor in manner paralleling that of dantrolene. Heart, Lung, Circ. *27*, S182. https://doi.org/10.1016/j.hlc.2018.06.313.

56. Iwabuchi, M. (2019). SAT-473 acromegaly and drug-induced parkinsonism were controlled by lanreotide. J. Endocr. Soc. *3* (*Suppl. 1*). SAT-473. https://doi.org/10.1210/js.2019-SAT-473.

57. Abe, K., Chiba, Y., Katsuse, O., and Hirayasu, Y. (2016). A case of Parkinson disease with both visual hallucination and pain improved by gabapentin. Clin. Neuropharmacol. *39*, 55–56. https://doi.org/10.1097/WNF.0000000000000122.

58. Fujishiro, H. (2014). Effects of gabapentin enacarbil on restless legs syndrome and leg pain in dementia with Lewy bodies. Psychogeriatrics *14*, 132–134. https://doi.org/10.1111/psyg.12043.

59. Bullinaria, J.A., and Levy, J.P. (2007). Extracting semantic representations from word co-occurrence statistics: a computational study. Behav. Res. Methods *39*, 510–526. https://doi.org/10.3758/BF03193020.

60. Devlin, J., Chang, M.W., Lee, K., and Toutanova, K. (2018). BERT: pre-training of deep bidirectional transformers for language understanding. arXiv, 1810.04805.

61. Wu, Y., Schuster, M., Chen, Z., Le, Q.V., Norouzi, M., Macherey, W., Krikun, M., Cao, Y., Gao, Q., and Macherey, K. (2016). Google's neural machine translation system: bridging the gap between human and machine translation. arXiv, 1609.08144.

62. Cen, Y., Zou, X., Zhang, J., Yang, H., Zhou, J., and Tang, J. (2019). Representation learning for attributed multiplex heterogeneous network. In Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining., pp. 1358–1368. https://doi.org/10.1145/3292500.3330964.

63. Hamilton, W., Ying, Z., and Leskovec, J. (2017). Inductive representation learning on large graphs. arXiv, 1706.02216.

64. Dong, Y., Chawla, N.V., and Swami, A. (2017). metapath2vec: scalable representation learning for heterogeneous networks. In Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining., pp. 135–144. https://doi.org/10.1145/3097983.3098036.