Title: A network paradigm predicts drug synergistic effects using downstream protein-protein interactions

**Authors:** Jennifer L. Wilson1,\*, Ethan Steinberg2, Rebecca Racz3, Russ B. Altman4,5, Nigam Shah2, Kevin Grimes6

**Affiliations:**

1 Department of Bioengineering, University of California Los Angeles; Los Angeles, California, USA.

2 Center for Biomedical Informatics Research, Stanford University; Palo Alto California, USA.

3 Division of Applied Regulatory Science, US Food and Drug Administration; Silver Spring Maryland, USA.

4 Department of Bioengineering, Stanford University; Palo Alto California, USA.

5 Department of Genetics, Stanford University; Palo Alto California, USA.

6 Department of Chemical and Systems Biology, Stanford University; Palo Alto California, USA.

\*To whom correspondence should be addressed:

Jennifer Wilson

420 Westwood Plaza

5121 Engineering V, Rm 4121D

Los Angeles CA 90095

703-969-3318

[jenniferwilson@ucla.edu](mailto:jenniferwilson@ucla.edu)

**Key words:** systems biology, protein network models, drug-drug interactions, real world evidence, drug pathways

**Conflict of interest:** RBA is a founder and stockholder in Personalis, and a stockholder in 23andMe; he declares no conflicts of interest. JLW was a consultant for Sanofi from July-Dec 2021; she declares no conflicts of interest. All other authors declared no competing interests for this work.

**Abstract:** In some cases, drug combinations affect adverse outcome phenotypes by binding the same protein; however, drug-binding proteins are associated through protein-protein interaction networks within the cell, suggesting that drug phenotypes may result from long-range, network effects. We first used protein-protein interaction network analysis to classify drugs based on proteins downstream of their targets and next predicted drug combination effects where drugs shared network proteins but had distinct binding proteins (e.g., targets, enzymes, or transporters). By classifying drugs using their downstream proteins, we had an 80.7% sensitivity for predicting rare drug combination effects documented in gold-standard datasets. We further measured the effect of predicted drug combinations on adverse outcome phenotypes using novel observational studies in the electronic health record. We tested predictions for 60 network-drug classes on 7 adverse outcomes and measured changes in clinical outcomes for predicted combinations. These results demonstrate a novel paradigm for anticipating drug synergistic effects using proteins downstream of drug targets.

**Main Text:**

**INTRODUCTION**

Many drug-drug interactions (DDIs) associated with adverse effects occur from a shared binding protein, where drugs share similar targets, enzymes, carrier, or transporter proteins. For example, one drug can inhibit an enzyme that is responsible for metabolism of another drug substrate. However, not all DDIs are explained by this mechanism. Regulatory guidances recommend clinical (*1*) and in vitro experiments(*2*) to evaluate a drug’s potential for drug interactions; these recommendations emphasize the study of drug metabolizing enzymes and transporters that are relevant to other marketed therapies. DrugBank’s drug-drug interaction database (*3*, *4*) curates DDIs based on shared protein mechanisms and PharmGKB (*5*) curates drug-gene interactions, but neither consider the effects of proteins downstream of drug targets. In contrast, evidence suggests that drugs synergize through pathway effects without shared binding proteins. For example, the combined use of the chemotherapeutic drugs paclitaxel and carboplatin reduced hematopoietic toxicity experienced with carboplatin alone yet the combination did not affect the pharmacokinetics of either single drug (*6*), suggesting a non-shared-protein mechanism.

Using *in silico* methods, such as protein-protein interaction (PPI) network models, to anticipate drug effects is attractive because of the relative ease and scale of these methods for making predictions. These approaches have successfully predicted opportunities for drug repurposing (*7*-*10*), for treating co-morbid conditions (*11*), for identifying drug-drug interactions (DDIs) (*12*, *13*) and for understanding disease mechanisms (*14*). Already, there is mounting evidence that single and combination drug effects propagate through protein networks. Yet, downstream PPIs are not routinely used to anticipate drug effects in regulatory and industry settings because of the propensity of these models to over-predict drug phenotypes. We recently developed a per-phenotype PPI network approach that improved prediction performance 50% and increased average precision 76-95% when anticipating single drug adverse events, compared to global approaches (*15*). Interestingly, downstream proteins, relative to drug targets, were highly weighted in predicting a drug’s adverse outcome. Further, downstream proteins distinguished true from false positive predictions and were integral to preventing over-prediction. Because drug effects propagate through networks and our previous discovery that downstream proteins were predictive of drug adverse outcomes, we hypothesized that downstream proteins could be predictive of DDIs when drugs did not share binding proteins.

We explored the extent to which protein-protein interactions downstream from the targets of two drugs were sufficient to predict DDIs in cases where the drugs had distinct binding proteins (motivated in **Fig. 1.**) (We refer to drug targets, enzymes, carriers, or transporters as “targets” in the rest of this analysis.) To complete this analysis, we generated a novel set of drug-adverse-reaction (ADR) pairs by extracting these relationships from the drugs’ labels. Informed by the success of our per-phenotype PPI approach, we used meta-analysis to prioritize proteins downstream of targets of drugs labeled with the same ADR and re-classified drugs using these network proteins. We then predicted DDIs for drugs using their network class. We validated predicted combinations using novel observational studies in the electronic health record and demonstrated an ability to detect rare DDIs using protein interactions downstream of their targets. Although we used ADRs as a case study, our network paradigm is broadly applicable to all biological phenotypes, suggesting a relatively simple, and useful approach for anticipating drug synergistic effects generally.

**MATERIALS AND METHODS**

*Data and code availability*

The data and code used in this manuscript and referenced in this section are available at <https://github.com/jenwilson521/Designated-Medical-Event-Pathways>. Note: In the computational analysis, we used “DME” as shorthand for “ADR”. For transparency, we noted each script used for each analysis in the results section. Patient data from the electronic health record analysis is not made available to respect patient privacy and data use agreements.

*Extracting adverse reaction phenotypes from drug labels*

An algorithm was built using Linguamatics, a natural language processing software, to extract designated medical events (DMEs) (adverse drug reactions, ADRs) as MedDRA Preferred Terms from the black box warning, warnings & precautions, and adverse reactions sections of FDA product labels. All available FDA product labels (as of December 2017) were obtained from DailyMed and indexed in Linguamatics. For each ADR, the related MedDRA Preferred Term, Lower Level Term, and colloquial terms were searched (i.e., “SJS” was an additional term searched for “Stevens-Johnson syndrome”). Drugs with one or more ADRs in their product label were exported for analysis in PathFX. The data from this analysis are included in **Sup. File 1.** (*supp1\_Drugs\_labeled\_for\_AEs.txt)*.

*PathFX modeling of marketed drugs and identification of pathway associations to ARs*

To find pathway associations to ADRs, we used the PathFX algorithm(*8*) to identify network relationships between drug targets and ADR-associated proteins. Compared to other methods, PathFX used a data-driven approach to discover network associations, the algorithm generated “white-box” predictions of drug associations, and demonstrated high specificity in predicting drug-ADR effects(*8*). We used drug targets from DrugBank(*3*) (version 5.1.0) as inputs to PathFX. This analysis yielded a dataset of drug-ADR associations and downstream proteins associated with ADRs from drug labels. A summary of the PathFX algorithm approach and detailed description of the analysis and results is included in the *Supplemental Materials and Methods.*

*Network meta-analysis*

We next used meta-analysis to identify downstream proteins shared between drugs with the same ADR. PathFX contained multiple phenotypes associated with ADRs from a drug label (e.g., “Hemolytic anemia, nonspherocytic, due to glucose phosphate isomerase deficiency” and “Hemolytic anemia” were both considered as a prediction of “anemia”), and we collapsed these phenotypes when investigating each ADR. For a full description of the meta-analysis and pathways considered, please see the *Supplemental Materials and Methods.*

*Considering hypothesis for clinical evaluation*

We leveraged data in TWOSIDES(*16-18*) as a filter for predicted drug combinations. TWOSIDES used data from the FDA Adverse Event Reporting (FAERs) system for identifying adverse outcomes that were statistically associated with combinations of drugs. We searched TWOSIDES for our predicted combinations to assess whether drug combinations were observed clinically and to get an estimate of the potential effect size of a drug combination on an adverse outcome. We used the scripts */char\_data/charac\_novel\_combinations.py* and */Code/* *charac\_novel\_combos\_using\_int.py* to investigate if TWOSIDES supported our predicted drug combinations for ARPs and SPs. We leveraged drug synonyms from DrugBank (*contained in /data/drugbank\_vocabulary.csv*) to find match drug combinations from TWOSIDES with our predicted DDIs. We later filtered drug combinations that overlapped from our predictions and TWOSIDES if the predicted ARs were synonymous.

We next aggregated predicted DDIs by network class. We used drug-drug-network protein-ADR data from */data/cotherapy/potential\_co\_therapies.xlsx* and the filtered drug-drug combinations from TWOSIDES to generate predictions for our expanded observational studies. These predictions are contained in /*char\_data/network\_mechanisms\_for\_ehr\_ml.xlsx* and are summarized in **Table S6**.

*Novel observation studies using the Optum Clinformatics Dataset*

We pursued 60 novel observational studies using the Optum Clinformatics dataset using best practices for propensity matching patients to control confounding. Importantly, we used a deidentified dataset that did not require IRB approval. A full description of the dataset and methods are provided in the *Supplemental Materials and Methods.*

**RESULTS**

**Network analysis of single drugs with the same ADR**

We first discovered PPI associations between a drug’s target(s) and adverse drug reactions listed in the drug product’s FDA-approved drug labeling. Specifically, we focused our investigation on a list of designated medical events which are adverse drug reactions (ADRs) of high priority in regulatory review. We used a natural language processing method to extract ADRs from the warnings, boxed warnings, adverse reactions, and precautions sections of the drugs’ labels. This analysis yielded associations between 1,970 drugs and 34 ARs. This provided a unique dataset for interrogating network proteins of drugs associated with ADRs.

For network analysis, we restricted our analysis to 1,136 drugs that had drug-binding proteins listed in DrugBank(*3*) and further restricted to 970 drugs whose targets were connected in our PPI network(*8*). We used the PathFX algorithm (*8*) to create networks for these drugs (**Fig. 2.,** **Sup. File 2.**). Compared to other PPI network models, PathFX used the amount and quality of evidence supporting PPIs around drug targets to prioritize downstream proteins and then used statistical enrichment to discover phenotypes enriched in the drug’s network. Importantly, PathFX was naïve to a drug’s true set of phenotypes (e.g., an ADR from the drug label or the drug’s intend-to-treat disease) and instead used the corpus of evidence to anticipate drug network associations (further discussed in methods). PathFX discovered network associations for 424 drugs to 24 ARs.

This analysis discovered downstream proteins that were common to multiple drug-ADR pairs and distinct to ADRs (**Fig 3A.**, **Sup. File 3.**, **Sup. File 4.**). For example, for drugs labeled with sepsis, their networks shared drug-binding and downstream proteins (**Fig. 3B**). Because of these patterns, we reclassified drugs based on shared downstream proteins (**Fig. 2B**). For instance, multiple drugs associated with sepsis contained the adrenoreceptor beta 2 (ADRB2) downstream of their drug targets; this yielded two new classes for sepsis-associated drugs: “ADBR2 network” drugs and “non-ADBR2 network” drugs. We repeated this reclassification for all shared downstream proteins across all 24 ARs and tracked two types of network proteins for classification – ADR-associated network proteins (ARPs) or any protein on a shortest-path between a drug target and ADR protein (SPs). We discovered 172 network classes (each corresponding to 172 “non-Gene-net” classes) across 12 ADRs or 1,623 classes across 24 ADRs using ARPs, or SPs, respectively.

Next, using non-ADR drugs, we predicted novel DDIs for each network class where non-ADR drugs had target proteins downstream in the network class. For instance, ADBR2 is a target for the drug albuterol and albuterol is not associated with sepsis on its label. We predicted that “ADBR2 network” drugs would interact with albuterol to affect sepsis outcomes. In total, we predicted 18,988 drug-drug-ADR combinations using ARPs (51,605 combinations using SPs) from network classification. We further removed predicted DDIs if the drugs shared any target proteins because we were motivated to understand DDI effects due to downstream proteins. This yielded 6,098 drug-drug-ADR triplets using ARPs (19,741 triplets using SPs) representing 5,246 unique drug-drug pairs using ARPs (11,904 unique pairs using SPs) for further consideration (some drug-drug pairs were associated with multiple ADRs).

**Literature, TWOSIDES evidence supports combination effects and suggests directionality**

We estimated the sensitivity of our method by using TWOSIDES(*17*-*18*), a well-regarded dataset for drug combination effects. TWOSIDES uses the FDA Adverse Event Reporting System (FAERS) to detect DDIs based on the relative reporting rates of combination drugs as compared to single drugs while controlling for confounding variables(*17*-*18*). Predicted DDIs in TWOSIDES indicated combinations prescribed in the real world. For reference, TWOSIDES contained 42,920,391 drug-drug-ADR sets reported for 211,990 unique drug-drug pairs. Of note, TWOSIDES contained DDIs for 12,726 unique ADRs and included many more and milder side effects than our analysis (e.g., diarrhea, headache). We next counted our total drug-drug predictions and drug-drug-ADR triplets tracking both ARPs and SPs (**Table 1**). We first filtered our predictions by drug-drug combinations documented in TWOSIDES, reasoning that if a drug combination was reported in TWOSIDES, the combination was likely prescribed in the real world. To estimate the sensitivity of our method, we counted predicted drug-drug-ADR triplets documented in TWOSIDES (**Table 1**). From these results, using ARPs relative to SPs generated a higher sensitivity for detecting DDIs (80.7% vs 50.2%).

Our PathFX analysis identified non-directional associations and motivated us to pursue complementary data sources. For instance, our network analysis discovered an association between the drug paroxetine that had sepsis on its label, a non-sepsis-labeled drug, albuterol, and the ADR, sepsis, but our analysis didn’t indicate whether co-administering albuterol with paroxetine would reduce or worsen sepsis. We used the literature to infer the directionality of drug effects. Specifically, we used sentences to identify if a combination drug would worsen or mitigate the drug-induced ADR. We searched for combo-drug-ADR relationships within PubMed abstracts using natural language processing. Using emerging results from the literature was important because it did not replicate data used in the network analysis. We manually curated sentences containing mention of combo drugs and ADRs to understand how the combo drug may affect the ADR. For instance, the beta-2 agonist albuterol, binds the ADRB2 protein, which is downstream in the interaction network of three drugs that are associated with sepsis on their drug labels (paroxetine, atropine, and cocaine). Albuterol is not associated with sepsis on its drug label, yet in our search of published abstracts, we discovered that albuterol is associated with sepsis in a rat model. Specifically, we discovered and manually validated the following sentence to support further consideration of these combinations study: *“This study showed for the first time that oral administration of albuterol exerted protective effects on CLP-induced sepsis and related lung injury in rats”* (*19*). The full list of predicted drug-drug-ADR combinations, their relevant network proteins, and manually curated literature evidence are provided in **Sup. File 7.**

We further investigated predicted drug-drug-ADR combinations with clinical data using published databases and novel observational studies. To consider feasibility of detecting drug combination effects in patient data, we again referenced TWOSIDES (*17-18*), and investigated the proportional reporting ratios (PRRs) documented for the 368 combinations discovered using ARPs. We used PRRs documented in the TWOSIDES database; importantly, TWOSIDES corrected for hidden and unmeasured covariates when calculating predicted drug combination effects. PRRs were a sufficient proxy for ADR severity of predicted DDIs. Indeed, predicted combination effects were discovered in TWOSIDES (**Table 2**, full results in **Sup. File 6**). Further, because PathFX networks do not contain directional pathway information, measuring a drug-combination effect in TWOSIDES suggested that drug combinations may increase risk for ADRs in the real world.

We also pursued multiple novel observational studies to test our hypotheses and used two approaches to conduct this analysis. For these analyses, we used the deidentified Optum Clinformatics dataset v7 that included over 88 million US patients, both privately insured and Medicare beneficiaries, largely under the age of 65. We accessed a version of the Optum dataset standardized to the Observational Medical Outcomes Partnership (OMOP) common data model (CDM); standardized data models have decreased heterogeneity between datasets and improved consistency in underlying data(*20*). We leveraged the Observational Health Data Sciences and Informatics (OHDSI) network tools, specifically CohortMethod to measure ADR outcomes for patients exposed to our predicted combinations. CohortMethod is a software package that facilitates extracting patient data from the EHR, conducting large-scale propensity matching for controlling confounding variables, and estimating outcome models, such as Cox regression (example applications *21,22*). The LEGEND study is one relevant application where CohortMethod measured cardiovascular outcomes across first-line anti-hypertensive drugs(*21*). Our approach was conceptually similar, however, we aggregated drugs into network classes instead of by their chemical structures or therapeutic use classes. We first used CohortMethod to test two network DDI predictions to validate our ability to detect DDIs in the real world. We used a customized pipeline to analyze the 58 network classes that encompassed the 368 drug-drug-ADR triplets documented in TWOSIDES.

**Altered sepsis outcomes for network-classified drugs in a novel observational study**

We first investigated the effect of albuterol on beta-2 adrenergic receptor 2 (ADRB2)network drugs. PathFX identified network associations for 29 drugs with sepsis listed on the drugs’ labels. From this 29-drug set, two drugs, paroxetine, and atropine, contained ADRB2, an albuterol (also known as salbutamol) drug target, downstream in their networks and did not share other target proteins with albuterol. Of the remaining 27 drugs, 18 drugs did not contain albuterol-binding proteins downstream in their networks nor share drug target proteins with albuterol. These 18 drugs were considered the “non-ADRB2-net” class (**Table S1**). We hypothesized that concomitant use of albuterol would alter risk of sepsis for ADBR2-network drugs relative to non-ADBR2-network drugs.

For the first measurement, we measured the risk of sepsis for patients on ADBR2-network (“target” cohort) or non-ADBR2-network (“comparator” cohort) drugs. For the second measurement, we measured the risk of sepsis for patients with an overlapping exposure to albuterol+ADBR2-network drugs (“target” cohort) or albuterol+ non-ADBR2-network (“comparator” cohort) drugs. To select patients with an overlapping exposure, we required patients have an albuterol “DRUG ERA” that started between the start and end of an exposure to either the ADBR2-network or non-ADBR2-network drugs and risk for sepsis was observed for 30 days following the start of the second drug exposure. The drug era is considered a sufficient proxy to estimate an exposure to an active ingredient and the details of this data table are further explained in *Materials and Methods.* We further used large-scale propensity matching to estimate confounding and then matched patients based on their propensity score to estimate risk. Propensity matching aggregates all available patient data in the health record including commonly-considered confounders such age, diagnoses, demographics as well as other data such as number of visits, time to visits, that also reflect patient characteristics

(*17-18*, *23*, *24*). After matching, we discovered good covariate balance between the target and comparator cohorts and sufficient patient attrition for measuring outcomes (**Fig. S1., Sup. File 6.**).

We measured the risk of sepsis between these two drug classes without a combination therapy and with co-administration of albuterol (**Fig. 4**., **Table 3**). The risk of sepsis occurring in the ADBR2-net class is increased compared to the non-ADBR2-net class when albuterol is used concurrently: Hazard Ratio (HR)=0.792 with the combination compared to HR=0.525 without the combination; this yielded an HR-ratio of 1.51. The risk of sepsis from paroxetine or atropine (ADBR2-network class) was less than non-ADBR2-network class drugs, however, the combined use of albuterol with paroxetine or atropine increased the risk of sepsis compared to non-ADBR2-network class. We did not discover literature evidence supporting sepsis outcomes in combined use of atropine or paroxetine with albuterol. A retrospective chart review supported that albuterol and atropine were both therapeutic options for systematic bradycardia (*25*), suggesting that patients may have overlapping exposures to these drugs. A clinical trial in infants suffering from chronic lung disease observed that salbutamol (a synonym of albuterol) had no observable effect on patient sepsis (*26*), further supporting that albuterol isn’t associated with sepsis when used alone.

**Altered pancreatitis outcomes for network-classified drugs in a novel observational study**

We additionally repeated this process, this time emphasizing the effect of aspirin (also known as acetylsalicylic acid) prescribed in combination with drug network classes associated with aspirin target proteins on the ADR, pancreatitis. We measured hazard ratios for patients in these groups (**Fig. 4**., **Table 3**) and observed a shift in hazard ratio for patients taking the predicted combination drug (full explanation included in *Supplemental Results).*

**Novel observational studies for 58 additional DDI classes discovered using ARPs**

We sought validation for an additional 58 network-class DDIs predicted from using ARPs because these predictions had higher sensitivity for anticipating effects in TWOSIDES (**Sup. File 6**). For these 58 predictions, we estimated a hazard ratio using Cox regression on a 1-1 propensity score matched cohort with a caliper of 0.1. We used a logistic regression propensity score model trained on low dimensional CLMBR patient representations

(*29*). Precomputed CLMBR representations enabled more rapid cohort definitions and HR-ratio estimation than CohortMethod. Like before, we included patients in our baseline/combo analysis if they had exposure to drugs in the network or non-network classes with/without the predicted combo drug, respectively. We measured HRs between network and non-network drug classes with and without drug combinations.

Like the two cases outlined above, we measured the relative risk between the net-GENE and non-net-GENE classes with and without the predicted combination drug for the 58 remaining class predictions (**Sup. File 5**). Not surprisingly because our drug interaction predictions are rare, we were unable to generate sufficient patient cohorts to measure HRs for all predicted classes. For 21 of the 58 total classes, we had sufficient patients to measure HRs between the net-GENE and non-net-GENE classes and for 8 of these cases, there were also sufficient patients to measure HRs between drug classes with the predicted combination drugs (**Fig. S3**). However, we removed one class prediction because the predicted combination drug’s indication was too similar to the side effect. For the 7 remaining classes we measured the change in Cox coefficient to estimate effects of the predicted DDIs. (**Table 4, Fig. S4**). These DDI effects were moderate with HR-ratios that ranged from 0.85-1.17. The highest HR-ratio, 1.17, was measured for drug-induced hypertension associated with proopiomelanocortin (POMC) network class drugs used in combination with loperamide, an ingredient used to treat diarrhea. The lowest HR-ratio, 0.85, was measured for from drug-induced hypertension associated with prostaglandin E receptor 4 (PTGER4) network class drugs used in combination with misoprostol, an ingredient once used to treat stomach ulcers. This result suggests a protective effect of misoprostol for the PTGER4 network drug class.

**DISCUSSION**

We predicted drug-drug interactions (DDIs) using network classification and validated our predictions using DDI databases and novel observational studies. We first extracted a novel dataset of drug-adverse reaction (ADR) pairs using data extracted from drug labels. This dataset was crucial to our analysis and will be valuable to other investigations of drug ADRs. We used network analysis to discover downstream proteins associated with ADRs, reclassified drugs by their downstream proteins, and predicted DDIs based on network classification. We demonstrated high sensitivity for detecting rare DDIs using ARPs for classification, further supporting that rare or emerging drug-drug effects may arise when drugs do not share protein targets. We validated DDI predictions for albuterol and aspirin based on network classification and 58 additional DDI effects using novel observational studies. Overall, these results provide evidence for investigating downstream proteins for anticipating DDIs and that protein-protein interactions between drugs’ targets are sufficient for identifying drug combination effects.

Compared to other network approaches, our analysis was, to our knowledge, unique in the requirement that we excluded drugs with shared protein targets. This allowed us to exclusively explore DDIs that resulted from downstream effects and not a shared protein mechanism. Other approaches to predicting DDIs are extensively reviewed in (*30*)and our approach is most like the network propagation technique of Park et al (*31*). PathFX, like their approach, begins with a diffusion-based approach to identify potential signaling cascades affected by binding a drug’ target(s). Instead of requiring two drugs to have similar signaling networks, we only required that an AR-associated drug “diffuse” to the target of a non-AR drug target. We prioritized DDIs where multiple AR drugs converged on similar non-AR drug targets (e.g., ADRB2).

Fortunately, rigorous regulatory review and good clinical practices prevent the use of many harmful drug combinations, and this limited our ability to extensively validate every prediction. We could not measure DDI effect sizes for all network predictions, yet we found evidence for rare drug combinations. Nonetheless, *in silico* network analysis is relatively cheap and efficient and could aid in therapeutic development where anticipation of ADRs is essential for therapeutic development. Further, our predictions are not documented in routinely used DDI data sources and integration of these predictions could inform clinical care or further research efforts. Our discovery of drug combinations that mitigated ADR outcomes suggested a new paradigm for managing drug induced ADRs; specifically, that mitigating therapies could be prescribed based on drug network class. Similarly, *in silico* network analysis which predicts therapies to mitigate side effects could also inform safety analysis plans for clinical development.

Our results also have implications for advancing protein-protein interaction (PPI) networks for anticipating drug effects. There is sufficient evidence that PPI networks can anticipate drug effects and be used predictively for identifying repurposing opportunities. However, our analysis is distinct because of our emphasis on attribution; we aimed to ascribe drug effects to specific downstream proteins. Classifying drugs by their downstream proteins and measuring relative ADR risk in the presence of secondary drugs is evidence that drug effects could be attributed to downstream proteins discovered from PathFX network analysis. Further experimental validation would be required to investigate these hypotheses. However, it suggests that PPI methods are useful not just for pattern discovery (e.g., drug A’s network is like drug B’s network) but also for predicting mechanistic effects (e.g., drug A’s ADR outcome is mediated by the downstream protein Y).

Our study expands a growing body of knowledge that drugs can exert synergistic effects without sharing drug-binding proteins, which may lead to a better understanding of adverse drug reactions and rational design of new therapeutic combinations. Drug synergy is a broad field where many frameworks are used to anticipate drug effects (*32*, *33*). Some approaches leverage “supra-additive” effects of drugs used in combination (*34*, *35*), yet these measurements often rely on complex and relatively costly high-throughput screens(*36*). While the performance of computational synergy prediction algorithms has increased, these effects have yielded little success in the clinic (*33*, *37*). A community competition for synergy prediction noted that drugs with high experimental synergy contained drug targets in the same pathway and further, that well-predicted drug synergies occurred when combination drug targets were downstream of a shared protein (*33*). While we used drug induced ADRs as the focus of this investigation, analysis of proteins downstream of drug targets could improve prediction of drug synergistic effects on disease outcomes.

**Study Highlights:** Current knowledge of drug-drug-interactions (DDIs) emphasize the drug target level by identifying shared transporters, enzymes, or pharmacodynamic targets, and don’t prioritize proteins downstream of targets. Here we sought to address if proteins downstream of drug targets were sufficient to predict DDIs; we used protein interaction network analysis and real-world evidence to predict and detected rare DDIs mediated by downstream proteins. These results provide evidence that downstream proteins are sufficient for anticipating drug-drug effects. From these results, anticipating drug side effects during drug discovery and clinical development may consider drug interactions mediated through downstream proteins.

**Acknowledgments:** The authors would like to thank Graham Erwin and Emily Flynn for reading the manuscript, and Oluseyi Adeniyi, Jielin Sun, and Michael Pacanowski for reading the manuscript and helpful discussions about the material.

**Funding:** JLW and KG were supported SPARK, JLW was supported by a Sanofi iDEA Award and Grant Number U01FD004979 from the FDA.

**Author contributions:**

J.L.W. wrote the manuscript; J.L.W. designed the research; J.L.W., E.S., R.R. performed the research; J.L.W., E.S., R.R. analyzed the data; E.S., N.S., R.B.A., K.G. contributed new reagents/analytical tools.

**Data and materials availability:** All data are available in the main text or the supplementary materials.

References

1. *In Vitro Drug Interaction Studies — Cytochrome P450 Enzyme- and Transporter-Mediated Drug Interactions Guidance for Industry* (U.S. Food and Drug Administration, Guidance for Industry, 2020).

2. *Clinical Drug Interaction Studies — Cytochrome P450 Enzyme- and Transporter-Mediated Drug Interactions Guidance for Industry* (U.S. Food and Drug Administration, Guidance for Industry, 2020).

3. Wishart, D. S. *et al.* DrugBank: a comprehensive resource for in silico drug discovery and exploration. *Nucleic Acids Research* **34**, D668-72 (2006). doi: 10.1093/nar/gkj067.

4. Law, V. *et al.* DrugBank 4.0: shedding new light on drug metabolism. *Nucleic Acids Research* **42**, D1091-7 (2014). doi: 10.1093/nar/gkt1068.

5. Whirl-Carrillo, M. *et al.* Pharmacogenomics Knowledge for Personalized Medicine. *Clinical pharmacology and therapeutics* **92**, 414–417 (2012). doi: 10.1038/clpt.2012.96.

6. McLeod, H. L. Clinically relevant drug-drug interactions in oncology. *Br J Clin Pharmacol* **45**, 539–544 (1998). doi: 10.1046/j.1365-2125.1998.00719.x.

7. Guney, E., Menche, J., Vidal, M. & Barabási, A.-L. Network-based in silico drug efficacy screening. *Nat Commun* **7**, 10331 (2016). doi: 10.1038/ncomms10331.

8. Wilson, J. L. *et al.* PathFX provides mechanistic insights into drug efficacy and safety for regulatory review and therapeutic development. *PLoS Comput Biol* **14**, e1006614-27 (2018). doi: 10.1371/journal.pcbi.1006614.

9. Cheng, F., Kovács, I. A. & Barabási, A.-L. Network-based prediction of drug combinations. *Nat Commun* **10**, 1–11 (2019). doi: 10.1038/s41467-019-09186-x.

10. Gysi, D. M. *et al.* Network medicine framework for identifying drug-repurposing opportunities for COVID-19. *Proc. Natl. Acad. Sci. U.S.A.* **118**, (2021). doi: 10.1073/pnas.2025581118.

11. Aguirre-Plans, J. *et al.* Proximal Pathway Enrichment Analysis for Targeting Comorbid Diseases via Network Endopharmacology. *Pharmaceuticals* **11**, 61–18 (2018). doi: 10.3390/ph11030061.

12. Zitnik, M., Agrawal, M. & Leskovec, J. Modeling polypharmacy side effects with graph convolutional networks. *Bioinformatics* **34**, i457–i466 (2018). doi: 10.1093/bioinformatics/bty294.

13. Yoo, S. *et al.* In silico profiling of systemic effects of drugs to predict unexpected interactions. *Sci Rep*  **8**, 1612 (2018). doi: 10.1038/s41598-018-19614-5.

14. Ruiz, C., Zitnik, M. & Leskovec, J. Identification of disease treatment mechanisms through the multiscale interactome. *Nat Commun* **12**, 1–15 (2021). doi: 10.1038/s41467-021-21770-8.

15. Wilson, J. L., Gravina, A. & Grimes, K. From random to predictive: a context-specific interaction framework improves selection of drug protein–protein interactions for unknown drug pathways. *Integr Biol* (2022). doi: 10.1093/intbio/zyac002.

16. Tatonetti, N. P. *et al.* Detecting Drug Interactions From Adverse-Event Reports: Interaction Between Paroxetine and Pravastatin Increases Blood Glucose Levels. *Clinical pharmacology and therapeutics* **90**, 133–142 (2011). doi: 10.1038/clpt.2011.83.

17. Tatonetti, N. P., Ye, P. P., Daneshjou, R. & Altman, R. B. Data-driven prediction of drug effects and interactions. *Sci Transl Med* **4**, 125ra31-125ra31 (2012). doi: 10.1126/scitranslmed.3003377.

18. Tatonetti, N. P., Fernald, G. H. & Altman, R. B. A novel signal detection algorithm for identifying hidden drug-drug interactions in adverse event reports. *J Am Med Inform Assoc* **19**, 79–85 (2012). doi: 10.1136/amiajnl-2011-000214.

19. Ozogul, B. *et al.* Comparative study on effects of nebulized and oral salbutamol on a cecal ligation and puncture-induced sepsis model in rats. *Drug Res (Stuttg)* **65**, 192–198 (2015). doi: 10.1055/s-0034-1375683.

20. Voss, E. A., Ma, Q. & Ryan, P. B. The impact of standardizing the definition of visits on the consistency of multi-database observational health research. *BMC Med Res Methodol* **15**, 13–10 (2015). doi: 10.1186/s12874-015-0001-6.

21. Suchard, M. A. *et al.* Comprehensive comparative effectiveness and safety of first-line antihypertensive drug classes: a systematic, multinational, large-scale analysis. *The Lancet* **394**, 1816–1826 (2019). doi: 10.1016/S0140-6736(19)32317-7.

22. Hripcsak, G. *et al.* Comparison of Cardiovascular and Safety Outcomes of Chlorthalidone vs Hydrochlorothiazide to Treat Hypertension. *JAMA Intern Med* **180**, 542–551 (2020). doi: 10.1001/jamainternmed.2019.7454.

23. Sekhon, J. S. Multivariate and Propensity Score Matching Software with Automated Balance Optimization: The Matching Package for R. *Jour of Statistical Soft.* **42**, 1-52 (2011). doi: 10.18637/jss.v042.i07.

24. Tian, W. *et al.* Leukotriene B4antagonism ameliorates experimental lymphedema. *Sci Transl Med* **9**, eaal3920 (2017). doi: 10.1126/scitranslmed.aal3920.

25. Rollstin, A., Carey, M. C., Doherty, G., Tawil, I. & Marinaro, J. Oral albuterol to treat symptomatic bradycardia in acute spinal cord injury. *Intern Emerg Med* **11**, 101–105 (2016). doi: 10.1007/s11739-015-1324-3.

26. Ng, G., Silva, O. da & Ohlsson, A. Bronchodilators for the prevention and treatment of chronic lung disease in preterm infants. *Cochrane Database Syst Rev* **12**, CD003214 (2016). doi: 10.1002/14651858.CD003214.pub3.

27. Hung, W. Y. & Lanfranco, O. A. Contemporary review of drug-induced pancreatitis: A different perspective. *World J Gastrointest Pathophysiol* **5**, 405–415 (2014). doi: 10.4291/wjgp.v5.i4.405.

28. Audia, P., Feinfeld, D. A., Dubrow, A. & Winchester, J. F. Metformin-induced lactic acidosis and acute pancreatitis precipitated by diuretic, celecoxib, and candesartan-associated acute kidney dysfunction. *Clin Toxicol (Phila)* **46**, 164–166 (2008). doi: 10.1080/15563650701355314.

29. Steinberg, E. *et al.* Language models are an effective representation learning technique for electronic health record data. *J Biomed Inform* **113**, 103637 (2021). doi: 10.1016/j.jbi.2020.103637.

30. Han, K. *et al.* A Review of Approaches for Predicting Drug–Drug Interactions Based on Machine Learning. *Front Pharmacol* 12, 814858 (2022). doi: 10.3389/fphar.2021.814858.

31. Park, K., Kim, D., Ha, S. & Lee, D. Predicting Pharmacodynamic Drug-Drug Interactions through Signaling Propagation Interference on Protein-Protein Interaction Networks. Plos One 10, e0140816 (2015). doi: 10.1371/journal.pone.0140816.

32. Bansal, M. *et al.* A community computational challenge to predict the activity of pairs of compounds. *Nature Biotechnology* **32**, 1213–1222 (2014). doi: 10.1038/nbt.3052.

33. Menden, M. P. *et al.* Community assessment to advance computational prediction of cancer drug combinations in a pharmacogenomic screen. *Nat Commun* **10**, 2674 (2019). doi: 10.1038/s41467-019-09799-2.

34. Geary, N. Understanding synergy. *Am J Physiol Endocrinol Metab* **304**, E237-53 (2013). doi: 10.1152/ajpendo.00308.2012.

35. Wooten, D. J., Meyer, C. T., Lubbock, A. L. R., Quaranta, V. & Lopez, C. F. MuSyC is a consensus framework that unifies multi-drug synergy metrics for combinatorial drug discovery. *Nat Commun* **12**, 4607–16 (2021). doi: 10.1038/s41467-021-24789-z.

36. Han, K. *et al.* Synergistic drug combinations for cancer identified in a CRISPR screen for pairwise genetic interactions. *Nature Biotechnology* **35**, 463–474 (2017). doi: 10.1038/nbt.3834.

37. Palmer, A. C. & Sorger, P. K. Combination Cancer Therapy Can Confer Benefit via Patient-to-Patient Variability without Drug Additivity or Synergy. *Cell* **171**, 1678-1691.e13 (2017). doi: 10.1016/j.cell.2017.11.009.

38. Villaveces, J. M. *et al.* Merging and scoring molecular interactions utilising existing community standards: tools, use-cases and a case study. *Database (Oxford)* **2015**, bau131–bau131 (2015). doi: 10.1093/database/bau131.

39. Ryan, P. B., Schuemie, M. J., Gruber, S., Zorych, I. & Madigan, D. Empirical performance of a new user cohort method: lessons for developing a risk identification and analysis system. *Drug Saf* **36 Suppl 1**, S59-72 (2013). doi: 10.1007/s40264-013-0099-6.

**Figure Captions**

**Fig. 1**. **A downstream-protein paradigm for detecting drug-drug interactions is distinct from a shared-protein paradigm**. In a shared protein mechanism, two drugs share a protein target (transporter, metabolizing enzyme, etc.), and this competition causes an adverse outcome. In a network protein mechanism, protein-protein interactions connect drugs’ target proteins. A DDI is caused when a second drug targets a protein downstream of the original drug. Downstream proteins can be used to reclassify drugs.

**Fig. 2.** **Project workflow used pathway modeling and electronic health record analysis to assess DDIs predicted by network class.** (**A**) Starting with drug-ADR relationships from drug labels, we used PPI modeling and network classification to predict DDIs (workflow demonstrated for ARPs). We filtered predicted DDIs using TWOSIDES for further validation. (**B**) The starting drug-ADR dataset comprised 1970 drugs associated with 34 different adverse drug reactions (ADRs). For the 970 drugs with targets connected to our interactome, we constructed networks and looked for downstream associations to ADRs. We used downstream AR-associated proteins (ARPs) to define network classes and predicted DDIs where non-ADR drugs targeted downstream proteins. The figure depicts a hypothetical “GENE A” class based on the downstream protein, “A”. (**C**) We validated predicted DDIs by measuring hazard ratios between “network” and “non-network” classes with (top row) and without (bottom row) the predicted combination drug and take the ratio of HRs to estimate the DDI effect. Hypothetical example shown to depict experimental set-up.

**Fig. 3.** **Drug networks have common target and downstream proteins across and within ADRs.** (**A**) The most common target (red) and downstream (grey) proteins (rows) for all drug networks associated with 24 ADRs (columns). Protein count indicates in how many drug networks the protein appears (**B**) Target (red) and downstream (grey) proteins (columns) for all sepsis-associated drugs networks (rows).

**Fig. 4.** **Hazard ratio estimates for ADBR2 and T-E-N network classes.** We estimated the between class effects with and without a combination drug for two predicted DDI effects.

**Supplementary Materials**

Materials and Methods

*Summary of PathFX analysis*

*Application of PathFX to understand drug-ADR relationships mediated by downstream proteins*

*Extended explanation of network meta-analysis*

*Network images and heatmaps*

*Identifying novel co-therapies and determining directionality of effect*

*Electronic Health Record Dataset: Optum Clinformatics Data Mart 7.0*

*Novel observational study for assessing aspirin and albuterol combinations using CohortMethod*

*Novel observational studies for additional network predictions*

Fig. S1. Attrition diagram after performing patient matching in sepsis study.

Fig. S2. Attrition diagram after performing patient matching in pancreatitis study.

Fig. S3. Cox coefficients measured for 21 of the 58 predicted class effects**.**

Fig. S4.Changes in the cox coefficient for 8 predicted classes with combination drugs compared to without the combination drug.

Table S1. Drugs used in sepsis DDI study.

Table S2. Drugs used in pancreatitis DDI study.

Data file S1. ADRs extracted from drug labels (supp1\_Drugs\_labeled\_for\_AEs.txt)

Data file S2. Drug-network ADR associations discovered from PathFX (supp2\_true\_positives\_summary.xlsx)

Data file S3. Analysis of common downstream, ADR-associated proteins (supp3\_DME\_merged\_node\_counts.xlsx)

Data file S4. Top ADR-associated network proteins used in heatmap (supp4\_DME\_heatmap\_top\_node\_counts.xlsx)

Data file S5. DDI predictions from all-shortest-path (SPs) (supp5\_all\_SP\_drug\_class\_predictions.xlsx)

Data file S6. DDI predictions from ADR-associated-proteins (ARPs) and EHR analysis results (supp6\_ARPs\_PRRs\_clin\_results.xlsx)

Data file S7. Summary from literature search of directionality of DDI effects. (supp7\_summary\_drug\_interactions.xlsx)