In Silico Off-Target Profiling for Enhanced Drug Safety Assessment

Authors: Jin Liu, Yike Gui, Jingxin Rao, Jingjing Sun, Gang Wang, Qun Ren, Ning Qu, Buying Niu, Zhiyi Chen, Xia Sheng, Yitian Wang, Mingyue Zheng, Xutong Li, Mingyue Zheng, Xutong Li

Date: 2024-02-20T00:00:00+00:00

Abstract

Ensuring drug safety in the early stages of drug development is crucial to avoid costly failures in subsequent phases. However, the economic burden associated with detecting drug off-targets and potential side effects through in vitro safety screening and animal testing is substantial. Drug off-target interactions, along with the adverse drug reactions they induce, are significant factors affecting drug safety. To assess the liability of candidate drugs, we developed an artificial intelligence model for the precise prediction of compound off-target interactions, leveraging multi-task graph neural networks. The outcomes of off-target predictions can serve as representations for compounds, enabling the differentiation of drugs under various ATC codes and the classification of compound toxicity. Furthermore, the predicted off-target profiles are employed in ADR enrichment analysis, facilitating the inference of potential ADRs for a drug. Using the withdrawn drug Pergolide as an example, we elucidate the mechanisms underlying ADRs at the target level, contributing to the exploration of the potential clinical relevance of newly predicted off-target interactions. Overall, our work facilitates the early assessment of compound safety/toxicity based on off-target identification, deduces potential ADRs of drugs, and ultimately promotes the secure development of drugs.

Full Text

Preamble

In Silico Off-Target Profiling for Enhanced Drug Safety Assessment

Jin Liu^{a,b,†}, Yike Gui^{d,b,†}, Jingxin Rao^{b,c}, Jingjing Sun^{b,c}, Gang Wang^{b,c}, Qun Ren^{d,b}, Ning Qu^{b,c}, Buying Niu^{b,c},

Zhiyi Chen^{e,b,c}, Xia Sheng^{b,c}, Yitian Wang^{b,c}, Mingyue Zheng^{a,b,c,d,e,}, *Xutong Li^{b,c,}*

^{a}College of Pharmaceutical Sciences, Zhejiang University, Hangzhou 310058, China
^{b}Drug Discovery and Design Center, State Key Laboratory of Drug Research, Shanghai Institute of Materia Medica, Chinese Academy of Sciences, 555 Zuchongzhi Road, Shanghai 201203, China
^{c}University of Chinese Academy of Sciences, No.19A Yuquan Road, Beijing 100049, China
^{d}Nanjing University of Chinese Medicine, 138 Xianlin Road, Nanjing 210023, China
^{e}School of Pharmaceutical Science and Technology, Hangzhou Institute for Advanced Study, Hangzhou 330106, China

^{†}These authors contributed equally to this work.

^{*}Corresponding authors. Tel.: +86-021-50271399; Tel.: +86-13821547729.

E-mail addresses: myzheng@simm.ac.cn (Mingyue Zheng), lixutong@simm.ac.cn (Xutong Li).

Acknowledgments

This work was supported by National Key Research and Development Program of China (2022YFC3400504 to MYZ), National Natural Science Foundation of China (T2225002 and 82273855 to MYZ, 82204278 to XTL), Lingang Laboratory (LG202102-01-02 to MYZ), SIMM-SHUTCM Traditional Chinese Medicine Innovation Joint Research Program (E2G805H to MYZ), and Shanghai Municipal Science and Technology Major Project.

Author contributions: XTL and MYZ conceived the project. JL implemented the off-target prediction model and conducted the computational analysis. YKG, JXR, JJS, GW, QR, NQ, BYN, ZYC, XS and YTW collected and analyzed the data. JL, YKG and JXR wrote the paper. All authors discussed the results and commented on the manuscript.

Conflicts of interest: The authors declare no competing interests.

Abstract

Ensuring drug safety in the early stages of drug development is crucial to avoid costly failures in subsequent phases. However, the economic burden associated with detecting drug off-targets and potential side effects through in vitro safety screening and animal testing is substantial. Drug off-target interactions, along with the adverse drug reactions they induce, are significant factors affecting drug safety. To assess the liability of candidate drugs, we developed an artificial intelligence model for the precise prediction of compound off-target interactions,

leveraging multi-task graph neural networks. The outcomes of off-target predictions can serve as representations for compounds, enabling the differentiation of drugs under various ATC codes and the classification of compound toxicity. Furthermore, the predicted off-target profiles are employed in ADR enrichment analysis, facilitating the inference of potential ADRs for a drug. Using the withdrawn drug Pergolide as an example, we elucidate the mechanisms underlying ADRs at the target level, contributing to the exploration of the potential clinical relevance of newly predicted off-target interactions. Overall, our work facilitates the early assessment of compound safety/toxicity based on off-target identification, deduces potential ADRs of drugs, and ultimately promotes the secure development of drugs.

KEY WORDS: Drug safety; off-target prediction; adverse drug reactions; toxicity; molecular representation; artificial intelligence

Running title: In Silico Off-Target Profiling for Enhanced Drug Safety Assessment

1. Introduction

Ensuring drug safety during the early stages of drug development is of paramount importance, as it not only safeguards patient well-being but also contributes to the overall success and viability of pharmaceutical endeavors [1-3]. Traditional approaches to safety evaluation and toxicity prediction for compounds have relied on costly in vitro methods (e.g., organ-on-a-chip) and in vivo methods (e.g., animal models) that may not accurately reflect human responses [4]. Notably, off-target toxicity is a significant contributor to drug attrition [5-7], highlighting the need to identify undesired drug targets that could lead to adverse drug reactions (ADRs) in humans [1, 8]. This identification process presents a relatively low-cost approach to evaluating drug safety.

Pharmaceutical companies commonly employ in vitro pharmacological assays to profile compounds against a comprehensive panel of unsafe off-targets to reduce the number of molecules tested in subsequent assays [1]. Based on the internal off-target panels of four pharmaceutical companies—AstraZeneca, GlaxoSmithKline, Novartis, and Pfizer—Bowes et al. proposed 44 early drug safety targets that include the toxicity of the central nervous system, immune system, gastrointestinal tract, and heart [1]. AbbVie obtained 70 safety-related targets via a literature search, most of which are included in Eurofins' safety panel [9]. Roche utilized experimental data based on the Bioprint® database and employed a statistical ranking method, resulting in a panel of 50 safety targets [10]. However, compared to screening compounds on known therapeutic targets, off-target screening of compounds is challenging due to the lack of boundaries. Conducting extensive experimental screening on numerous targets can be cost-prohibitive.

Therefore, employing in-silico predictions to assess compound-target interac-

tions provides a cost-effective approach to investigating off-target compound safety [10, 11]. The traditional target prediction methods rely on chemical similarity search, where many studies use multi-target SAR (Structure-Activity Relationship) models to retrieve putative targets of compounds. With the advent of the big data era, integrating artificial intelligence (AI) methods into this process offers further cost reduction opportunities. Mayr et al. [12] employed data collected from ChEMBL to construct a series of supervised binary classification models, such as Random Forest (RF), K-Nearest Neighbor (KNN), and Deep Neural Network. In a recent study by Roche, researchers proposed a suite of off-target prediction models, including Neural Networks, RF, Auto-Sklearn, AutoGluon, and H2O. These models were assessed using a dataset of 4,000 compounds from the company, enabling a thorough exploration and comparison of neural networks and machine learning methods in constructing off-target prediction models for 50 distinct targets, each with varying dataset sizes and imbalances [13]. Lunghini et al. [14] introduced ProfhEX, a platform that utilizes tree-based Gradient Boosting (GB) and RF algorithms to establish prediction models for 46 off-targets. The platform also provides a comprehensive mechanistically-driven liability profile of small molecules.

Previous research related to compound safety has predominantly concentrated on singular aspects. Apart from off-target prediction, ADR prediction and toxicity prediction are also commonly employed for evaluating compound safety [15]. Various machine learning algorithms have been applied to ADR prediction, leveraging features such as drug phenotype, chemical and biological information, and target proteins to model the complex drug-ADR relationships [16-18]. Liu et al. integrated phenotype, chemical and biological information of drugs and tested various classifiers such as Logistic Regression, Naive Bayes, KNN, RF and SVM for each ADR [19]. Zhang et al. modeled drugs and their side effects as a multi-label task, using various information such as chemical substructures, target proteins, and indications to represent drugs [20]. Zhang et al. constructed a KG containing four types of nodes (drug, indication, target, and side effect), and proposed a novel knowledge graph embedding method combined with a logistic regression classification model to predict whether a given drug has a certain ADR [21]. For compound toxicity prediction, computational toxicology, an emerging field, offers numerous models for large-scale virtual screening to identify candidates for subsequent experimental testing [22-24]. These models can be expert-designed, involving techniques like structural alerts [25, 26] or read-cross [27], or they can be created automatically using machine learning techniques. While expert-designed rules provide some guidance for toxicity prediction, they often exhibit excessive sensitivity, leading to numerous false-positive outcomes. Machine learning methods primarily rely on quantitative structure-activity relationship (QSAR) [28], which characterizes a drug' s chemical structure and combines it with relevant supervised learning algorithms such as RF, XGBoost, and SVM, among others, for toxicity modeling [5, 22, 29].

However, ADR and toxicity data are primarily derived from limited clinical sources, posing challenges to traditional prediction methods, especially those re-

lying on marketed and clinical compound structures, which could exhibit poor generalizability. Moreover, considering the inherent link between drug off-target effects and ADRs, as well as toxicity [3, 30], characterizing the compound' s off-targets becomes a critical determinant of its safety. In light of this, we propose predicting a drug' s off-target profile and utilizing it as a compound representation for subsequent tasks, including ATC classification, toxicity prediction, and ADR enrichment analysis. Initially, using a comprehensive compound-protein interaction database, we construct a multi-task graph neural network model to predict compounds' off-target profiles based on their chemical structures and compare it with several previous off-target prediction models to demonstrate its performance.

The off-target prediction results for any given molecule can then serve as a molecular representation, capturing the molecule' s off-target and subsequent ADR or toxicity effects. We explored the use of these representations in drug ATC classification and toxicity prediction, comparing their performance with that of ECFP-based models to showcase their effectiveness. Furthermore, ADR enrichment analysis is employed to leverage the off-target profile, identifying crucial ADRs at the target level, particularly severe ADRs leading to drug withdrawal. Using Pergolide as a case study, we predicted its off-target profile and subsequently utilized the off-target representation to elucidating its ADR mechanisms, attempting to provide the potential explanations to drug-target-ADR correlations of Pergolide. Hence, initiating from the molecular structures, we can obtain the molecules' off-target representations, which provide valuable information for safety-related prediction tasks. This early safety assessment protocol can steer a rational drug development process, facilitating the discovery of safe compounds.

2.1. Collection and processing of compound-target interaction datasets

As indicated in Supporting Information Text S1, our project created an off-target panel consisting of 90 targets. More information about these targets can be found in Supporting Information Table S1. According to the gene names of the targets, we followed the steps outlined in Supporting Information Text S2 to collect compounds associated with the corresponding targets from the ChEMBL [31] and PubChem [32] databases. The databases used in our study are presented in Supporting Information Table S2. We eliminated experiment indicators with insufficient data and mainly retained the following six indicators: Ki, Kd, IC50, EC50, %activity, and %inhibition. To classify compounds under a specific target as active or inactive, we applied the threshold settings introduced in the Illuminating the Druggable Genome (IDG) project [33, 34]. Supporting Information Table S3 provides details on the different threshold settings. To facilitate subsequent training needs, we merged the data obtained from ChEMBL and PubChem and only selected targets containing at least 10 positive compounds. In total, 242 multi-species targets were retained, including

90 human targets, screening against approximately 320,000 unique compounds. The statistic of the processed data is shown in Table 1 .

To expose the model to a larger number of negative samples and facilitate a more comprehensive understanding of the negative sample space [35], mitigate false positives and enhance the model' s usability, we employed a negative sampling approach to augment the negative samples. Under each target class, we selected compounds as negative sampled decoys based on their similarity to positive samples in terms of physical and chemical properties and molecular fingerprint similarity of less than 0.6. The sampling ratio differed among target classes to maintain a positive-to-negative compound ratio of approximately 1:5 in the data after negative sampling. This ensured that the model had enough exposure to negative samples. The data amounts after negative sampling are detailed in Table 1. Supporting Information Fig. S1 depicts the chemical spatial distribution of the sampled decoys and positive molecules under different target classes. The similarity in chemical spatial distribution indicates the reasonableness of the sampling approach. Employing these decoys, we aimed to demonstrate the model' s ability to distinguish between the two molecule types based on their structural characteristics, rather than solely considering their physical and chemical properties [36].

The compound-target interaction data division used in this project adopts the drug-blind mode, which does not consider the target and groups all small molecules together for classification. We performed random stratified partitioning with a ratio of 0.1 to obtain the test set first and then performed five-fold cross-validation on the remaining data to obtain train and validation datasets for each fold. We then trained and validated the model five times and evaluated the model by computing the average of the results obtained from the five experiments.

2.2. Construction of multi-task GNN models

Multitask learning is a machine learning approach that utilizes a shared representation, training multiple related tasks simultaneously [37, 38]. We employ a hard parameter sharing GNN model for multitask learning, in which the same underlying parameters are shared across all final tasks, while each model maintains independent top-level parameters [39].

A graph neural network, Attentive FP, is used for molecular representation, which leverages the graph attention mechanism to represent molecules and learn related tasks [40]. To reduce the bias caused by the imbalance label (negative samples are much more than positive samples), we used a weighted cross-entropy loss function, where the class weights were set as the inverse of class frequencies [41]. Early stopping and hyper-parameter search strategy were used for model optimization. The hyper-parameter search range of the multi-task GNN can be found in Supporting Information Table S4.

2.3. Execution of ADR enrichment analysis

In research conducted by Novartis [42], Jeffrey et al. [3] and other drug safety researchers [1, 43], comprehensive connections between ADRs and off-targets were established through rigorous testing and analysis on commercially available drugs. Leveraging their data, we created a mapping linking each ADR with its related off-targets. To ensure precision in subsequent ADR enrichment analysis, ADRs corresponding to off-targets fewer than 3 were excluded, and those with severity scores less than 0.1 were filtered to retain more hazardous ADRs. ADRs severity scores ranging from 0 to 1 were obtained from Stanford' s study, which ranked 2929 ADRs through crowdsourcing [15]. Consequently, we obtained 358 ADR terms associated with 193 off-targets (all included in our off-target panel). These ADR-targets mappings served as an annotation database like gene ontology (GO) used in gene enrichment analysis, providing prior information for ADR enrichment analysis.

The enrichment analysis utilized the hypergeometric distribution. For a given ADR, the probability that predicted off-targets for a drug are included in the ADR-related off-target set is calculated by Eq. (1):

$$p(k) = P(X = k) = \frac{\binom{N-K}{k}\binom{n-k}{K}}{\binom{N}{n}}$$

Here, N represents the total number of targets in the annotation database (194 targets); n denotes the number of predicted positive targets for a drug, and K denotes the number of targets belonging to the specific ADR term, of which k is predicted as positive targets for the drug. In essence, this equation signifies the probability that the predicted off-target profile for a drug is enriched in a specific ADR. For enrichment analysis, p-values are calculated based on the binomial approximation of the hypergeometric distribution, with Bonferroni correction and multiple hypothesis testing performed by FDR adjustment [44, 45] (See Supporting Information Text S3 for detailed instructions). ADR enrichment analysis was conducted using the enrichr function in GSEApy Python package (version 1.0.6).

2.4. Evaluation metrics

The classification model' s performance is assessed using AUROC, Balanced Accuracy (BACC), Matthews Correlation Coefficient (MCC), and F1 score—critical indicators for evaluating classifiers in imbalanced data scenarios. These metrics are computed from the confusion matrix: True Positives (TP), False Positives (FP), False Negatives (FN), and True Negatives (TN).

Specifically, BACC, MCC, and F1 are calculated using Eq. (2), (3), and (4) respectively:

$$Balanced\ accuracy = \frac{TPR + TNR}{2}$$

$$MCC = \frac{TP \times TN - FP \times FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}}$$

$$F1 = \frac{2 \times TP}{2 \times TP + FP + FN}$$

The Area Under the Receiver Operating Characteristic Curve (AUROC) is the area beneath the ROC curve, which consists of the True Positive Rate (TPR) (Eq. (5)) against the False Positive Rate (FPR) (Eq. (6)) at different thresholds:

$$TPR(recall) = \frac{TP}{TP + FN}$$

$$FPR = \frac{FP}{FP + TN}$$

$$AUROC = \int TPR(FPR)d(FPR)$$

For multi-label learning, two prominent ranking-based metrics hold significance: Mean Average Precision (mAP) and Rank Loss. mAP, particularly, functions as an indicator of the ranking quality reflected in prediction outcomes. In the context of C classes of labels, mAP is calculated as the average of the Average Precision (AP) across all classes (Eq. (8)):

$$mAP = \frac{1}{C}\sum_{i=1}^{C} AP_i$$

where AP is determined by calculating the area under the Precision-Recall curve for each class i (Eq. (9)):

$$AP = \sum_{k=1}^{n}(Precision_k \times (Recall_k - Recall_{k-1}))$$

where Precision and Recall are computed at each threshold k, i.e., a specific point along the Precision-Recall curve; n is the number of thresholds, i.e., the total number of positive instances.

Rank Loss is a metric employed in multi-label classification to quantify the quality of the ranking assigned to positive labels for each instance. The calculation

involves sorting the predicted probabilities of positive labels in descending order for each instance, forming a ranked list. Subsequently, the algorithm counts the number of inversions, representing instances where the predicted ranking contradicts the true ranking of positive labels. This count is then divided by the total number of possible pairs of positive labels, calculated as Eq. (10):

$$Rank\ loss = \frac{Number\ of\ Inversions}{Total\ Number\ of\ Possible\ Pairs}$$

where the total number of possible pairs is calculated as $n \times (n-1)/2$, and n is the number of positive labels for the instance. The process is repeated for all instances, and the average Rank Loss is determined by averaging the obtained values. A lower Rank Loss value signifies superior performance, with an optimal score of 0, indicating perfect agreement between predicted and true rankings.

3.1. The overall workflow of drug safety analysis based on off-target prediction

Figure 1 [Figure 1: see original paper] illustrates the off-target prediction model for drugs and the utilization of the off-target representation. The workflow of drug safety analysis based on off-target prediction is shown in Fig. 1. Initially, we constructed an off-target panel and curated corresponding target-compound interaction data. Based on these data, we built 7 ligand-based off-target prediction models through multi-task GNNs, corresponding to 7 target families: GPCR, ion channel, enzyme, kinase, transporter, nuclear receptor, and others. Consequently, the binding probabilities against the off-target panel can be obtained for each compound. The predicted off-target profiles can then be employed as molecular representations for the subsequent classification of a drug' s ATC, toxicity, and ADR enrichment analysis.

3.2. Classification performance of off-target prediction models

We constructed multi-task GNNs (MTGNN) for the ligand-based prediction of off-target profiles. The panel for off-target analysis comprises 90 protein targets of Homo sapiens collected from previous research [1, 9, 10], comprising 50 GPCR targets, 16 ion channel targets, 9 enzyme targets, 6 kinase targets, 4 nuclear receptor targets, 4 transporter targets, and 2 other targets. Subsequently, seven distinct multi-task GNN models were created, one for each of the seven target families. The multi-task strategy was employed to leverage shared information among tasks within each target family [39], resulting in enhanced model quality and robustness while simultaneously preventing cross-family negative transfer. Additionally, corresponding protein targets from other species were included in the training data to improve the multitask model, as these targets could serve as valuable sources of inductive bias. The hyperparameter values for

these seven multi-task GNN models were individually optimized through grid searching (see Supporting Information Table S5). As a result, for each compound, the interaction probabilities with the 242 targets can be inferred from these seven multi-task GNNs (see Method; the information for the off-targets can be found in Supporting Information Table S1).

We compared the performance of our off-target prediction MTGNN model with Roche' s off-target prediction methods, including NeuralNetworks, RandomForest, and Auto-Sklearn. Detailed information on the models and their parameter settings can be found in Roche' s work [13] and Supporting Information Table S6. Fig. 2A [Figure 2: see original paper] provides an overview of the highest-scoring models based on AUROC, BACC, MCC, and F1 scores. In terms of AUROC and BACC, MTGNN scored higher than the other models for 134 and 143 targets, respectively. In terms of MCC, MTGNN outperformed NeuralNetworks and Auto-Sklearn. Regarding F1 score, RandomForest had the best MCC for up to 94 targets, but its overall performance was suboptimal due to poor AUROC and BACC. MTGNN outperformed RandomForest and Auto-Sklearn but was inferior to NeuralNetworks. Given the importance of distinguishing true positives and true negatives in off-target modelling, we considered that BACC serves as the primary evaluation indicator [13].

As can be observed from the BACC results in Fig. 2B, MTGNN demonstrates the highest average BACC compared to the other three models for each target family. Furthermore, when considering the average performance for all targets together while disregarding their target type, MTGNN outperformed all Roche' s off-target models in all indicators (Supporting Information Fig. S2). Further details and evaluation metric values for these methods under each target type are available in Supporting Information Table S7.

In the specific context of this study, it is evident that MTGNN exhibits superior overall performance, thus positioning it as a top-ranked method. The robust predictive capacity of MTGNN is attributed to its integration of domain information from related tasks, thereby enhancing its performance on tasks characterized by limited data availability. As illustrated in Fig. 2C, MTGNN consistently outperforms Neural Networks across a range of dataset sizes, which was identified as the best-performing model by Roche. This is particularly pronounced when dealing with data insufficiency, as seen in the dataset size intervals of [1, 100] and [100, 500]. This observation is of paramount significance for the extension of target range research aimed at broad and comprehensive predictions of off-target effects.

It' s noteworthy that experts advocate the utilization of human targets, rather than animal homologs, in constructing off-target panels for predicting ADRs in humans [1]. Thus, we focus on the predicted outcomes for 90 human targets. As depicted in Fig. 2D, these human targets consistently exhibit robust classification performance, with a significant majority achieving BACC scores exceeding 0.7. In Fig. 2E, a detailed exploration of precision and recall metrics for each of the 90 human targets reveals high recall values (>0.8) for most tasks,

coupled with moderate precision levels ranging from 0.4 to 0.6. Despite the low positive rate (<25%) for these tasks due to augmented negative samples, the models maintain their sensitivity to potentially unsafe compound-target interactions. Given the paramount importance of avoiding false negatives in early-stage drug development to prevent the oversight of unsafe molecules during safety assessments, our model holds considerable value in mitigating research and development failures arising from unsafe off-target interactions.

3.3. Application of off-target prediction panel

The off-target prediction model offers the capability to infer interaction probabilities against 242 targets for any given molecule. This can be regarded as a 242-dimensional representation that characterizes the off-target-related molecular features. We explore the utilization of these representations in drug ATC classification, toxicity prediction, and ADR enrichment analysis.

3.3.1. Drug ATC classification

The Anatomical Therapeutic Chemical (ATC) classification system was proposed by the World Health Organization (WHO) in 1981 (https://www.whocc.no/atc/structure_{{{and}}}_{{p
Researchers leverage the hierarchical structure of ATC codes as features to improve the performance of ADR prediction models. Drugs categorized under different ATC codes often manifest specific ADRs, influenced by their off-target profiles [46].

To assess how the off-target representation characterizes the ATC code, as described in Supporting Information Text S4, we modeled the ATC classification as a multi-label problem, where each compound corresponds to 14 ATC labels. From the ATC-SMILES dataset, a benchmark collection designed for the ATC classification task [47], we curated a total of 3491 compounds spanning 14 categories. For precise counts of compounds under each ATC code, refer to Supporting Information Table S8. The impact of off-target representation on ATC classification was demonstrated through a comparative experiment involving two models: MLKNN and ECFP_{MLKNN}, using the compound' s off-target representation and molecular fingerprint features (1024-dimensional ECFP4 fingerprint) as characteristics, respectively. Conducting five-fold cross-training and evaluation on the same test set revealed that MLKNN outperformed ECFP_{MLKNN}, as indicated by superior AUROC, mAP, and Rank Loss metrics (Fig. 3A [Figure 3: see original paper], and Supporting Information Table S9). This underscores the efficacy of the off-target representation-based multi-label model in accurately ordering ATC codes for compounds, surpassing the performance of conventional molecular fingerprint features.

It' s noteworthy that compounds classified under "Nervous system (N)" manifest a higher frequency of off-target bindings (Fig. 3B), as depicted in Fig. 3C, in contrast to other drug categories, evident in denser and darker points on the heatmap. Prior investigations have highlighted that in drugs exhibiting fatal

toxicity, 78.6% acted on the Nervous system (N) [48]. This occurrence can be attributed to the fact that many drugs in the N category often display pharmacological promiscuity, targeting GPCR receptors [49, 50], including adenosine receptors, acetylcholine receptors, serotonin receptors, along with potassium ion channels, voltage-gated sodium ion channels, and specific transporter targets [51-54]. These characteristics are reflected in the off-target representation of these drugs.

3.3.2. Toxicity prediction

The off-target representation is considered a crucial feature for toxicity prediction, forming the basis for an off-target-based toxicity prediction approach applicable to any given drug. To conduct the experiment, we curated a drug toxicity dataset from different sources: (1) The Clintox dataset, obtained from MoleculeNet [55], comprises 108 toxic and 1365 non-toxic compounds after data cleaning. (2) From DrugBank and online resources, we collected 107 drugs withdrawn due to toxic side effects. (3) Onakpoya et al. [56] compiled 462 drugs withdrawn globally for toxic side effects, resulting in 390 small molecules. (4) ChEMBL provided 865 compounds flagged with 'Black Box Warnings' , identifying 505 toxic compounds. After merging and deduplication, the dataset includes 877 toxic (labeled as 1) and 1229 non-toxic compounds (labeled as 0).

UMAP visualization was employed to illustrate the relationship between off-target representation and compound toxicity. As shown in Fig. 4A [Figure 4: see original paper], the off-target representation exhibited clearer discrimination between safe and unsafe compounds, outperforming molecular fingerprint features (Fig. 4B). This observation may be attributed to the evidently fewer bound safety-related off-targets for most safe drugs, a phenomenon illustrated by the denser and darker points in the heatmap for toxic drugs in contrast to non-toxic drugs (Fig. 4C). Moreover, to gauge the impact of off-target and structural representation on the compound toxicity prediction model, we employed a toxicity classifier using LightGBM based on the off-target representation. We also implemented ECFP_{LightGBM}, where the off-target representation was substituted with a molecular fingerprint feature (1024-dimensional ECFP4 fingerprint). Compared to LightGBM, the performance of ECFP_{LightGBM} exhibited a noticeable decline on the same test set (Fig. 4D, Supporting Information Table S11). These findings underscore the critical role of off-target representation in the assessment of drug toxicity.

3.3.3. ADR enrichment analysis

The relationship between an Adverse Drug Reaction (ADR) and its corresponding off-target can be analogized to that of a biological pathway and its associated gene set. Upon obtaining the predicted off-target profile of a queried drug, potential ADRs can be inferred through enrichment analysis, employing an annotation database that correlates each ADR with its corresponding off-target set. Our study established mappings between 358 ADR terms and 193 off-targets,

serving as an annotation database for prior information in ADR enrichment analysis, conducted using the hypergeometric distribution (refer to Method).

We conducted ADR enrichment analysis on four withdrawn drugs due to safety concerns—Pergolide, Phenylpropanolamine, Sibutramine, and Sertindole—evaluating the efficacy of the enriched ADRs compared to the known relevant ADRs of these drugs (Supporting Information Table S12). Positive off-target predictions for these drugs were obtained, analogous to "differential genes" , with prediction values exceeding 0.3 to broaden the scope of enriched ADRs beyond the conventional threshold of 0.5.

Fig. 5A [Figure 5: see original paper] illustrates the ranking of ADR enrichment analysis results for each drug based on Adjusted P-value. Detailed enrichment analysis ranking results for each drug can be found in Supporting Information Tables S13-S16. (1) Pergolide is a dopamine receptor agonist commonly utilized in the treatment of Parkinson' s disease and other conditions [57], which has been associated with an increased risk of cardiac valvulopathy, leading to its withdrawal from the US and Canadian markets in 2007. Among the 42 pertinent ADRs for Pergolide, 16 are significantly enriched ($p < 0.05$) out of a total of 358 ADRs. These include high-frequency ADR terms associated with pergolide such as Orthostatic hypotension (frequent, 9%), Extrapyramidal disorder (frequent, 1.6%), Insomnia (frequent, 7.9%), and Dyskinesia (frequent, 62.4%). Notably, consistent with Pergolide' s withdrawal due to cardiotoxicity, several cardiotoxic-related ADR terms were significantly enriched, including Tachycardia, Cardiac failure congestive, and Heart rate increased. (2) Similarly, Phenylpropanolamine has 25 relevant ADRs, and 9 of them are significantly enriched (Fig. 5B). Neurotoxicity and Central nervous system stimulation, two significantly enriched ADR terms, are associated with hemorrhagic stroke [58-60], the primary reason for the withdrawal of Phenylpropanolamine from the market. Apart from its impacts on the nervous system, Phenylpropanolamine also induced a series of cardiac side effects [61], including significantly enriched Arrhythmia, Cardiac failure, and Prolonged electrocardiogram QT. The recognition of critical and potentially fatal adverse drug reactions, such as prolonged Electrocardiogram QT, in the early stages of drug discovery is imperative for prioritizing human safety and identifying potential risks. (3) Sibutramine was withdrawn from the Canadian and U.S. markets due to the increased risk of heart attacks and strokes in patients with a history of heart disease. It is associated with 35 known ADRs, and the enrichment results reveal 9 significant ADRs among them (Fig. 5C). Notably, Tachycardia, Anticholinergic syndrome, and Cerebrovascular disorder [62, 63], which are related to its withdrawal, are significantly enriched. (4) Sertindole, withdrawn from the market due to cardiotoxicity, did not show significant enrichment in cardiotoxic-related Tachycardia. However, 9 out of 25 ADRs associated with this drug were significantly enriched, with all of the top six enrichment results being known ADRs of this drug (Fig. 5D). These examples underscore the effectiveness of off-target prediction results in enriching relevant ADRs, particularly severe ADRs leading to drug withdrawal, thereby highlighting the validity of off-target representation

in characterizing ADRs.

3.4. Drug-target-ADR networks

Through off-target prediction and subsequent ADR enrichment analysis, we can establish correlations between a drug' s off-targets and the corresponding ADRs, providing an off-target-based explanation for ADRs. Using Pergolide as a case study, we predicted its off-target profile and correlated the known ADRs of the drug with the predicted off-targets to create a drug-target-ADR correlation network.

Firstly, the accurate prediction of Pergolide' s off-target profile was achieved (known off-targets are obtained from databases such as ChEMBL, PubChem, and DrugBank, overlapping with our off-target panel). Fig. 6 [Figure 6: see original paper] demonstrates that out of the 10 known targets of Pergolide predictable by the off-target model, 8 were appropriately included in the predicted off-target profile. Subsequently, using the off-target representation, the ADR enrichment analysis correctly identified its crucial ADRs related to Cardiotoxicity, Orthostatic hypotension, Insomnia, etc. (see ADR enrichment analysis). Furthermore, the off-target model indicated the presence of additional potential off-targets related to Pergolide' s ADRs, providing potential explanations for its drug-target-ADR correlation.

For instance, cardiac-related toxicities associated with Pergolide were linked to both its known targets (HTR2A, ADRA2B, and HTR2B [64]) and predicted targets (CHRM1 and CACNA1C). Diarrhoea, a prevalent side effect of Pergolide, can be ascribed not only to Pergolide's known target ADRA2A but also possibly to the predicted targets SLC6A4 [65] and TACR2 [66]. For Insomnia, besides the known target HTR1A of Pergolide, two predicted new targets, SLC6A4 [67] and HTR7 [68], were correlated with this ADR. Similarly, the drug-target-network diagram for Sertindole can be found in Supporting Information Fig. S3, and the relationship between side effects and targets can be analyzed in a similar manner as for Pergolide.

3.5. Code availability

All data and scripts to build the models are available in the GitHub repository: https://github.com/myzhengSIMM/Offtarget_{drugsafety}.

4. Conclusions

Off-target interactions frequently occur with drug usage and are a major cause of drug side effects and candidate failure during drug discovery. We employed a multi-task GNN to accurately predict these off-target interactions derived from molecular graphs. These predictions were then utilized to comprehensively assess drug safety from multiple aspects, including ATC catalogs, toxicity, and ADRs, providing a valuable supplement to traditional, time-consuming, and

labor-intensive safety pharmacology experiments. Notably, in ADR enrichment analysis, based on differential targets for each drug derived from off-target prediction results, severe ADRs leading to drug withdrawal were significantly enriched in our cases, further illustrating the effectiveness of the off-target panel for drug safety assessment.

One limitation of our study lies in the reliance on ligand similarity for off-target prediction, excluding protein-related information. The variability in predictive performance across different protein targets and families indicated potential disparities driven by data or biological factors. Therefore, incorporating protein-related insights into off-target predictions holds promise for enhancing the accuracy of our predictions. Additionally, it' s important to note that our model is not universally applicable for off-target prediction, and re-modeling may be necessary when the number of targets changes. Inspired by the application of pre-training and transfer learning in the field of medicine [69-71], we consider building a general paradigm for off-target panels prediction by simultaneously incorporating information from both proteins and compounds in future work. In terms of safety assessment, the free plasma concentration of drugs significantly impacts adverse reactions and drug safety [3, 72]. Compounds predicted to have multiple off-target bindings pose lower risks when they have a low free plasma concentration. Conversely, compounds with fewer target bindings but high free plasma concentrations can be more hazardous [10]. Hence, evaluating drug safety and explaining ADRs should consider the free plasma concentration during therapeutic use. It' s noteworthy that drugs used to treat severe, refractory diseases might result in more frequent and severe side effects, which are deemed acceptable [29].

Overall, our work aims to predict drug off-target interactions to assess drug safety. Utilizing a compound' s off-target representation to deduce ATC catalogs, toxicity, and ADR offers a valuable framework and methodology for the preclinical identification of compound toxicity. Future steps include expanding the off-target panel, optimizing off-target and ADR prediction models, and refining the safety prediction model to contribute to the development of safer pharmaceuticals.

References

1. Bowes J, Brown AJ, Hamon J, Jarolimek W, Sridhar A, Waldron G, Whitebread S: Reducing safety-related drug attrition: the use of in vitro pharmacological profiling. Nat Rev Drug Discov 2012; 11(12):909-922.

2. Whitebread S, Hamon J, Bojanic D, Urban L: Keynote review: in vitro safety pharmacology profiling: an essential tool for successful drug development. Drug Discov Today 2005; 10(21):1421-1433.

3. Sutherland JJ, Yonchev D, Fekete A, Urban L: A preclinical secondary pharmacology resource illuminates target-adverse drug reaction associations of marketed drugs. Nat Commun 2023; 14(1):4323.

4. Hao Y, Moore JH: TargetTox: A Feature Selection Pipeline for Identifying Predictive Targets Associated with Drug Toxicity. J Chem Inf Model 2021; 61(11):5386-5394.

5. John L, Mahanta HJ, Soujanya Y, Sastry GN: Assessing machine learning approaches for predicting failures of investigational drug candidates during clinical trials. Comput Biol Med 2023; 153:106494.

6. Siramshetty VB, Nickel J, Omieczynski C, Gohlke BO, Drwal MN, Preissner R: WITHDRAWN–a resource for withdrawn and discontinued drugs. Nucleic Acids Res 2016; 44(D1):D1080-1086.

7. Vo AH, Van Vleet TR, Gupta RR, Liguori MJ, Rao MS: An Overview of Machine Learning and Big Data for Drug Toxicity Evaluation. Chem Res Toxicol 2020; 33(1):20-37.

8. Cook D, Brown D, Alexander R, March R, Morgan P, Satterthwaite G, Pangalos MN: Lessons learned from the fate of AstraZeneca's drug pipeline: a five-dimensional framework. Nat Rev Drug Discov 2014; 13(6):419-431.

9. Lynch JJ, 3rd, Van Vleet TR, Mittelstadt SW, Blomme EAG: Potential functional and pathological side effects related to off-target pharmacological activity. J Pharmacol Toxicol Methods 2017; 87:108-126.

10. Bendels S, Bissantz C, Fasching B, Gerebtzoff G, Guba W, Kansy M, Migeon J, Mohr S, Peters JU, Tillier F et al: Safety screening in early drug discovery: An optimized assay panel. J Pharmacol Toxicol Methods 2019; 99:106609.

11. Li Y, Fan Z, Rao J, Chen Z, Chu Q, Zheng M, Li X: An overview of recent advances and challenges in predicting compound-protein interaction (CPI). Medical Review 2023; 0(0).

12. Mayr A, Klambauer G, Unterthiner T, Steijaert M, Wegner JK, Ceulemans H, Clevert DA, Hochreiter S: Large-scale comparison of machine learning methods for drug target prediction on ChEMBL. Chem Sci 2018; 9(24):5441-5451.

13. Naga D, Muster W, Musvasva E, Ecker GF: Off-targetP ML: an open source machine learning framework for off-target panel safety assessment of small molecules. J Cheminform 2022; 14(1):27.

14. Lunghini F, Fava A, Pisapia V, Sacco F, Iaconis D, Beccari AR: ProfhEX: AI-based platform for small molecules liability profiling. J Cheminform 2023; 15(1):60.

15. Gottlieb A, Hoehndorf R, Dumontier M, Altman RB: Ranking adverse drug reactions with crowdsourcing. J Med Internet Res 2015; 17(3):e80.

16. Ietswaart R, Arat S, Chen AX, Farahmand S, Kim B, DuMouchel W, Armstrong D, Fekete A, Sutherland JJ, Urban L: Machine learning guided

association of adverse drug reactions with in vitro target-based pharmacology. EBioMedicine 2020; 57:102837.

17. Perez-Nueno VI, Souchet M, Karaboga AS, Ritchie DW: GESSE: Predicting Drug Side Effects from Drug-Target Relationships. J Chem Inf Model 2015; 55(9):1804-1823.

18. Dey S, Luo H, Fokoue A, Hu J, Zhang P: Predicting adverse drug reactions through interpretable deep learning framework. BMC Bioinformatics 2018; 19(Suppl 21):476.

19. Liu M, Wu Y, Chen Y, Sun J, Zhao Z, Chen XW, Matheny ME, Xu H: Large-scale prediction of adverse drug reactions using chemical, biological, and phenotypic properties of drugs. J Am Med Inform Assoc 2012; 19(e1):e28-35.

20. Zhang W, Liu F, Luo L, Zhang J: Predicting drug side effects by multi-label learning and ensemble learning. BMC Bioinformatics 2015; 16:365.

21. Zhang F, Sun B, Diao X, Zhao W, Shu T: Prediction of adverse drug reactions based on knowledge graph embedding. BMC Med Inform Decis Mak 2021; 21(1):38.

22. Hemmerich J, Ecker GF: In silico toxicology: From structure-activity relationships towards deep learning and adverse outcome pathways. Wiley Interdiscip Rev Comput Mol Sci 2020; 10(4):e1475.

23. Raies AB, Bajic VB: In silico toxicology: computational methods for the prediction of chemical toxicity. Wiley Interdiscip Rev Comput Mol Sci 2016; 6(2):147-172.

24. Lo YC, Rensi SE, Torng W, Altman RB: Machine learning in chemoinformatics and drug discovery. Drug Discov Today 2018; 23(8):1538-1546.

25. Allen TEH, Goodman JM, Gutsell S, Russell PJ: Using 2D Structural Alerts to Define Chemical Categories for Molecular Initiating Events. Toxicol Sci 2018; 165(1):213-223.

26. Metivier JP, Lepailleur A, Buzmakov A, Poezevara G, Cremilleux B, Kuznetsov SO, Le Goff J, Napoli A, Bureau R, Cuissart B: Discovering structural alerts for mutagenicity using stable emerging molecular patterns. J Chem Inf Model 2015; 55(5):925-940.

27. Mellor CL, Marchese Robinson RL, Benigni R, Ebbrell D, Enoch SJ, Firman JW, Madden JC, Pawar G, Yang C, Cronin MTD: Molecular fingerprint-derived similarity measures for toxicological read-across: Recommendations for optimal use. Regulatory Toxicology and Pharmacology 2019; 101:121-134.

28. Cherkasov A, Muratov EN, Fourches D, Varnek A, Baskin, II, Cronin M, Dearden J, Gramatica P, Martin YC, Todeschini R et al: QSAR modeling:

where have you been? Where are you going to? J Med Chem 2014; 57(12):4977-5010.

29. Gayvert KM, Madhukar NS, Elemento O: A Data-Driven Approach to Predicting Successes and Failures of Clinical Trials. Cell Chem Biol 2016; 23(10):1294-1301.

30. Nguyen DA, Nguyen CH, Mamitsuka H: A survey on adverse drug reaction studies: data, tasks and machine learning methods. Brief Bioinform 2021; 22(1):164-177.

31. Mendez D, Gaulton A, Bento AP, Chambers J, De Veij M, Felix E, Magarinos MP, Mosquera JF, Mutowo P, Nowotka M et al: ChEMBL: towards direct deposition of bioassay data. Nucleic Acids Res 2019; 47(D1):D930-D940.

32. Kim S, Chen J, Cheng T, Gindulyte A, He J, He S, Li Q, Shoemaker BA, Thiessen PA, Yu B et al: PubChem in 2021: new data content and improved web interfaces. Nucleic Acids Res 2021; 49(D1):D1388-D1395.

33. Rodgers G, Austin C, Anderson J, Pawlyk A, Colvis C, Margolis R, Baker J: Glimmers in illuminating the druggable genome. Nat Rev Drug Discov 2018; 17(5):301-302.

34. Oprea TI, Bologa CG, Brunak S, Campbell A, Gan GN, Gaulton A, Gomez SM, Guha R, Hersey A, Holmes J et al: Unexplored therapeutic opportunities in the human genome. Nat Rev Drug Discov 2018; 17(5):317-332.

35. Nath A, Leier A: Improved cytokine-receptor interaction prediction by exploiting the negative sample space. BMC Bioinformatics 2020; 21(1):493.

36. Imrie F, Bradley AR, Deane CM: Generating property-matched decoy molecules using deep learning. Bioinformatics 2021; 37(15):2134-2141.

37. Chen D, Gao K, Nguyen DD, Chen X, Jiang Y, Wei GW, Pan F: Algebraic graph-assisted bidirectional transformers for molecular property prediction. Nat Commun 2021; 12(1):3521.

38. Crawshaw M: Multi-task learning with deep neural networks: A survey. arXiv preprint arXiv:200909796 2020.

39. Ruder S: An overview of multi-task learning in deep neural networks. arXiv preprint arXiv:170605098 2017.

40. Xiong Z, Wang D, Liu X, Zhong F, Wan X, Li X, Li Z, Luo X, Chen K, Jiang H et al: Pushing the Boundaries of Molecular Representation for Drug Discovery with the Graph Attention Mechanism. J Med Chem 2020; 63(16):8749-8760.

41. Wang Z, Clark NR, Ma' ayan A: Drug-induced adverse events prediction with the LINCS L1000 data. Bioinformatics 2016; 32(15):2338-2345.

42. Lounkine E, Keiser MJ, Whitebread S, Mikhailov D, Hamon J, Jenkins JL, Lavan P, Weber E, Doak AK, Cote S et al: Large-scale prediction and testing of drug activity on side-effect targets. Nature 2012; 486(7403):361-367.

43. Huang LH, He QS, Liu K, Cheng J, Zhong MD, Chen LS, Yao LX, Ji ZL: ADReCS-Target: target profiles for aiding drug safety research and application. Nucleic Acids Res 2018; 46(D1):D911-D917.

44. Rappoport N, Fromer M, Schweiger R, Linial M: PANDORA: analysis of protein and peptide sets through the hierarchical integration of annotations. Nucleic Acids Res 2010; 38(Web Server issue):W84-89.

45. Korthauer K, Kimes PK, Duvallet C, Reyes A, Subramanian A, Teng M, Shukla C, Alm EJ, Hicks SC: A practical guide to methods controlling false discoveries in computational biology. Genome Biol 2019; 20(1):118.

46. Henriksson A, Zhao J, Dalianis H, Bostrom H: Ensembles of randomized trees using diverse distributed representations of clinical events. BMC Med Inform Decis Mak 2016; 16 Suppl 2(Suppl 2):69.

47. Cao Y, Yang ZQ, Zhang XL, Fan W, Wang Y, Shen J, Wei DQ, Li Q, Wei XY: Identifying the kind behind SMILES-anatomical therapeutic chemical classification using structure-only representations. Brief Bioinform 2022; 23(5).

48. Ojanpera I, Kriikku P, Vuori E: Fatal toxicity index of medicinal drugs based on a comprehensive toxicology database. Int J Legal Med 2016; 130(5):1209-1216.

49. Hauser AS, Attwood MM, Rask-Andersen M, Schioth HB, Gloriam DE: Trends in GPCR drug discovery: new agents, targets and indications. Nat Rev Drug Discov 2017; 16(12):829-842.

50. Wendell SG, Fan H, Zhang C: G Protein-Coupled Receptors in Asthma Therapy: Pharmacology and Drug Action. Pharmacol Rev 2020; 72(1):1-49.

51. Dolphin AC, Insel PA, Blaschke TF, Meyer UA: Introduction to the Theme "Ion Channels and Neuropharmacology: From the Past to the Future" . Annu Rev Pharmacol Toxicol 2020; 60:1-6.

52. Belvisi MG, Birrell MA: The emerging role of transient receptor potential channels in chronic lung disease. Eur Respir J 2017; 50(2).

53. Ayala-Lopez N, Watts SW: Physiology and Pharmacology of Neurotransmitter Transporters. Compr Physiol 2021; 11(3):2279-2295.

54. Pera T, Penn RB: Bronchoprotection and bronchorelaxation in asthma: New targets, and new ways to target the old ones. Pharmacol Ther 2016; 164:82-96.

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

56. Onakpoya IJ, Heneghan CJ, Aronson JK: Post-marketing withdrawal of 462 medicinal products because of adverse drug reactions: a systematic review of the world literature. BMC Med 2016; 14:10.

57. Clarke CE, Speller JM: Pergolide for levodopa-induced complications in Parkinson' s disease. Cochrane Database Syst Rev 2000(2):CD000235.

58. Puy L, Parry-Jones AR, Sandset EC, Dowlatshahi D, Ziai W, Cordonnier C: Intracerebral haemorrhage. Nat Rev Dis Primers 2023; 9(1):14.

59. Walker JS: Phenylpropanolamine potentiates caffeine neurotoxicity in rats. Journal of Pharmaceutical Sciences 1989; 78(12):986-989.

60. Dietz AJ, Jr: Amphetamine-like Reactions to Phenylpropanolamine. JAMA 1981; 245(6):601-602.

61. Lake CR, Gallant S, Masson E, Miller P: Adverse drug effects attributed to phenylpropanolamine: a review of 142 case reports. Am J Med 1990; 89(2):195-208.

62. Caprio FZ, Sorond FA: Cerebrovascular Disease: Primary and Secondary Stroke Prevention. Med Clin North Am 2019; 103(2):295-308.

63. Liu L, Chen W, Zhou H, Duan W, Li S, Huo X, Xu W, Huang L, Zheng H, Liu J et al: Chinese Stroke Association guidelines for clinical management of cerebrovascular disorders: executive summary and 2019 update of clinical management of ischaemic cerebrovascular diseases. Stroke Vasc Neurol 2020; 5(2):159-176.

64. Hutcheson JD, Setola V, Roth BL, Merryman WD: Serotonin receptors and heart valve disease–it was meant 2B. Pharmacol Ther 2011; 132(2):146-157.

65. Camilleri M: Is there a SERT-ain association with IBS? Gut 2004; 53(10):1396-1399.

66. Tack J, Schumacher K, Tonini G, Scartoni S, Capriati A, Maggi C: The neurokinin-2 receptor antagonist ibodutant improves overall symptoms, abdominal pain and stool pattern in female patients in a phase II study of diarrhoea-predominant IBS. Gut 2017; 66(8):1403-1413.

67. Perlis ML, Corbitt CB, Kloss JD: Insomnia research: 3Ps and beyond. Sleep Med Rev 2014; 18(3):191-193.

68. Hedlund PB, Huitron-Resendiz S, Henriksen SJ, Sutcliffe JG: 5-HT7 receptor inhibition and inactivation induce antidepressantlike behavior and sleep pattern. Biol Psychiatry 2005; 58(10):831-837.

69. Theodoris CV, Xiao L, Chopra A, Chaffin MD, Al Sayed ZR, Hill MC, Mantineo H, Brydon EM, Zeng Z, Liu XS et al: Transfer learning enables predictions in network biology. Nature 2023; 618(7965):616-624.

70. Kroll A, Ranjan S, Engqvist MKM, Lercher MJ: A general model to predict small molecule substrates of enzymes based on machine and deep learning. Nature Communications 2023; 14(1):2787.

71. Dalkıran A, Atakan A, Rifaioğlu AS, Martin MJ, Atalay RÇ, Acar AC, Doğan T, Atalay V: Transfer learning for drug–target interaction prediction. Bioinformatics 2023; 39(Supplement_1):i103-i110.

72. Dudley JT, Deshpande T, Butte AJ: Exploiting drug-disease relationships for computational drug repositioning. Brief Bioinform 2011; 12(4):303-311.

Appendix A. Supplementary data

Multimedia component 1

Text S1. Establishment of off-target panel.
Text S2. The compound-protein interaction data collection and processing process.
Text S3. The principle of enrichment analysis.
Text S4. The introduction of MLKNN and the data split method for ATC classification.
Text S5. The introduction of LightGBM, the data process and hyperparameter search process.
Text S6. The drug-target-network analysis of Sertindole.

Fig. S1. Physico-chemical spatial distribution of positive molecules and relative decoys under different target classes.
Fig. S2. The comparison of the performance of the four off-target prediction models across all targets on the AUROC, MCC, BACC, and F1 metrics.
Fig. S3. Drug-target-ADR association diagram for Sertindole.

Table S2. The list of databases used in the work.
Table S3. Activity thresholds corresponding to different types of targets.
Table S4. Hyperparametric search range for MTGNN.
Table S5. Hyperparameter settings of the models corresponding to seven target types.
Table S6. The parameter settings of Roche' s off-target prediction model in work.
Table S7. Model performance of different off-target prediction models under different target types.
Table S8. Drug data statistics under the first-level ATC code.
Table S9. Comparison of model effects of MLKNN and its ablation experiment for ATC classification.
Table S10. The hyperparameter search range and best value of LightGBM and

ECFP_{LightGBM}.
Table S11. Comparison of model effects of LightGBM and its ablation experiments.
Table S12. Drugs and their associated ADRs.

Multimedia component 2

Table S1. The detailed description of off-targets and their side effects.

Multimedia component 3

Tables contained ADR enrichment analysis results for drugs.
Table S13. The ADR enrichment analysis results for pergolide.
Table S14. The ADR enrichment analysis results for phenylpropanolamine.
Table S15. The ADR enrichment analysis results for sibutramine.
Table S16. The ADR enrichment analysis results for sertindole.

Note: Figure translations are in progress. See original paper for figures.