APRILE: Exploring the Molecular Mechanisms of Drug Side Effects with Explainable Graph Neural Networks

A Preprint

Hao Xu
Queen's University
xu.hao@queensu.ca

Shengqi Sang
University of Waterloo
s4sang@uwaterloo.ca

Herbert Yao
Queen's University
16hy16@queensu.ca

Alexandra I. Herghelegiu
University of Sheffield
aiherghelegiu1@sheffield.ac.uk

Haiping Lu
University of Sheffield
h.lu@sheffield.ac.uk

Laurence Yang*
Queen's University
laurence.yang@queensu.ca

July 2, 2021

Abstract

With the majority of people 65 and over taking two or more medicines (polypharmacy), managing the side effects associated with polypharmacy is a global challenge. Explainable Artificial Intelligence (XAI) is necessary to reliably design safe polypharmacy. Here, we develop APRILE: a predictor-explainer framework based on graph neural networks to explore the molecular mechanisms underlying polypharmacy side effects by explaining predictions made by the predictors. For a side effect and its associated drug pair, or a set of side effects and their drug pairs, APRILE gives a set of proteins (drug targets or non-targets) and Gene Ontology (GO) items as the explanation. Using APRILE, we generate such explanations for 843,318 (learned) + 93,966 (novel) side effect–drug pair events, spanning 861 side effects (472 diseases, 485 symptoms and 9 mental disorders) and 20 disease categories. We show that our two new metrics, pharmacogenomic information utilization and protein-protein interaction information utilization, provide quantitative estimates of mechanism complexity. Explanations were significantly consistent with state of the art disease-gene associations for 232/239 (97%) side effects. Further, APRILE generated new insights into molecular mechanisms of four diverse categories of ADRs: infection, metabolic diseases, gastrointestinal diseases, and mental disorders, including paradoxical side effects. We demonstrate the viability of discovering polypharmacy side effect mechanisms by learning from an AI model trained on massive biomedical data. Consequently, it facilitates wider and more reliable use of AI in healthcare.

1 Introduction

Of people aged 65 and over, 75% are on two or more medicines, while 7-35% are prescribed ten or more medicines [1]. The use of multiple medications simultaneously (polypharmacy) increases the risk of adverse drug reactions (ADRs). Therefore, polypharmacy management is a core element of the global challenge for medication safety that was launched by the World Health Organization (WHO) in 2017[1]. By 2050, 22% of the global population will be over 65 years old (double the proportion in 2010). Therefore, managing polypharmacy is especially critical to aging-associated disease management. Due to the exponential rise in complexity of drug-drug interactions for polypharmacy, Artificial Intelligence (AI)-assisted methods are required to improve polypharmacy safety. However, advances in AI are required as healthcare increasingly recognizes the need for interpretable AI. Interpretability is defined as "the ability to explain or to provide the meaning in understandable terms to a human" [2]. Thus, interpretable AI helps users (including

*To whom correspondence may be addressed.

clinicians, patients, drug developers) to identify incorrect AI advice and avert unfair biases [3], and to bridge scientific fields by providing new information to human decision-makers [4].

Thus, in recent years, increasing numbers of interpretable AI methods have been designed to improve the reliability of model predictions [5, 6]. Moreover, explainer methods have been developed to explain the predictions made by particular classes of machine/deep learning models, including even those that are non-interpretable (e.g., an explainer for all graph neural networks [7]) [8, 9]. These explainers can reveal new relations between predictions and features (e.g., new mechanistic associations between genes and ADRs).

In this paper, we demonstrate that explainable machine learning can effectively reveal the mechanisms underlying polypharmacy side effects. We present APRILE (**A**dverse **P**olypharmacy **R**eaction **I**ntelligent **L**earner and **E**xplainer), an explainable framework in the format of predictor-explainer, that leverages pharmacogenomic information for discovering the molecular mechanisms of interactions between drugs. Pharmacogenomic information plays an essential role for predicting and explaining the drug interactions and ADRs. This information includes drug targets, functional associations or physical connections between or within drugs and proteins, all of which can be naturally represented by a knowledge graph (KG). The workflow of APRILE is as follows: first it trains a machine learning model, APRILE-Pred, that is capable of predicting drug combination's side effects from pharmacogenomic data. Then, given a side effect prediction or a set of side effect predictions, APRILE trains another machine learning model, APRILE-Exp. This model produces a tentative molecular mechanism level explanation to those predictions—it queries APRILE-Pred to identify the most important pharmacogenomic information required to make the predictions.

2 APRILE: a general framework to explain polypharmacy side effects

The APRILE framework is shown in Fig 1 and described in the Methods section. It is composed of two modules, APRILE-Pred and APRILE-Exp, for predicting and explaining adverse drug reactions (ADRs) respectively.

APRILE-Pred is inspired by recent successful polypharmacy side effect prediction models [10, 11]. Given a pharmacogenomic KG (Fig 1a), APRILE-Pred learns representation vectors, i.e. node embeddings, for proteins and drugs by extracting knowledge on drug targets and underlying protein function association (Fig 1b). Then, it uses the drug embeddings to predict all possible side effects that are caused by every drug combination (Fig 1c). APRILE-Pred is trained end-to-end to ensure that the embeddings are optimized toward the ADR prediction task (see Methods). The correctly predicted ADRs are then able to be explained by APRILE-Exp.

APRILE-Exp explains a known side effect by deciphering how a APRILE-Pred model makes prediction(s). Given a trained APRILE-Pred model and a prediction as inputs, APRILE-Exp generates an explanation by identifying an optimal KG subgraph such that the side effect(s) is still predicted correctly (i.e., the side effect prediction is the same as when the full KG was used) (Fig 1d).

While explaining, APRILE-Exp revises the edge weights of the KG, which control the information propagated among protein-protein and protein-drug relations in the trained APRILE-Pred. Weights are revised to maximize the accuracy of the input prediction(s) while minimizing the size of the set of involved drug targets and relations between proteins. Basically, APRILE-Exp takes a trained APRILE-Pred model and the side effect prediction(s) as inputs, and returns an explanation in the form of a small group of drug targets and protein interactions. Then, the functional roles of APRILE-Exp's explanations (i.e. a set of drug targets and non-target proteins) are investigated using Gene Ontology (GO) enrichment analysis (see Methods and Fig 1f).

Both APRILE-Pred and APRILE-Exp consist of multiple layers of graph neural networks (GNNs). GNNs recursively incorporate information through the information propagation paths in the graph, leading to the node intermediary embeddings capturing both graph structure and node features. However, these rich embeddings learned by APRILE-Pred prevent APRILE-Exp from finding important information sources on drug targets and non-target proteins, as the pharmacogenomic information and ADRs have been learned into drug embeddings already during the back-propagation. This makes APRILE-Pred's ADR predictions difficult to be explained by APRILE-Exp. Therefore, to build a more interpretable model and reduce explanation redundancy for the side effects that have molecular origins, information on drug targets, protein relations and ADRs should be prevented from being learned in the trainable embedding matrices. Accordingly, we do not update the embedding layer of both protein and drug features in APRILE-Pred during training (Fig 1b,e). This strategy is referred to as lazy training. It improves the downstream explanation by keeping the contribution of relationships between {drug, protein} and protein relatively independent from learned drug representations for making ADR predictions.

Polypharmacy can cause ADRs in a multitude of ways: drugs having opposite effects, acting on the same target protein, or enhancing effects through sequential or complementary effects [12]. As this information is not known from the drug information alone, we require a metric to indicate whether the ADR can be explained by only the drug-target

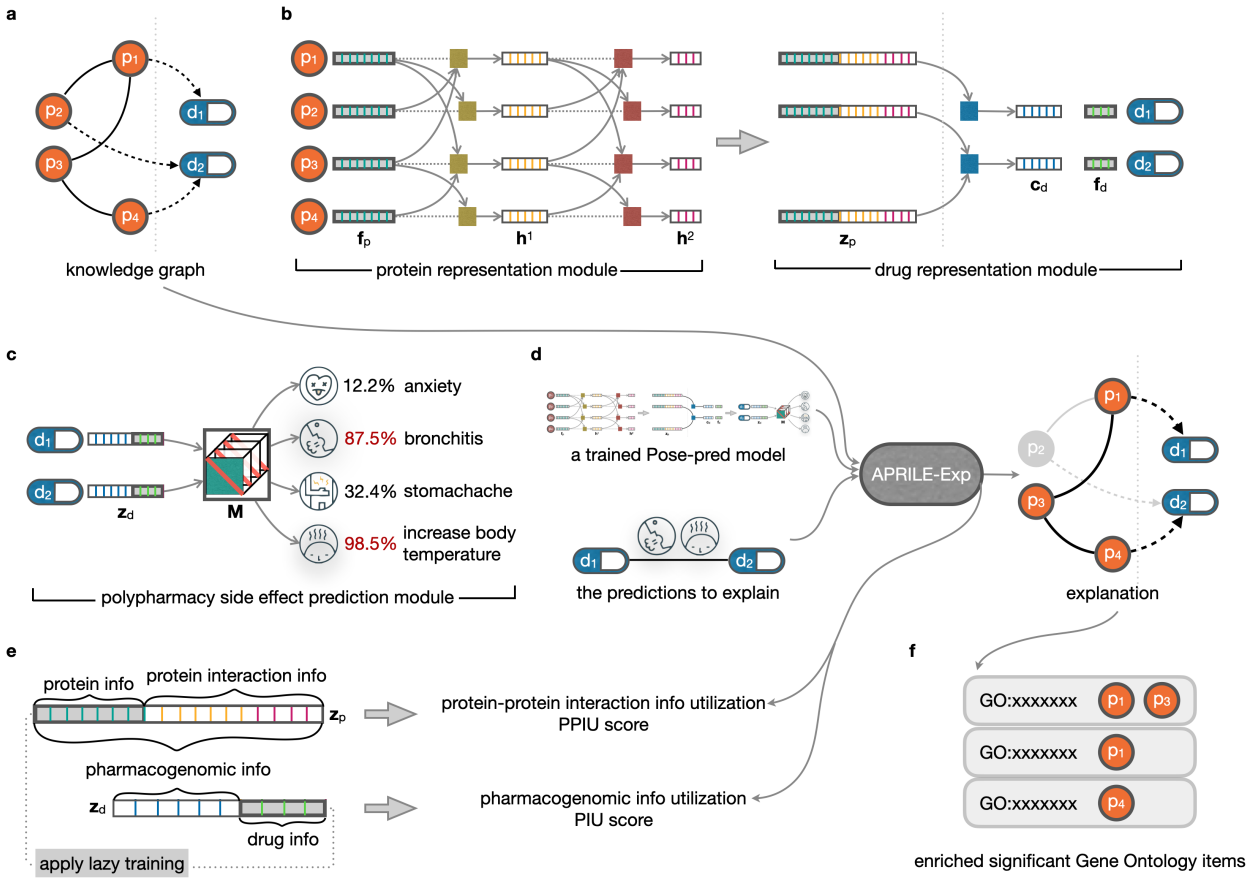

Figure 1: Overview of the APRILE framework - a workflow to predict and explain ADRs **a**, the pharmacogenomic knowledge graph which contains drug-gene and gene-gene relations. **b**, protein and drug representation modules apply graph convolutional neural networks to extract pharmacogenomic information from **a**. **c**, APRILE prediction module uses drug representations to predict the probability of that all drug pairs cause different side effects. The modules in **b** and **c** constitute the predictor of APRILE, APRILE-Pred, and are trained end-to-end with polypharmacy adverse event data. **d**, APRILE-Exp gives a small subgraph of **a** as an explanation for given prediction(s) e.g. polypharmacy adverse event(s), which are the most important drug targets and protein associations in making the given prediction(s), accompanied by functional roles of these proteins. PPIU and PIU score are also provided to measure the utilization of information on drug targets and protein interaction by APRILE-Pred. **e**, lazy training fixes the parameters $\mathbf{f}_p$ and $\mathbf{f}_d$ when they were initialized to prevent the relational information from being learned into the embeddings of protein, drug and side effect. **f**, gene ontology (GO) enrichment analysis is used to investigate the functional roles of the genes in APRILE-Exp's explanations.

proteins or if a larger set of protein-protein interactions is likely to be involved. To address this need, we designed two metrics: pharmacogenomic information utilization (PIU) and protein-protein interaction information utilization (PPIU) (see Methods). These metrics quantify how important the information on drug targets or non-target proteins is respectively for APRILE-Pred's predictions. Therefore, they are used to evaluate APRILE-Exp's interpretability and provide quantitative estimates of the complexity of molecular mechanisms for a side effect (in Section 3.1).

Moreover, the prediction(s) can be either an ADR, which is represented as a triplet of (drug_1, drug_2, side effect), or a set of such triplets. Thus, depending on the prediction(s) we are interested in, APRILE-Exp provides flexible explanations for different research questions. Here, we use APRILE-Pred to discover the mechanisms of an ADR by explaining the corresponding single prediction (in Section 3.2). We then explore the common mechanisms of a side effect, a disease category, or a group of drug interactions by explaining a set of related ADRs simultaneously (in Section 3.3).

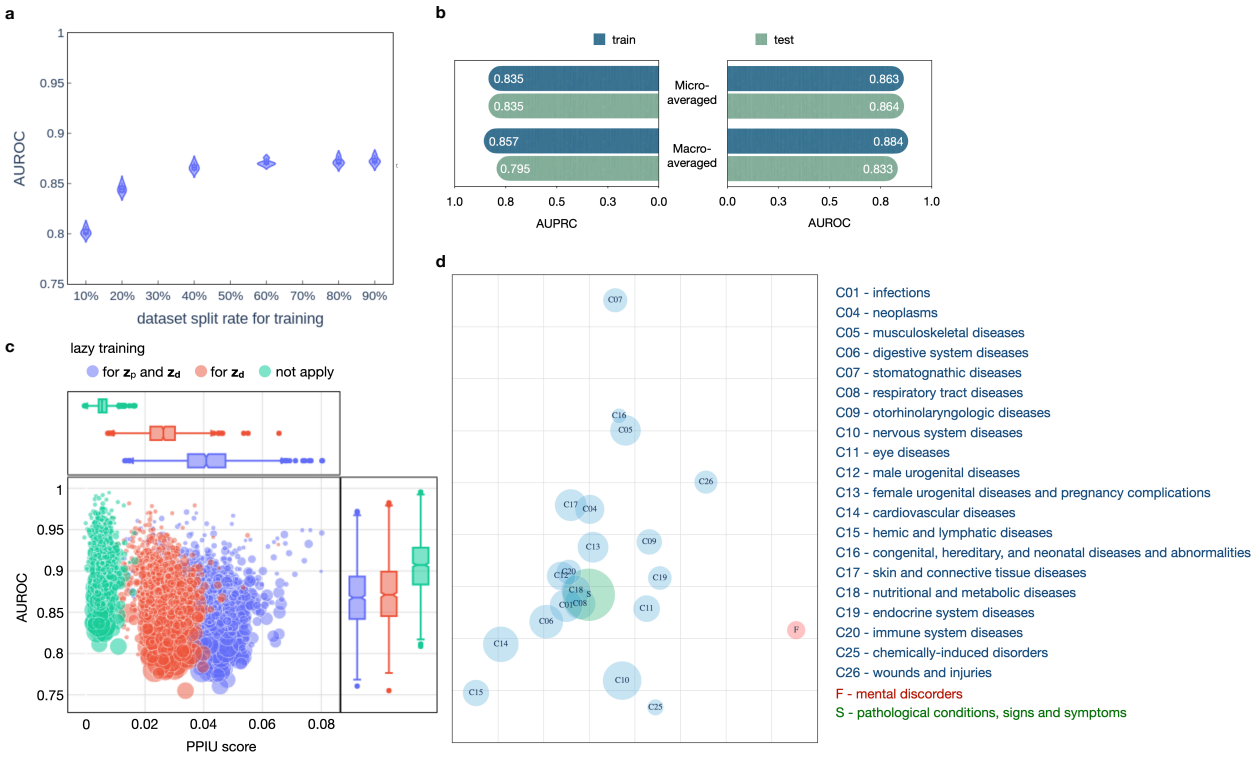

Figure 2: **a**, Sensitive evaluation for lazy training to different dataset split rates and parameter initial values. **b**, performance comparison of the selected trained APRILE-Pred model on the training and testing sets. **c**, ablation study on lazy training, which compares the distribution of AUROC for each side effect when applying no, partial and full lazy training. Marker size indicates the number of drug pairs that cause the corresponding side effect. **d**, UMAP of the category-average side effect embeddings learned by APRILE-Pred. Marker sizes are the log-scaled numbers of side effects in each categories.

3 Results

3.1 Building an explainable graph representation learning model

The interpretability of APRILE-Pred is a fundamental determinant of the effectiveness of APRILE-Exp, as it affects how much drug target and non-target protein information can be directly accessed by APRILE-Exp when explaining the ADR predictions we are interested in. To evaluate this ability, we use PIU and PPIU metrics. As interpretability indicators of model prediction, the PIU and PPIU scores are calculated for each prediction and side effect, and suggest the importance of these two types of information. First, we trained and tested APRILE-Pred on $\sim$4.6 million drug-drug interactions (DDIs), each associated with an ADR. We then computed both PIU and PPIU scores for each ADR and DDI in both training and testing sets (see Methods).

We found that lazy training improves the interpretability of APRILE-Pred (Fig 2c), which means it successfully retains all kinds of relational information relatively independent of the intermediary node embeddings during training. When applying lazy training, the parameters in the feature embedding layer of both proteins and drugs will no longer be updated once initialized (see Methods). Thus, we assessed how sensitive the model is to the initialization of these fixed matrices and the size of samples for training in detail. To this end, we trained the APRILE-Pred models with five different parameter initialization settings and five different dataset splitting rates (the proportion of positive training samples to the whole positive samples for each side effect). The results show that the standard deviation of the AUROC score on the corresponding testing set for different splitting rates is 0.038 ± 0.0004 (see Fig. 2a and Supplementary Table B.1).

The ablation study on lazy training also showed that there is a trade-off between model prediction accuracy and the PPIU score (see Fig. 2c). Sacrificing a little accuracy can significantly improve the potential interpretation of the model. In addition, APRILE-Pred model tends to better predict rare side effects than common side effects (the marker size

indicates the number of positive instances of the corresponding side effects in Fig. 2c). This observation is as expected because APRILE-Pred emphasizes learning from pharmacogenomic information and weakens learning from drug features and direct drug interactions.

Model selection is based on the model's generalization performance. The selected APRILE-Pred model is trained with 90% positive samples, tested with the remaining 10%, and has the most balanced performance on the training and testing set (see Fig. 2b). To examine what it learned, we interrogated the prediction model from side effect to category level and examined its learned embeddings. Based on the Medical Subject Headings (MeSH) database [13], we classify the side effects into three main categories: *disease* (C), *mental disorder* (F03) and *symptom*. We found that a two-dimensional UMAP [14] projection of side effect embeddings learned by the selected trained APRILE-Pred contained meaningful clusters for 20 disease subcategories and the other two categories (see Fig. 2d). Clusters F (mental disorder), C07 (stomatognathic disease) and C13 (female urogenital diseases and pregnancy complications) were located at the extremities, indicating that their embeddings differed the most from other MeSH categories and from each other. Clusters close to each other are interpreted as having similar embeddings: e.g., C16 (congenital, hereditary, and neonatal disease and abnormalities) and C05 (musculoskeletal diseases). However, comparing the Euclidean distance and the Jaccard similarity of the side effects of these 22 side effect categories shows that the overlapping side effects among different categories does not dominate the clustering of these categories (Supplementary Figure A.2). Rather, the embeddings are based on extracting information from integrating the pharmacogenomic knowledge graph with drug adverse events.

The correct predictions are regarded as the knowledge learned or discovered by the model, depending on if the model has seen them during training. The true positive rate of the selected model on both training and testing set is 0.8. This means the model successfully learned 843,318 pieces of ADR knowledge and discovered 93,966 pieces of novel ADR cases. In the next section, we start evaluating the reliability of APRILE-Exp by exploring its explanations for these correctly predicted knowledge pieces, and afterwards investigate its reliability for explaining more complex cases.

3.2 Discovering the toxicity mechanism of combination therapy by explaining single model prediction

When explaining single predictions made by APRILE-Pred, APRILE-Exp answers why the prediction model decides that a pair of drugs causes a side effect. In this section, we focus on exploring these single predictions. We perform a literature-based evaluation of explanation for both learned ADR knowledge and novel predictions based on the selected trained APRILE-Pred model. To this aim, we ask APRILE-Pred to make a prediction for every drug pair and every side effect in the knowledge graph. We then divided these predictions into two groups – learned knowledge and new prediction – according to whether predictions are in the training set or not, and constructed the corresponding ranked lists based on predicted probability for each prediction group. Then, we use APRILE-Exp to obtain a subset of drug targets, protein interactions and significant GO items for each of the twenty highest ranked predictions in both rank lists, and afterwards search biomedical literature to see if we can find supporting evidence. In addition, as PPIU scores indicate the importance of gene interactions for making corresponding predictions, when using APRILE-Exp to explain predictions, we apply auto-tuning to force APRILE-Exp to take non-target proteins into account for the predictions with a PPIU score higher than 0.08. This threshold is determined based on the maximum of all the averaged PPIU score for side effects (see Fig. 2c).

Sepsis. We investigated sepsis as an example of a side effect for which a high probability (> 0.95) was predicted but with alternative explanations having low or high PPIU scores. Explanations with low PPIU (< 0.1) mostly included genes that were direct drug targets (64 genes), except for two genes (HSH2D and SLURP1) (Supplementary Data). Transferrin, which is responsible for transporting iron, was a direct drug target. Alternations in transferrin's ability to transport iron has been observed in sepsis patients [15]. Thus, perturbing transferrin by a drug may increase the chances of experiencing sepsis as a side effect.

Meanwhile, explanations for sepsis with high PPIU (> 0.9) included six target genes and one common non-target gene. The non-target gene CHRNB1 (cholinergic receptor nicotinic Beta 1 subunit) codes for a subunit of the nicotinic acetylcholine receptor (nAchR), which is part of the cholinergic anti-inflammatory pathway. This pathway is a neuroimmune mechanism involving complex cross-talk between the nervous and immune systems [16]. As the two explanations with PPIU> 0.9 also had high PIU (0.45 and 0.67), we also examined the target genes. The six target genes comprised cholinergenic receptor subunits (CHRFAM7A, CHRNA10, CHRNB2, CHRNB4, CHRND) and a choline transporter (SLC5A7). SLC5A7 is also involved in neuromuscular transmission. This example shows that APRILE-Exp identified complex mechanisms of sepsis involving neuroimmune responses.

Literature evidence was found for additional explanations for ADRs in the training set (Table 1). Out of 843,318 predictions (Supplementary Data), we manually searched literature that supported explanations for the top five predictions ranked by prediction score. We found evidence for 2/3 genes for scotoma, one for (CD7) for melanocytic naevus and

none for elevated prostate specific antigen. Indirect evidence was found for bacterial pneumonia and post traumatic stress disorder.

Table 1: Literature evidence for APRILE-Exp explanations for training data.

Side effect	Drug target?	Gene	Supported	Literature evidence
scotoma	target	PDE6B	yes	PDE6B mutations cause Retinitis Pigmentosa (RP) [17]. Scotoma is a symptom for RP patients. The polypharmacy may affect PDE6B similar to mutations.
scotoma	target	PDE6C	yes	PDE6C mutations cause achromatopsia [18]. Scotoma is a symptom of achromatopsia.
elevated prostate specific antigen	target	PDE11A	no	—
melanocytic naevus	non-target	CD7	yes	Patient with melanocytic nevus exhibited low CD7 expression levels [19, 20]
bacterial pneumonia	target	CA4	indirect	CA4 is expressed in the lung and kidney and its inhibition alters CO_2 equilibrium in tissues [21]. In lambs with pneumonia, an insignificant decrease of carbonic anhydrase activity was observed [22].
post traumatic stress disorder (PTSD)	target	PDE6B	indirect	No direct evidence but PDE6B was differentially expressed in hippocampi of alcoholics, potentially affecting cortisol pathway dysregulation, which is observed in PTSD [23].

3.3 Discovering the side effect molecular mechanism by explaining model predictions

When studying the general mechanism of a side effect, APRILE-Exp takes the set of all predictions on this side effect as input and gives the reasons why APRILE-Pred makes these predictions. For the example of pleural pain, the input to APRILE-Exp is $\{(\text{drug}_1, \text{drug}_2, \text{pleural pain}) \in \mathbf{A}^R\}$ (see Methods), which includes all drug pairs causing pleural pain. We do this for every side effect (861 in total) and get a set of important genes, gene-gene interactions and enriched GO items (see Fig 3a and Supplementary Data) We found that 6 GO items: 'GO:0007186', 'GO:0007268', 'GO:0007204', 'GO:0007165', 'GO:0006954', 'GO:0007200' are associated with all side effects. Meanwhile, 250 GO items are associated with only one side effect, while there are 159 side effects with unique GO items.

We then investigated the significance of the explanations by comparing them against disease-gene associations from CTD [24], for 239 side effects that were found in CTD. We computed the precision and Jaccard index based on comparing the sets of genes in the APRILE-Exp explanation versus the disease-associated genes in CTD. We then compared APRILE-Exp accuracy with 10,000 random explanations (i.e., gene sets) (see Methods). We compared two types of explanations: all genes in a APRILE-Exp subgraph, or the subset of these genes that are in enriched GO terms. Additionally, two types of random models were tested: (A) one that preserves the number of drug target vs. non-target genes, and (B) one that samples from all genes in the KG.

Overall, APRILE-Exp explanations for 97% (232/239) of side effects were better than 10,000 random explanations ($P < 0.05$) (Supplementary Data). On average, the precision of APRILE-Exp explanations was 33% to 55% higher than random models, with the in comparing GO-enriched explanations versus drug target-agnostic random models (random model B).

For the 5 to 7 side effects where $P > 0.05$, the number of disease-associated genes in CTD was typically small with most < 60, one case having 629 genes, and one case having 10,620 genes. For cases with very few CTD genes, we propose that APRILE-Exp explanations may contain new hypotheses for side effect-gene associations. In the last case (proctitis), the random models had a mean precision of 0.5. Thus, by randomly sampling genes from the KG, half are expected to be in the disease-associated gene list. This case suggests that proctitis may be highly complex, involving a multitude of genes, which makes a mechanistic explanation particularly challenging.

Next, we applied affinity propagation [25] to cluster these side effects, based on the negative squared Euclidean distances among the side effects (using multi-hot representation based on their associated GO items). We then tuned the input preferences to minimize the number of clusters until all clusters contain more than one side effect. Finally, we obtained 25 side effect clusters in total (Fig. 3b).

We hypothesized that side effects clustered by biological function may also be related by MeSH disease categories. Enrichment analyses of the 25 clusters for 22 MeSH categories (see Methods) was used to rank order clusters by odds ratio, subject to an unadjusted enrichment $P < 0.05$ (Supplementary Figure A.1). The top 5 enriched clusters and their MeSH categories were panic disorder-mental disorders, embolism pulmonary-hemic and lymphatic diseases, cough-respiratory tract diseases, palpitation-cardiovascular disease, and incisional hernia-eye diseases. While all five clusters had unadjusted $P < 0.05$ and odds ratio> 4, FDR-adjusted P all exceeded 0.05 (Supplementary Data).

Therefore, we concluded that mechanisms based on biological processes may explain mechanistically related side effects in a meaningful way that complements conventional disease categorization based on MeSH. We thus performed GO enrichment (see Methods) of each cluster and found 1 to 124 enriched GO biological processes (FDR-adjusted $P < 0.05$) and odds ratio > 4.34 (the median odds ratio across all clusters) (Supplementary Data). We further identified 0 to 64 GOs that were uniquely enriched for each cluster.

The bursitis cluster includes 20 side effects including breast cancer, sleep apnea, renal cyst, and gastroduodenal ulcers. These side effects span 10 diverse disease categories. Only one GO term was uniquely enriched: aminergic neurotransmitter loading into synaptic vesicle. In the context of breast cancer, drugs that dysregulate aminergic neurotransmitters affect immune function and cancer progression [26]. Clustering of the side effects sleep apnea and muscle disorder was also supported by literature: aminergic neurotransmitter factors or imbalances can cause sleep apnea [27], or masticatory muscle disorder [28]. Thus, non-intuitive, mechanistic similarities were captured by APRILE-Exp.

The panic disorder cluster includes 54 side effects spanning 19 disease categories, including mental disorders, nervous system diseases, and cardiovascular diseases (Supplementary Data). It is enriched for 11 unique GO terms, including GOs related to regulation of signal transduction, regulation of ion transmembrane transporter activity, tryptophan catabolism, and fatty acid metabolism and nucleoside salvage (Supplementary Data). Consistent with this clustering, tryptophan metabolism is disturbed in cardiovascular disease patients [29] and also in panic disorder [30]. In both cases, cytokine-based signal transduction alterations cause disrupted tryptophan metabolism.

3.3.1 Case Study I: Infectious Diseases

Global analysis. In this case study, we show how to use APRILE-Exp to explore mechanisms of side effects which are infectious diseases. In the KG, there are 67 side effects that are infectious diseases. We ran APRILE-Exp on all the predictions involving these side effect with a confidence threshold of 0.9, to identify common mechanisms among them. We identify 1373 enriched significant Gene Ontology (GO) terms that are common to all these infections (Fig 3c). Of the genes involved in these GO terms, we investigated two further: CXCR4 and CXCL12.

CXCR4 appears frequently (51 GO terms) in enriched GO terms and is co-enriched with CXCL12 in the majority of cases (33 GO terms). CXCR4 is a cellular chemokine receptor that plays central roles in development, hematopoiesis, and immune surveillance through signaling induced by its ligand, CXCL12. Several viruses have been found to modulate CXCR4 expression or alter its functional activity, with direct effects on cell trafficking, immune responses, cell proliferation, and cell survival [31]. The cellular component GO term, endoplasmic reticulum membrane, was frequently (24/50, 48%) enriched for infection side effects: The ER is the nexus of innate and adaptive immune response pathways. The ER activates multiple signalling pathways to maintain protein homeostasis (e.g., proper folding), which if unresolved contributes to inflammation. Thus, the ER (stress) can explain multiple side effects involving inflammation, and pain induced by inflammation.

Pleural pain. We next investigated the side effect, pleural pain. In particular, we investigated new drug combinations that were not in the training set but predicted with score > 0.99. Hence, the explainability of this high-confidence prediction is of strong interest. APRILE-Pred predicted that pleural pain is a polypharmacy side effect of a new drug pair: omeprazole (an H+/K+ ATPase-binding proton-pump inhibitor) and gabapentin (a VGCC blocker).

According to APRILE-Exp, pleural pain was enriched for the GO terms voltage-gated calcium channel (VGCC) complex (cellular component) and VGCC activity (molecular function). Voltage-gated calcium channels (VGCC) are required to generate and transmit pain via neurons. Despite gabapentin being used to treat pain by blocking VGCC, its interaction with omeprazole is predicted to cause pleural pain. APRILE-Exp suggests that the drug interaction causes VGCC blocking activity to be altered. Indeed, blocking of Na+/K+ ATPase activity disrupts the sodium gradient. Consequently, the sodium-calcium exchange system is disrupted, leading to intracellular calcium accumulation. Excess intracellular calcium can be toxic to neurons, and even associated with seizures. This example clearly demonstrates that APRILE-Exp can explain ADRs by identifying key biological mechanisms of drug interactions. Pleural pain is an example of an ADR where individual drug effects are altered when combined. Here, inhibiting the respective drug targets causes imbalance between two ion concentrations, leading to an ADR (Fig. 4a).

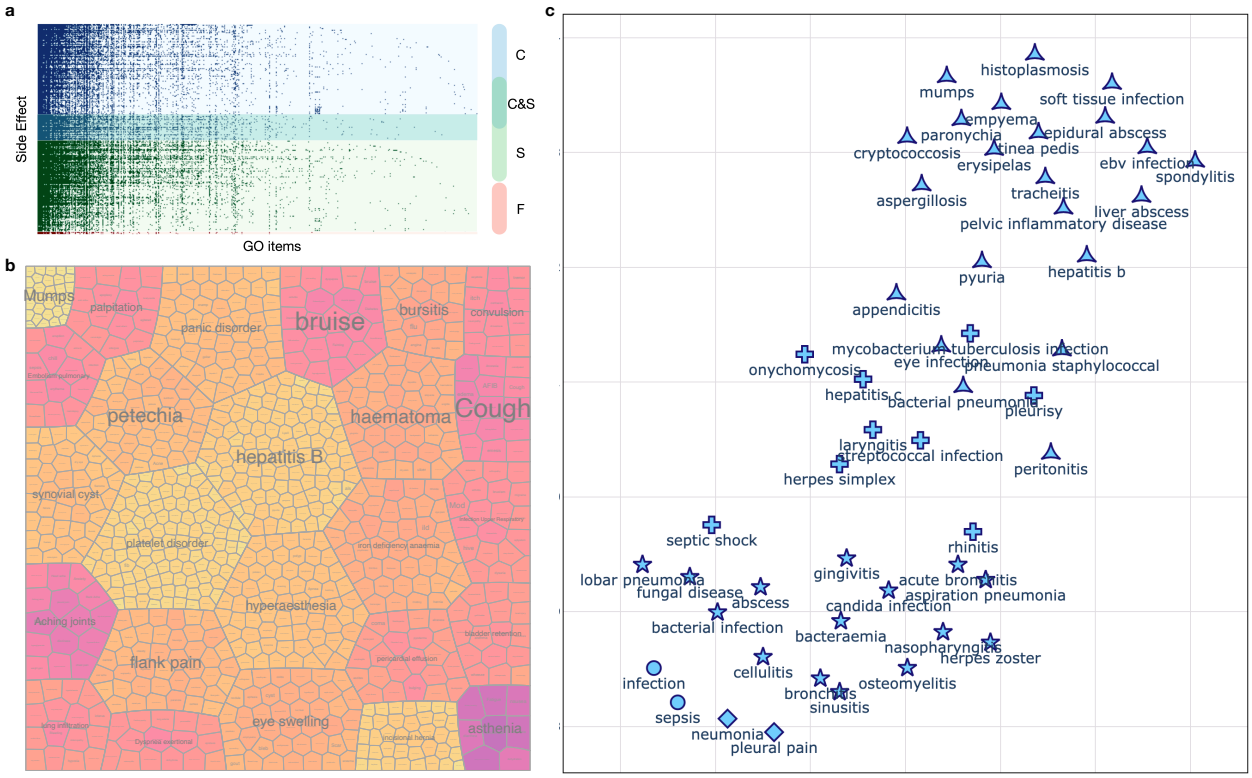

Figure 3: **a**, Association matrix for 861 side effects and 1,787 biological process GO items. C, S, F are side effect categories are the same as those in Fig 2d. **b**, side effect clustering based on the APRILE-Exp's explanation. The name of the example for each cluster (25 in total) are represented in large fonts. Each small cell is a side effect, and its colour indicts the number of GO items associated with it. (See sup figure for details). **c**, clustering side effects in the infection disease category based on associated GO items. The side effects in the same cluster use the same marker symbol.

3.3.2 Case Study II: Mental Disorders

Anxiety. We sought explanations for anxiety-causing drug pairs that include nicotine. These ADRs are interesting as nicotine can either cause or reduce anxiety depending on the anxiety model tested. Additionally, nicotine treatment can both desensitize and activate nAChRs.

In one example, ondansetron was paired with nicotine. Nicotine is an agonist of (activates) CHRNA7. Meanwhile, validated targets of ondansetron include 5-hydroxytryptamine receptors, cytochrome P450 proteins, and mu-type opioid receptors. It is also predicted to bind to CHRFAM7A[32]. CHRFAM7A is a chimeric gene, partially duplicated from CHRNA7, the human alpha 7 nicotinic acetylcholine receptor (nAChR). CHRFAM7A itself is a negative regulator of CHRNA7. Additionally, APRILE-Exp identified two genes to be important despite not being direct targets of either drug: Lynx-1 and SLURP-1, which are both endogenous allosteric modulators of nAChRs. SLURP-1 inhibits alpha 7 nAChR activity (ion influx through the receptor channel). Similary, Lynx1 inhibits nAChRs, by binding to them directly and reducing their sensitivity to acetylcholine.

Because both drugs activate CHRNA7, we hypothesize that polypharmacy of nicotine and Ondansetron induces nAChR hypersensitivity, resulting in increased anxiety. Consistent with this hypothesis, hypersensitive nAChR has shown to increase anxiety in mice [33]. A simple description of the mechanism is that multiple drug-protein and protein-protein activities combine in a new context, resulting in the polypharmacy side effect. Here, we propose that hypersensitivity to polypharmacy arises as the respective drug target proteins also inhibit a common protein that is not directly targeted by either drug (Fig. 4b).

Panic disorder. APRILE-Exp revealed that three drugs could cause panic disorder in any pairing of the three: Fexofenadine, Hydroxyzineand Loratadine.All three drugs are inverse agonists that reduce histamine receptor H1 (HRH1) activity. They do so by binding to specific sites of HRH1 to stabilize the inactive form of HRH1 [34]. HRH1 is a G protein-coupled receptor expressed in over 20 tissues or organs, including the heart. Of HRH1's many functions

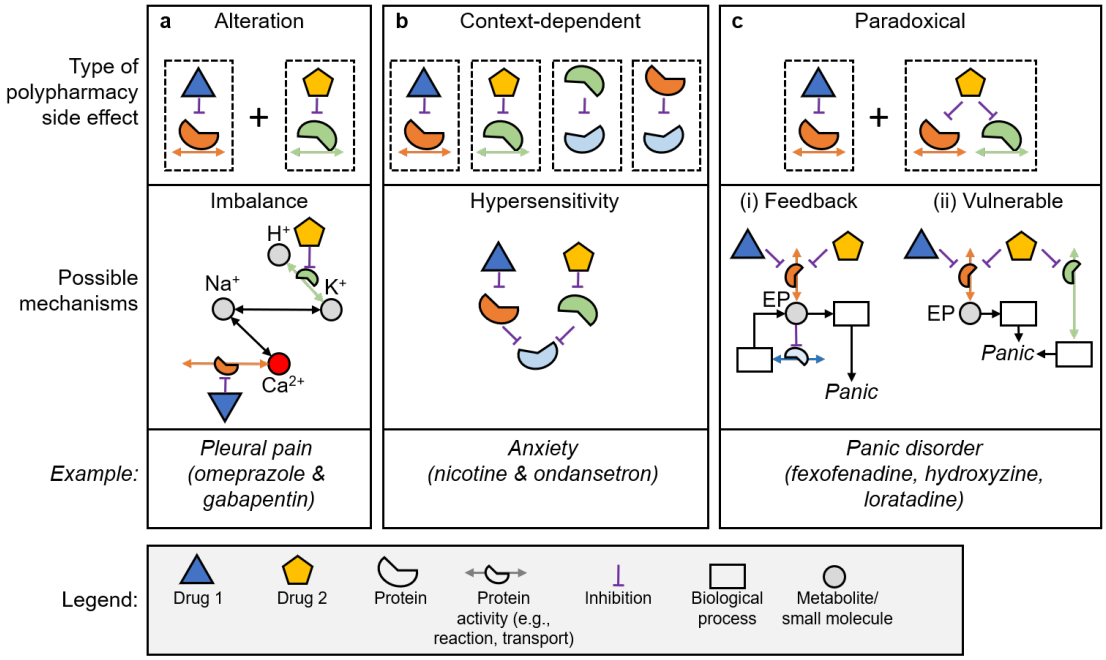

Figure 4: Mechanisms of ADRs in polypharmacy. **a**, polypharmacy alters drug effects resulting in molecular imbalance. **b**, drug effects change with context, such as drug targets inhibiting a common non-target protein. **c**, paradoxical ADRs may be caused by (i) inverse response within feedback regulatory processes, or by increased (ii) vulnerability to disruption of a biological process due to the action of a second drug.

[34], its mediation of catecholamine secretion from the adrenal medulla is apparently relevant to panic disorder. Panic disorder patients release epinephrine (a catecholamine) from the heart, possibly due to uptake of the epinephrine secreted at higher rates during panic attacks [35]. Alternatively, epinephrine may be synthesized directly by the heart PNMT (phenylethanolamine-N-methytransferase) activity [36].

Thus, we hypothesize that two inverse agonists (deactivators) of HRH1 interact, having the opposite effect activating HRH1, which increases release of epinephrine from the adrenal medulla. This hypothesized mechanism is an example of a paradoxical adverse reaction [37], wherein the drug effect is opposite to the expected outcome. Paradoxical adverse drug events are common. Up to a quarter of epilepsy patients experience increased seizures when treated with antiepileptic drugs (AEDs) at excessive dosages [38]. AEDs used to treat mood disorders can have the paradoxical adverse event of causing depression [38].

The mechanisms for the paradoxical side effect includes pharmacological effects that appear only at high drug dosages, blockade of inhibitory pathways leading to release of excitation, and indirect effects mediated by electrolyte disturbances or impaired alertness, in the case of seizures.

Combining AEDs acting by the same mechanism can cause adverse pharmacodynamic interactions [38], in which additive or super-additive effects may occur. Thus, we hypothesize that the mechanism for the paradoxical side effect of panic disorder from polypharmacy of anxiolytic (anxiety reducing) drugs may be similar to that of excessive dosage of a single anxiolytic drug. The exact mechanism of the paradoxical adverse reaction caused by pairs of the three drugs needs examined here needs to be investigated further. The value of APRILE-Exp is that it identified the side effects for which gene interactions are likely to be important for explaining the interaction. Furthermore, the low PPIU score suggests that the direct drug targets may be sufficient to explain the side effect.

Based on this guidance, we hypothesize that interactions between multiple known targets may cause the side effect. Namely, the interaction between HRH1 and KCNH2. Loratadine and Fexofenadine both inhibit KCNH2 (potassium voltage-gated channel subfamily H member 2). Inhibiting KCNH2 can prolong the action potential [39], which is consistent with increased metrics of action potential variability in panic disorder patients: QT dispersion, Tp-e (time between T wave peak and end), and Tp-e/QT ratio). Furthermore, all of these metrics were positively correlated with the severity of panic disorder [39].

Thus, we hypothesize that patients become more vulnerable to panic disorder due to increased QT dispersion caused by KCNH2 inhibition. In this vulnerable state, panic disorder may be caused if the antihistamine drugs increase epinephrine transport. Alternatively, NET (norepinephrine transporter) may be inhibited to reduce re-uptake of norepinephrine (NE), which is required for normal heart beat pulses. Indeed, reduced reuptake of NE was observed in patients with panic disorder [40], and NET was part of a larger APRILE-Exp explanation involving pairs from 10 possible drugs (Extended Data Fig A.3).

Overall, panic disorder is an example of a paradoxical ADR. We propose two broad mechanisms: (i) feedback, or (ii) vulnerability (Fig. 4c). For feedback, the drugs share a common target, whose abundance is regulated in one or more loops. Excessive inhibition by polypharmacy causes an inverse response, leading to panic attack. For vulnerability, polypharmacy disrupts a biological process (e.g., action potential) that increases vulnerability to another biological process being disrupted by the second drug (epinephrine release or NE reuptake).

3.3.3 Case Study III: Hypophosphatemia

Hypophosphatemia is a metabolic disease referring to low serum phosphate level (< 2.5mg/dL) [41]. Despite being only about 0.5% of the total phosphorus in the body, inorganic serum phosphate plays an essential role in keeping phosphorus homeostasis [42]. The shortage of serum phosphate usually leads to muscle weakness, reduced immunity, cardiac and respiratory failures, etc. [43, 44]. The cause of hypophosphatemia consists of three main factors: insufficient intake, increased excretion, and internal redistribution of phosphorus [45]. Although hypophosphatemia is not common in the general public, it is more prevalent in individuals with sepsis, alcohol addiction, and patients in the intensive care unit, leading to a severe worsening of their health conditions [46, 47] Therefore, predicting and preventing drug-induced hypophosphatemia is vital for patients' health and wellness. We used APRILE-Exp to predict the likelihood and mechanism of hypophosphatemia caused by the drug-drug interaction between glipizide and lansoprazole.

Glipizide is a second-generation oral hypoglycemic agent [48]. Its only indication is type 2 diabetes. It works by stimulating pancreatic beta cells to increase insulin production and reduce the clearance of insulin [48, 49]. The primary gene targets are ABCC8, PPARG, and the CYP family [48].

Lansoprazole is a proton pump inhibitor drug (PPI) commonly used to reduce gastric acid and treat gastric ulcers [50]. Main targets include ATP4A, MAPT, CYP family, ABCG2, and PHOSPHO1. APRILE-Exp was executed on the triplet 'glipizide' - 'lansoprazole' - 'hypophosphatemia' with a regulation score of 1.0 and the threshold probability of 95%. Based on the predictor and explainer, two drugs have a 96.95% chance of causing a combined adverse effect of hypophosphatemia.

Based on APRILE-Exp, glipizide targets gene CYP3A4, which interacts with gene ABCB11 and gene SLCO1B3. Lansoprazole has only one relevant target of the gene PHOSPHO1. Dropping gene CYP3A4 results in the lowest probability of 94.8%, followed by dropping PHOSPHO1 and ABCB11, which results in a probability of 96.0% and 98.3%, respectively. Dropping SLCO1B3 has no significant impact on predictive probability.

The GO enrichment analysis shows three relevant GO terms: 'metabolic drug process,' 'bile acid and bile salt transport,' and 'bile acid transmembrane transporter activity.' Firstly, bile acid transfer and relocation appear in the GO enrichment due to genes ABCB11 and SCLO1B3, which regulate bile acid transfer proteins, appearing in the subgraph. However, repeating the analysis with these genes excluded did not significantly impact the prediction score. Therefore, we focused instead on the remaining GO terms.

PHOSPHO1 is largely involved in hydration decompositions of choline phosphate and ethanolamine phosphate. These reactions will produce inorganic phosphate from organic compounds, which is a part of the phosphorus cycle. High-level lansoprazole inhibition of PHOSPHO1 disturbs the cycle, causing a drop in serum phosphate level. Lansoprazole is metabolized by the liver using the P450 pathway, a common drug and xenobiotic metabolic pathway. Glipizide inhibits CYP3A4, which is a member of the P450 pathway, causing a delayed excretion and increased exposure to lansoprazole [51]. The excess lansoprazole greatly reduces the amount of serum phosphate produced, leading to hypophosphatemia. This mechanism is considered the primary cause of the drug pair adverse effect and can be verified by the APRILE-Exp result, which shows that CYP3A4 and PHOSPHO1 are critical genes for the reverse effect prediction.

Another possible mechanism arises from glipizide's indication. Insulin therapies may lead to a redistribution of phosphate due to insulin-stimulated cells actively acquiring phosphate from the blood [52]. Glipizide functions by inhibiting gene ABCC8 from promoting insulin production by the liver [48]. Thus, it may have the same effect as direct insulin therapies. This mechanism is not sufficient to cause hypophosphatemia but may worsen an existing condition [41]. In addition, people who take glipizide are likely to suffer from diabetes. Diabetes patients have more renal phosphate excretion than average due to osmotic diuresis [42]. This may also contribute to an existing hypophosphatemia adverse effect. Overall, APRILE-Exp helped delineate multiple hypothesized mechanisms for hypophosphatemia as an ADR (Supplementary Figure A.4).

3.3.4 Case study IV: peptic ulcers

Duodenal ulcer, gastric ulcer and esophageal ulcer are all peptic ulcers that affect the gastrointestinal system. We ran APRILE-Exp on each of these three ulcers to identify common mechanisms. Fourteen enriched GO terms were common to all three of these diseases. The GO terms comprised 7 biological processes, 6 molecular functions, and 1 cellular component. We manually reviewed literature and found supporting evidence for these enriched GO terms (Supplementary Table B.2). As a result, we could delineate the following key mechanisms (we note the enriched GO terms in parentheses below).

First, *Helicobacter pylori* infection is a common cause of ulcers [53]. In fact, *H. pylori* infects the gastric epithelium of over half the world's population [53]. Of these people, nearly 20% manifest peptic ulcers or gastric ulcers [53]. Guided by APRILE-Exp explanations, we found that *H. pylori* virulence is affected by heme hinding (GO:0020037) proteins [54], and is associated with lowered iron binding (GO:0005506) capacity [55]. Further, *H. pylori* infection is associated with dyslipidemia–i.e., elevated regulation of cholesterol biosynthesis (GO:0045540).

Second, metabolic processes including arachidonic acid metabolism, geranyl diphosphate biosynthesis, and phospholipase A2 activity (GO:0004623) play a role in breaking or maintaining the mucosal barrier, which normally defends against infection [56]. Additional autophagic (i.e., self-degradation) processes also impact *H. pylori* infection [53]. These processes involve formation of autophagosome vesicles, which nucleate at the endoplasmic reticulum (GO:0005789) [53]. Overall, APRILE-Exp helped delineate mechanisms involving both pathogen and host factors that were common to three peptic ulcers.

4 Discussion

Here, we developed a predictor-explainer framework, APRILE, to predict and explain the molecular mechanisms of adverse drug reactions caused by polypharmacy.

PPIU indicates mechanism complexity. APRILE-Exp identified mechanisms of sepsis involving both direct drug targets and broader gene interactions, including neuroimmune responses. Furthermore, the utility of PPIU scores is evident: the complex mechanism (including genes that are not drug targets) was consistent with a high PPIU, whereas the direct (drug target-based) mechanism was consistent with a low PPIU. Therefore, PPIU scores can indicate when complex mechanisms may underlie a side effect.

APRILE-Exp provides significant and accurate explanations. We found that APRILE-Exp explanations were significantly consistent with CTD disease-gene associations [24] for 97% of 239 side effects, with 33% to 55% higher precision compared to 10,000 random explanations ($P < 0.05$). This result held even when random explanations preserved the proportion of drug-target vs. non-target genes. Thus, APRILE-Exp explanations are highly likely to comprise validated drug-protein and protein-protein interactions.

APRILE-Exp delineates non-intuitive polypharmacy side effect mechanisms. To illustrate how APRILE-Exp can be used by healthcare researchers, we used APRILE-Exp to guide our investigation into the mechanisms underlying polypharmacy side effects spanning four disease categories: infection, mental disorders, metabolic disease,and gastrointestinal disease. APRILE-Exp allowed us to formulate hypothesized mechanisms for non-intuitive polypharmacy side effects. In each case, we found supporting evidence in literature. For example, nicotine can both cause or reduce anxiety and either desensitize or activate its drug target (in a context-dependent way). APRILE-Exp narrowed down the most likely mechanism to be hypersensitivity of the nicotinic acetylcholine receptor, while involving the effects of two non-drug target genes. Furthermore, panic disorder was a paradoxical side effect, where polypharmacy induces the opposite effect of using each drug alone. Among many possible mechanisms, APRILE-Exp explanations were most consistent with two possible mechanisms that were also supported by literature evidence (Fig. 4). These molecular mechanisms directly guide the design of validation experiments. These case studies illustrate that APRILE-Exp can guide experiment design to pinpoint molecular mechanisms underlying non-intuitive (even paradoxical) polypharmacy side effects.

Lazy training encourage APRILE-Pred's interpretability. The application of lazy training significantly improves the overall utilization of protein interaction information in the APRILE-Pred models. On the one hand, lazy training improves the interpretability of the APRILE-Pred and benefits the optimization process in APRILE-Exp, on the other hand, it enables GCNs to maintain the independence of relational information. Therefore, lazy training is essential to find the information propagation paths that are important for APRILE-Pred to make predictions.

The design of loss in APRILE-Exp provides meaningful and flexible explanations. There are three items in the APRILE-Exp's loss, which are used to encourage APRILE-Exp to make correct prediction(s) when predicting the interested polypharmacy adverse event(s) based on a small subgraph of pharmacogenomic knowledge graph. In addition,

two hyper-parameters are designed to control the contribution proportion of drug targets and non-target information, which can be adjusted with reference to the PIU and PPIU scores. We also shared ideas on how to design different items for the explainer's loss. Exploring the loss function designing space may help to find more meaning drug targets and non-targets. These ideas may also help adapt GCN-based explainers [7] to different application scenarios.

This work presents a first step toward Artificial Intelligence (AI)-guided polypharmacy side effect explanation, as shown here by ample, domain-specific literature evidence of AI explanations. APRILE-Exp's high interpretability makes it useful when designing safe polypharmacy interventions, whose complex side effects can only be delineated with assistance from AI but whose predictions must also be trustworthy. Thus, interpretable AI such as APRILE-Exp will play an increasingly critical role for polypharmacy management in an aging population (22% over 65 years by 2050 [1]), and where polypharmacy is currently the norm (75% of people 65 and older on two or more medicines [1]).

5 Acknowledgements

This work was supported by Queen's University, and the Natural Sciences and Engineering Research Council of Canada (NSERC) [RGPIN-2020-06325]. This research was enabled in part by support provided by Compute Ontario (www.computeontario.ca) and Compute Canada (www.computecanada.ca). Computations were performed on resources and with support provided by the Centre for Advanced Computing (CAC) at Queen's University in Kingston, Ontario. The CAC is funded by: the Canada Foundation for Innovation, the Government of Ontario, and Queen's University.

6 Methods

6.1 Pharmacogenomic Knowledge Graph Construction

We use two subdatasets for polypharmacy-associated ADR modelling from the BioSNAP-Decagon dataset [57] to prepare the pharmacogenomic knowledge graph : **PP (protein-protein)**: relations between proteins depend on whether they have functional association or physical connections. **GhG (drug-gene)**: relations between drugs and their targeted proteins. In summary, the pharmacogenomic knowledge graph contains 19,365 vertices (284 drugs (D) and 19,081 proteins (P)), 1,449,820 edges. Among the edges, there are 1,431,224 protein interaction edges (P-P) and 18,596 drug protein targeting edges (P-D). We use the following data structures to represent the graph we constructed.

- $\mathbf{A}^p \in \mathbb{R}^{|P|\times|P|}$: symmetric adjacency matrix for the undirected subgraph within proteins.
- $\mathbf{A}^t \in \mathbb{R}^{|P|\times|D|}$: adjacency matrix for the directed bipartite subgraph between the proteins and drugs, whose edges are directed from proteins to drugs.

6.2 APRILE-Pred

APRILE-Pred is a Graphical Neural Network (GNN) model that takes protein-protein and protein-drug interaction graphs as input and predicts the side-effects caused by drug combinations. Specifically, we first learn the protein representation in the protein interaction subgraph, and then generate the drug representation through the protein-drug interaction subgraph with drug attributes (See Algorithm 1). In addition, The input drug/protein node features/attributes $\{\mathbf{v}_{d/p}|d \in D, p \in P\}$ are set up to one-hot vectors, which are binary vectors representing the ith drug/protein if its ith element is 1. Protein and drugs are indexed respectively.

6.2.1 Learning Drug Representation

Firstly, we use a 2-layer graph convolutional neural network (GCN) [58] to capture protein attributes and relations on the protein interaction subgraph $\mathbf{A}^p$. The input is the protein node features $\mathbf{h}_p^0 = \mathbf{v}_p, \forall p \in P$. For each protein node $p \in P$, the relation between two hidden layers is given by

$$\mathbf{h}_p^{k+1} = \text{ReLU}(\sum_{p' \in \mathcal{N}_p} \frac{1}{C_{p,p'}} \mathbf{W}_P^k \mathbf{h}_{p'}^k), \tag{1}$$

where $\mathcal{N}_p$ denotes the union of protein node p and its neighbors on the subgraph $\mathbf{A}^p$, $C_{p,p'} = |\mathcal{N}_p||\mathcal{N}_{p'}|$ is coefficient calculated by the product of the degree of p and $p' \in \mathcal{N}_p$, and $\mathbf{W}_P$s are parameters. The final learned protein embeddings/representation $\mathbf{z}_p = \mathbf{h}_p^1 \oplus \mathbf{h}_p^2, \forall p \in P$ are the concatenation of the outputs of all the GCN layers.

Then, we use a graph-to-graph information propagation module [11] to transform learned protein embeddings and drug node features into feature representation of drugs via the bipartite protein-drug integration subgraph $\mathbf{A}^t$. The protein's

contribution to the drug d's representation for all $d \in D$ is:

$$\mathbf{c}_d = \text{ReLU}(\frac{1}{C_d} \sum_{p' \in \mathcal{N}_d} \mathbf{W}_T \mathbf{z}_{p'}), \tag{2}$$

where $\mathcal{N}_d = \{p' | p' \in P, \mathbf{A}_{p',d} = 1\}$ is the set of drug d's protein neighbours, $C_d = |\mathcal{N}_d|$, and $\mathbf{W}_T$ is a trainable parameter. Besides, A linear transformation followed by an activation function is used to map the original drug node features $\mathbf{v}_d$ into the same space that protein's contributions $\mathbf{c}_d$ live in for each drug $d \in D$:

$$\mathbf{f}_d = \text{ReLU}(\mathbf{W}_F \mathbf{v}_d). \tag{3}$$

The final representation of drug $d \in D$ is: $\mathbf{z}_d = \mathbf{z}_p \oplus \mathbf{c}_d \oplus \mathbf{f}_d$, where $\oplus$ denotes concatenation.

6.2.2 Polypharmacy Side Effect Prediction

We predict if drug combinations can cause any side effect. Here, DistMult factorization [59] is used as the scoring function, because it is a well-known model that models rich relationships well on standard benchmarks alone or as a decoder. The probability that drug pairs $(d, d'), \forall d, d' \in D$ cause side effects $r, \forall r \in R$ is obtained by acting the sigmoid function on the score tensor $\mathbf{G} = \{g^r_{d,d'}\} \in \mathbb{R}^{|D| \times |D| \times |R|}$ computed by the scoring function:

$$p^r_{d,d'} = \sigma(g^r_{d,d'}) = \sigma(\mathbf{z}_d^\top \mathbf{M}_r \mathbf{z}_{d'}), \tag{4}$$

where $\mathbf{M}_r$ is a trainable diagonal matrix which is associated with the side effect $r \in R$. We call it a single-instance prediction or a single prediction. In addition to calculating the probability that the drug pair (d, d') causes the side effect r, the proportion of variant protein related information are also evaluated leading to the model to make such a prediction: pharmacogenomic information utilization (PIU) and protein-protein interaction information utilization (PPIU) score:

$$\text{PIU}^r_{d,d'} = 1 - \frac{p^r_{d,d'}(\mathbf{z}_d^{PIU})}{p^r_{d,d'}(\mathbf{z}_d)}, \tag{5}$$

$$\text{PPIU}^r_{d,d'} = 1 - \frac{p^r_{d,d'}(\mathbf{z}_d^{PPIU})}{p^r_{d,d'}(\mathbf{z}_d)}, \tag{6}$$

where $\mathbf{z}_d^{PIU} = \mathbf{z}_p \oplus (\mathbf{c}_d \cdot \mathbf{0}) \oplus \mathbf{f}_d$ and $\mathbf{z}_d^{PPIU} = (\mathbf{z}_p \cdot \mathbf{0}) \oplus \mathbf{c}_d \oplus \mathbf{f}_d$.

6.2.3 APRILE-Pred Training Details

Dataset. We use the database of polypharmacy side effects compiled by the BioSNAP-Decagon dataset [57]. This database consists of 964 polypharmacy side effects from 4,651,131 drug-drug interactions, which is a subset of the TWOSIDES database [60]. These polypharmacy side effects already exclude side effects attributed to an individual

Algorithm 1: APRILE-Pred's Forward Propagation

Input: $\mathbf{V}_p$, $\mathbf{V}_d$, $\mathbf{A}^p$, $\mathbf{A}^t$
Parameter: $\mathbf{W}_P$, $\mathbf{W}_T$, $\mathbf{W}_F$
Output: $\mathbf{P}$

$\mathbf{H}^0 \leftarrow \mathbf{V}_p$
for $l \in L_p$ **do**
  $\mathbf{H}^{l+1} \leftarrow \text{ReLU}(\ GCN^l_{pp}(\mathbf{H}^l, \mathbf{A}^p;\ \mathbf{W}^l_P)\ )$
end
$\mathbf{Z}_p \leftarrow [\mathbf{H}^0, ..., \mathbf{H}^L]$
$\mathbf{C} \leftarrow \text{ReLU}(\ GCN_{pd}(\mathbf{H}, \mathbf{A}^t;\ \mathbf{W}_T)\ )$
$\mathbf{F} \leftarrow \text{ReLU}(\mathbf{V}_d \mathbf{W}_F^T)$
$\mathbf{Z}_d \leftarrow [\mathbf{Z}_p, \mathbf{C}, \mathbf{F}]$
$\mathbf{P} \leftarrow Sigmoid(\mathbf{Z}_d^T \mathbf{M} \mathbf{Z}_d)$

Algorithm 2: APRILE-Exp's Forward Propagation

Input: $S = \{(d_n, d'_n, r_n)\}_n$, $\mathbf{V}_p$, $\mathbf{V}_d$, $\mathbf{W}_P$, $\mathbf{W}_T$, $\mathbf{W}_F$
Parameter: $\hat{\mathbf{A}}^p$, $\hat{\mathbf{A}}^t$
Output: $\mathbf{P}_S$

$\mathbf{H}^0 \leftarrow \mathbf{V}_p$
for $l \in L_p$ **do**
  $\mathbf{H}^{l+1} \leftarrow \text{ReLU}(\ GCN^l_{pp}(\mathbf{H}^l, \hat{\mathbf{A}}^p;\ \mathbf{W}^l_P)\ )$
end
$\mathbf{Z}_p \leftarrow [\mathbf{H}^0, ..., \mathbf{H}^L]$
$\mathbf{C} \leftarrow \text{ReLU}(\ GCN_{pd}(\mathbf{H}, \hat{\mathbf{A}}^t;\ \mathbf{W}_T)\ )$
$\mathbf{F} \leftarrow \text{ReLU}(\mathbf{V}_d \mathbf{W}_F^T)$
$\mathbf{Z}_d \leftarrow [\mathbf{Z}_p, \mathbf{C}, \mathbf{F}]$
$\mathbf{P} \leftarrow Sigmoid(\mathbf{Z}_d^T \mathbf{M} \mathbf{Z}_d)$
$\mathbf{P}_S = \{P_{d_n, d'_n, r_n} | (d_n, d'_n, r_n) \in S\}$

drug. In this work, we focus on the drugs that have at least a target protein, and the side effects that each occurred in at least 250 drug combinations. Therefore, there are 4,625,608 drug interaction edges (D-D) labeled by 861 side effects (R). We used as the positive polypharmacy adverse events to train and test the model.

Dataset split. The training and testing sets were stratified split based on side effects. To train and test the model, we need both positive and negative ADR knowledge instance in the form of (drug$_1$, drug$_2$, side effect). Positive instances are the observed polypharmacy side effect information, while negative ones are sampled with categorized negative sampling strategy [11].

Evaluation metrics. When comparing overall performance among trained models, we use the {Micro, Macro}-averaged receiver operator characteristic area under the curve (AUROC) and the area under the precision-recall curve (AUPRC) score for all side effects. The PIU and PPIU score of each side effect are computed based on the Macro-AUROC with the same $\mathbf{z}_d^{PIU}$ and $\mathbf{z}_d^{PPIU}$ for those of each prediction.

Lazy training. In order to weaken the contribution of protein and drug attributes to APRILE prediction, we control the trainability of the model parameters. Specifically, after initializing all the model parameters by Xavier initialization [61], we fix the parameters matrices $\mathbf{W}_P^1$ (in Eq.1) and $\mathbf{W}_F$ (in Eq.3), which are supposed to be trainable parameters, to their initial values. The fixed parameters are referred to as lazy learners.

APRILE-Pred's forward propagation is shown in Algorithm 1. We use the Adam optimizer with a training rate of 0.01 to minimize the cross-entropy loss. We applied full-batched end-to-end training for 80 epochs on a NVIDIA GV100GL graphics processing units (GPUs).

6.3 APRILE-Exp: Optimal Explanations Generation with GCNs Explainer

We develop APRILE-Exp to explore the molecular mechanisms of side effects. The APRILE-Exp explains the existence of known knowledge by understanding APRILE-Pred model prediction. It is primarily motivated as an adaption of previous work on GNNExplainer [7] for large-scale and tri-graph-like data. Given a trained APRILE-Pred model and a set of predictions we want to explain $S = \{(d_n, d'_n, r_n)\}_n$, the APRILE-Exp model generates a subgraph of the original graph (the dataset) that are most influential for the APRILE-Pred model's prediction. Therefore, the APRILE-Exp model transforms the problem of explaining why drug pairs cause side effects to finding out which protein interactions and drug interactions are the most important to predict side effects, which in turn helps us to understand the molecular mechanism of side effects.

We assigned an importance score to each edge in the P-P and P-D subgraphs by replacing the adjacency matrices $\mathbf{A}^p$ and $\mathbf{A}^t$ to $\hat{\mathbf{A}}^p$ and $\hat{\mathbf{A}}^t$, where $\hat{\mathbf{A}}^p$ and $\hat{\mathbf{A}}^t$ have the same non-zero elements as $\mathbf{A}^p$ and $\mathbf{A}^t$ but with tunable values $\hat{\mathbf{A}}^p_{ij} = \hat{\mathbf{A}}^p_{ji} \in [0, 1]$ and $\hat{\mathbf{A}}^t_{ij} \in [0, 1]$. As shown in Algorithm 2, the APRILE-Exp model calculates scores of drug pairs for each side effect with the same pipeline as the prediction model but with trainable $\mathbf{X}\hat{\mathbf{A}}^p$ and $\hat{\mathbf{A}}^t$, fixing the trained parameters $\mathbf{W}_P$, $\mathbf{W}_T$, and $\mathbf{W}_F$. We minimize the following loss during training:

$$\begin{aligned} Loss = &\sum_n \log(1 - P(d_n, d'_n, r_n)) \\ &+ \alpha(\sum_{i,j} \hat{A}^p_{i,j} + \sum_{i,j} \hat{A}^t_{i,j}) \\ &+ \beta(\sum_{i,j} \hat{A}^p_{i,j}(1 - \hat{A}^p_{i,j}) + \sum_{i,j} \hat{A}^t_{i,j}(1 - \hat{A}^t_{i,j})), \end{aligned} \tag{7}$$

where α and β are hyper-parameters. For the loss function, the first item encourages P to take a value close to 1, which can be replaced with any functions that decrease sharply near 1; the second term encourages the size of subgraph given by $\hat{A}_{i,j}$ to be as small as possible, and it can be replaced by any monotonically increase function; and the last term = $\hat{A}_{i,j}$ to be either 0 or 1, which can be replaced with any function that is 0 at points $x = 0$ and $x = 1$.

6.4 Gene Ontology Enrichment

To investigate the functional roles of the genes that are most influential for the interested predictions, we perform Gene Ontology (GO) enrichment analysis by using the GOATOOLS package [62]. The Fisher's exact test is used to compute the p-values for each of: biological process, molecular function and cellular component, which evaluates significant enrichment of certain GO terms. We choose the non-negative Benjamini-Hochberg multiple test correction method to correct the test, which controls the false discovery rate (FDR), and use a FDR set the significance cut-off to 0.05 [63]. We applied this methodology also to analyze enrichment of MeSH categories in clusters of side effects.

6.5 Permutation test

To assess the significance of APRILE-Exp explanations, we compared them against 10,000 random explanations (i.e., gene sets). For APRILE-Exp and random explanations, we computed the precision and Jaccard index using the CTD[24] disease-associated genes as validation. We compared two types of explanations: all genes in a APRILE-Exp subgraph, or the subset of these genes that are in enriched GO terms. Additionally, two types of random models were tested: (A) one that preserves the number of drug target vs. non-target genes, and (B) one that samples from all genes in the KG. Permutation test p-values were determined as the number of random explanations whose precision (or Jaccard index) was equal to or greater than that of the APRILE-Exp explanation.

6.6 Data Availability

Code and Data to train, test and select a APRILE-Pred model are available at `https://github.com/NYXFLOWER/APRILE-Pred`. Code to build and use APRILE-Exp is available at `https://github.com/NYXFLOWER/APRILE-Exp`.

References

1. Mair, A. *et al.* Polypharmacy management by 2030: a patient safety challenge. (2017).
2. Arrieta, A. B. *et al.* Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. *Information Fusion* **58,** 82–115 (2020).
3. Rajkomar, A., Dean, J. & Kohane, I. Machine learning in medicine. *New England Journal of Medicine* **380,** 1347–1358 (2019).
4. Jiménez-Luna, J., Grisoni, F. & Schneider, G. Drug discovery with explainable artificial intelligence. *Nature Machine Intelligence* **2,** 573–584 (2020).
5. Hastie, T., Tibshirani, R. & Wainwright, M. *Statistical learning with sparsity: the lasso and generalizations* (Chapman and Hall/CRC, 2019).
6. Roscher, R., Bohn, B., Duarte, M. F. & Garcke, J. Explainable machine learning for scientific insights and discoveries. *IEEE Access* **8,** 42200–42216 (2020).
7. Ying, Z., Bourgeois, D., You, J., Zitnik, M. & Leskovec, J. *Gnnexplainer: Generating explanations for graph neural networks* in *Advances in neural information processing systems* (2019), 9244–9255.
8. Hinton, G., Vinyals, O. & Dean, J. Distilling the knowledge in a neural network. *NIPS Deep Learning and Representation Learning Workshop* (2015).
9. Du, M., Liu, N. & Hu, X. Techniques for interpretable machine learning. *Communications of the ACM* **63,** 68–77 (2019).
10. Zitnik, M., Agrawal, M. & Leskovec, J. Modeling polypharmacy side effects with graph convolutional networks. *Bioinformatics* **34,** i457–i466 (2018).
11. Xu, H., Sang, S. & Lu, H. Tri-graph Information Propagation for Polypharmacy Side Effect Prediction. *NeurIPS Workshop on Graph Representation Learning* (2020).
12. Peterson, M. E. & Talcott, P. A. *Small Animal Toxicology-E-Book* (Elsevier Health Sciences, 2013).
13. Lipscomb, C. E. Medical subject headings (MeSH). *Bulletin of the Medical Library Association* **88,** 265 (2000).
14. McInnes, L., Healy, J., Saul, N. & Grossberger, L. UMAP: Uniform Manifold Approximation and Projection. *The Journal of Open Source Software* **3,** 861 (2018).
15. Piagnerelli, M. *et al.* Rapid alterations in transferrin sialylation during sepsis. *Shock* **24,** 48–52 (2005).
16. Kanashiro, A. *et al.* Therapeutic potential and limitations of cholinergic anti-inflammatory pathway in sepsis. *Pharmacological research* **117,** 1–8 (2017).
17. Karali, M. *et al.* Clinical and genetic analysis of a European cohort with pericentral retinitis pigmentosa. *International journal of molecular sciences* **21,** 86 (2020).
18. Chang, B. *et al.* A homologous genetic basis of the murine cpfl1 mutant and human achromatopsia linked to mutations in the PDE6C gene. *Proceedings of the National Academy of Sciences* **106,** 19581–19586 (2009).
19. Brazzelli, V. *et al.* Recurrence of mycosis fungoides on multiple melanocytic nevi: a case report and review of the literature. *Case reports in dermatology* **4,** 92–97 (2012).
20. Cerroni, L., Gatter, K. & Kerl, H. *An illustrated guide to skin lymphoma* (John Wiley & Sons, 2008).
21. Zhu, X. & Sly, W. Carbonic anhydrase IV from human lung. Purification, characterization, and comparison with membrane carbonic anhydrase from human kidney. *Journal of Biological Chemistry* **265,** 8795–8801 (1990).

22. Ekin, S., Berber, I., Kozat, S. & Gunduz, H. Selected trace elements and esterase activity of carbonic anhydrase levels in lambs with pneumonia. *Biological trace element research* **112,** 233–239 (2006).
23. McClintick, J. N. *et al.* Stress–response pathways are altered in the hippocampus of chronic alcoholics. *Alcohol* **47,** 505–515 (2013).
24. Davis, A. P. *et al.* Comparative Toxicogenomics Database (CTD): update 2021. *Nucleic acids research* **49,** D1138–D1143 (2021).
25. Frey, B. J. & Dueck, D. Clustering by passing messages between data points. *Science* **315,** 972–977 (2007).
26. Fleischer, L. M., Somaiya, R. D. & Miller, G. M. Review and meta-analyses of TAAR1 expression in the immune system and cancers. *Frontiers in pharmacology* **9,** 683 (2018).
27. Dempsey, J. A., Veasey, S. C., Morgan, B. J. & O'Donnell, C. P. Pathophysiology of sleep apnea. *Physiological reviews* (2010).
28. Alóe, F. Sleep bruxism neurobiology. *Sleep science* **2,** 40–48 (2009).
29. Mangge, H. *et al.* Disturbed tryptophan metabolism in cardiovascular disease. *Current medicinal chemistry* **21,** 1931–1937 (2014).
30. Quagliato, L. A. & Nardi, A. E. Cytokine alterations in panic disorder: A systematic review. *Journal of affective disorders* **228,** 91–96 (2018).
31. Arnolds, K. L. & Spencer, J. V. CXCR4: a virus's best friend? *Infection, Genetics and Evolution* **25,** 146–156 (2014).
32. Szklarczyk, D. *et al.* STITCH 5: augmenting protein–chemical interaction networks with tissue and affinity data. *Nucleic acids research* **44,** D380–D384 (2016).
33. Labarca, C. *et al.* Point mutant mice with hypersensitive $\alpha 4$ nicotinic receptors show dopaminergic deficits and increased anxiety. *Proceedings of the National Academy of Sciences* **98,** 2786–2791 (2001).
34. Wishart, D. S. *et al.* DrugBank: a comprehensive resource for in silico drug discovery and exploration. *Nucleic acids research* **34,** D668–D672 (2006).
35. Wilkinson, D. J. *et al.* Sympathetic activity in patients with panic disorder at rest, under laboratory mental stress, and during panic attacks. *Archives of General Psychiatry* **55,** 511–520 (1998).
36. Esler, M. *et al.* Cardiac sympathetic nerve biology and brain monoamine turnover in panic disorder. *Annals of the New York Academy of Sciences* **1018,** 505–514 (2004).
37. Smith, S. W., Hauben, M. & Aronson, J. K. Paradoxical and bidirectional drug effects. *Drug safety* **35,** 173–189 (2012).
38. Perucca, E. Overtreatment in epilepsy: adverse consequences and mechanisms. *Epilepsy research* **52,** 25–33 (2002).
39. Afsin, A., Asoğlu, R., Orum, M. H. & Cicekci, E. Evaluation of TP-E Interval and TP-E/QT Ratio in Panic Disorder. *Medicina* **56,** 215 (2020).
40. Esler, M. *et al.* in *Primer on the autonomic nervous system* 391–394 (Elsevier, 2004).
41. Liamis, G., Milionis, H. & Elisaf, M. Medication-induced hypophosphatemia: a review. *QJM: An International Journal of Medicine* **103,** 449–459 (2010).
42. Moe, S. M. Disorders involving calcium, phosphorus, and magnesium. *Primary Care: Clinics in Office Practice* **35,** 215–237 (2008).
43. Aubier, M. *et al.* Effect of hypophosphatemia on diaphragmatic contractility in patients with acute respiratory failure. *New England Journal of Medicine* **313,** 420–424 (1985).
44. Gravelyn, T. R., Brophy, N., Siegert, C. & Peters-Golden, M. Hypophosphatemia-associated respiratory muscle weakness in a general inpatient population. *The American journal of medicine* **84,** 870–876 (1988).
45. Amanzadeh, J. & Reilly, R. F. Hypophosphatemia: an evidence-based approach to its clinical consequences and management. *Nature clinical practice Nephrology* **2,** 136–148 (2006).
46. Cohen, J. *et al.* Hypophosphatemia following open heart surgery: incidence and consequences. *European journal of cardio-thoracic surgery* **26,** 306–310 (2004).
47. Salem, R. R. & Tray, K. Hepatic resection-related hypophosphatemia is of renal origin as manifested by isolated hyperphosphaturia. *Annals of surgery* **241,** 343 (2005).
48. Lebovitz, H. E. & Feinglos, M. N. Mechanism of action of the second-generation sulfonylurea glipizide. *The American journal of medicine* **75,** 46–54 (1983).
49. Barzilai, N., Groop, P.-H., Groop, L. & DeFronzo, R. A novel mechanism of glipizide sulfonylurea action: decreased metabolic clearance rate of insulin. *Acta diabetologica* **32,** 273–278 (1995).

50. Zimmermann, A. E. & Katona, B. G. Lansoprazole: a comprehensive review. *Pharmacotherapy: The Journal of Human Pharmacology and Drug Therapy* **17,** 308–326 (1997).
51. Riley, R., Parker, A., Trigg, S. & Manners, C. Development of a generalized, quantitative physicochemical model of CYP3A4 inhibition for use in early drug discovery. *Pharmaceutical research* **18,** 652–655 (2001).
52. Kjeldsen, S. E., Moan, A., Petrin, J., Weder, A. B. & Julius, S. Effects of increased arterial epinephrine on insulin, glucose and phosphate. *Blood pressure* **5,** 27–31 (1996).
53. Deen, N. S., Huang, S. J., Gong, L., Kwok, T. & Devenish, R. J. The impact of autophagic processes on the intracellular fate of Helicobacter pylori: more tricks from an enigmatic pathogen? *Autophagy* **9,** 639–652 (2013).
54. Worst, D., Otto, B. & De Graaff, J. Iron-repressible outer membrane proteins of Helicobacter pylori involved in heme uptake. *Infection and immunity* **63,** 4161–4165 (1995).
55. Qujeq, D., Sadogh, M. & Savadkohi, S. Association between helicobacter pylori infection and serum iron profile. *Caspian journal of internal medicine* **2,** 266 (2011).
56. Berstad, A., Berstad, K. & Berstad, A. pH-Activated Phospholipase A 2: an Important Mucosal Barrier Breaker in Peptic Ulcer Disease. *Scandinavian journal of gastroenterology* **37,** 738–742 (2002).
57. Marinka Zitnik Rok Sosic, S. M. & Leskovec, J. *BioSNAP Datasets: Stanford Biomedical Network Dataset Collection* `http://snap.stanford.edu/biodata`. Aug. 2018.
58. Kipf, T. N. & Welling, M. Semi-supervised classification with graph convolutional networks. *International Conference on Learning Representations* (2017).
59. Yang, B., Yih, W.-t., He, X., Gao, J. & Deng, L. Embedding entities and relations for learning and inference in knowledge bases. *arXiv preprint arXiv:1412.6575* (2014).
60. Tatonetti, N. P., Patrick, P. Y., Daneshjou, R. & Altman, R. B. Data-driven prediction of drug effects and interactions. *Science translational medicine* **4,** 125ra31–125ra31 (2012).
61. Glorot, X. & Bengio, Y. *Understanding the difficulty of training deep feedforward neural networks* in *Proceedings of the thirteenth international conference on artificial intelligence and statistics* (2010), 249–256.
62. Klopfenstein, D. *et al.* GOATOOLS: A Python library for Gene Ontology analyses. *Scientific reports* **8,** 1–17 (2018).
63. Noble, W. S. How does multiple testing correction work? *Nature biotechnology* **27,** 1135–1137 (2009).

Supplementary Information

A Supplementary Figures

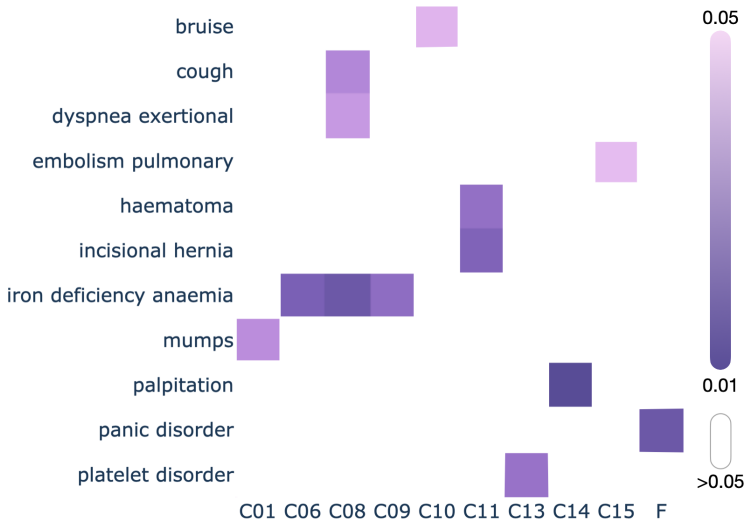

Figure A.1: Significant relations ($p < 0.05$) among side effect clusters and MeSH disease categories.

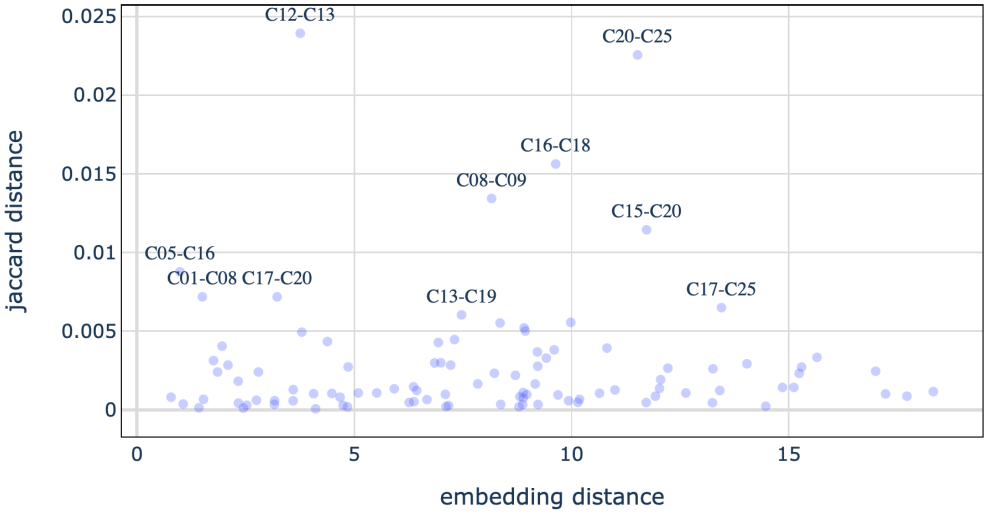

Figure A.2: Jaccard distance between side effects vs. Euclidean distance between embeddings learned by APRILE-Pred for each pair of categories (22 in total) in Fig. 2d. The markers are shown if their Jaccard distance > 0. The labels of markers are shown if their Jaccard distance > 0.006.

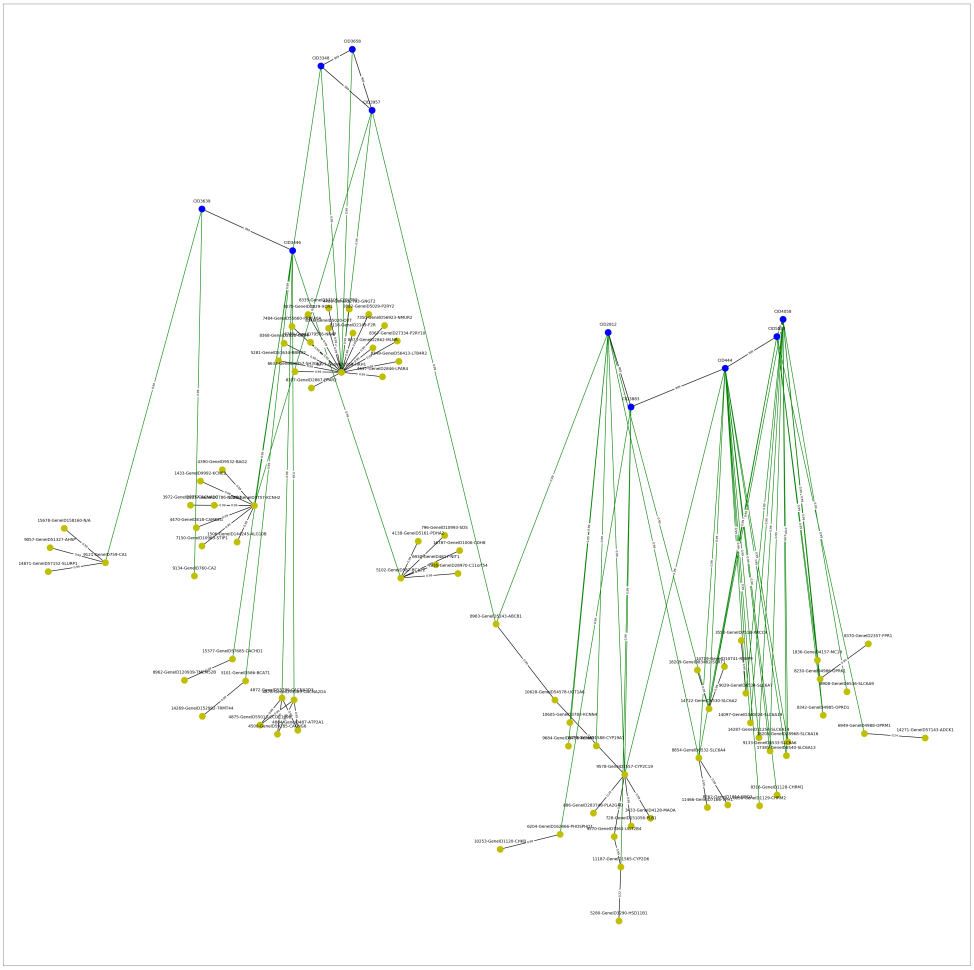

Figure A.3: APRILE explanation for panic disorder caused by drug pairs (10 possible drugs). Parameters used: regulation weight of 2.0, and probability threshold of 0.99.

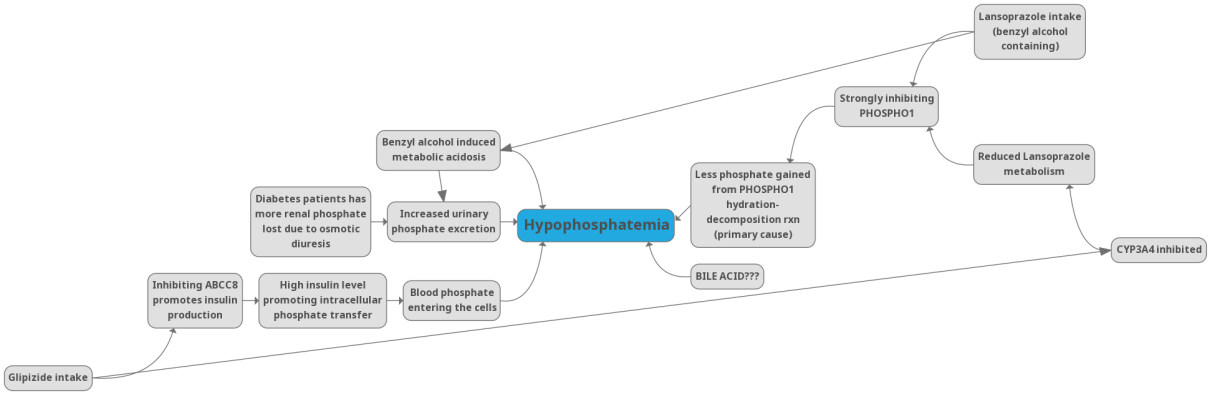

Figure A.4: Hypothesized mechanism for hypophosphatemia caused by glipizide-lansoprazole interaction.

641 # B Supplementary Tables

Table B.1: Macro-averaged AUROC for applying lazy training with different split rates and random states for parameter initialization

Split rate	Random-1	Random-2	Random-3	Random-4	Random-5
0.1	0.800364	0.800671	0.803084	0.797250	0.807832
0.2	0.844435	0.841982	0.846250	0.839958	0.850821
0.4	0.865966	0.861887	0.867489	0.864221	0.871715
0.6	0.870628	0.867482	0.869772	0.869908	0.875696
0.8	0.870904	0.867210	0.872412	0.870314	0.877834
0.9	0.872218	0.868020	0.873475	0.871330	0.878638

Table B.2: Commonly enriched GO terms identified by APRILE-Exp for three peptic ulcers

GO	GO description	Literature evidence
GO:0045540	Regulation of cholesterol biosynthetic process.	Cholesterol increases the resistance of *H. pylori*, and the clinical manifestations of this bacteria include peptic ulcer disease [1]. *Helicobacter pylori* is associated with dyslipidemia but not with other risk factors of cardiovascular disease.
GO:0055114	Oxidation-reduction process.	Derangements of redox homeostasis in the gastrointestinal tract are believed to lead to peptic ulcer [2].
GO:0019369	Arachidonic acid metabolic process.	Arachidonic acid metabolites play an important role in gastrointestinal homeostasis and have a significant impact in preventing or treating peptic ulcers [3].
GO:0033384	Geranyl diphosphate biosynthetic process.	Farnesyl diphosphate and geranyl diphosphate are both affected by the direct interaction between bisphosphonates and a biosynthetic enzyme [4]. Bisphosphonates are drugs used as treatment for osteoporosis that can cause serious inflammation and ulcers in the GI tract [5].
GO:0006805	Xenobiotic metabolic process.	Xenobiotics are widely used in the treatment of Parkinson disease, and Parkinson patients under treatment are prone to develop peptic ulcers, hence it is thought that there is a correlation [6].
GO:0005789	Endoplasmic reticulum membrane.	Research suggests that the endoplasmic reticulum would be the main source for autophagosomes' membrane, which is formed for autophagic processes during an ulcerative infection [7].
GO:0020037	Heme binding.	Some proteins that are involved in heme binding are thought to be one important virulence factor of *H. pylori* [8].
GO:0005506	Iron ion binding.	Iron ion binding capacity is thought to be lower in people with *H. pylori* infection [9].
GO:0102568	Phospholipase A2 activity consuming 1,2-dioleoylphosphatidylethanolamine.	This is directly correlated to the arachidonic acid metabolic process and to gastrointestinal ulceration [10].
GO:0004623	Phospholipase A2 activity.	Both this term and the previous one refer to the activity of phospholipase A2, which is a mucosal barrier breaker in peptic ulcer disease [11].
GO:0008395	Steroid hydroxylase activity.	Steroid hydroxylase is involved in steroid biosynthesis and a too high concentration of steroid can lead to gastrointestinal ulcers and bleeding [12].

C Supplementary Data

Supplementary data is available at

https://github.com/NYXFLOWER/APRILE-Exp/tree/master/supplementary-data

References

1. McGee, D. J. *et al.* Cholesterol enhances Helicobacter pylori resistance to antibiotics and LL-37. *Antimicrobial agents and chemotherapy* **55,** 2897–2904 (2011).
2. Cherkas, A. & Zarkovic, N. 4-Hydroxynonenal in redox homeostasis of gastrointestinal mucosa: Implications for the stomach in health and diseases. *Antioxidants* **7,** 118 (2018).
3. Isselbacher, K. The role of arachidonic acid metabolites in gastrointestinal homeostasis. *Drugs* **33,** 38–46 (1987).
4. Tsoumpra, M. K. *et al.* The inhibition of human farnesyl pyrophosphate synthase by nitrogen-containing bisphosphonates. Elucidating the role of active site threonine 201 and tyrosine 204 residues using enzyme mutants. *Bone* **81,** 478–486 (2015).
5. Cadarette, S. *et al.* Comparative gastrointestinal safety of weekly oral bisphosphonates. *Osteoporosis international* **20,** 1735–1747 (2009).
6. Carmody, R. N., Turnbaugh, P. J., *et al.* Host-microbial interactions in the metabolism of therapeutic and diet-derived xenobiotics. *The Journal of clinical investigation* **124,** 4173–4181 (2014).
7. Deen, N. S., Huang, S. J., Gong, L., Kwok, T. & Devenish, R. J. The impact of autophagic processes on the intracellular fate of Helicobacter pylori: more tricks from an enigmatic pathogen? *Autophagy* **9,** 639–652 (2013).
8. Worst, D., Otto, B. & De Graaff, J. Iron-repressible outer membrane proteins of Helicobacter pylori involved in heme uptake. *Infection and immunity* **63,** 4161–4165 (1995).
9. Qujeq, D., Sadogh, M. & Savadkohi, S. Association between helicobacter pylori infection and serum iron profile. *Caspian journal of internal medicine* **2,** 266 (2011).
10. Hornbeck, P. V. *et al.* PhosphoSitePlus, 2014: mutations, PTMs and recalibrations. *Nucleic acids research* **43,** D512–D520 (2015).
11. Berstad, A., Berstad, K. & Berstad, A. pH-Activated Phospholipase A 2: an Important Mucosal Barrier Breaker in Peptic Ulcer Disease. *Scandinavian journal of gastroenterology* **37,** 738–742 (2002).
12. Narum, S., Westergren, T. & Klemp, M. Corticosteroids and risk of gastrointestinal bleeding: a systematic review and meta-analysis. *BMJ open* **4** (2014).