CLINICAL DATA ANALYSIS

# Data-Driven Prediction of Drug Effects and Interactions

Nicholas P. Tatonetti,[1,2] Patrick P. Ye,[2] Roxana Daneshjou,[3] Russ B. Altman[2,3]*

Adverse drug events remain a leading cause of morbidity and mortality around the world. Many adverse events are not detected during clinical trials before a drug receives approval for use in the clinic. Fortunately, as part of postmarketing surveillance, regulatory agencies and other institutions maintain large collections of adverse event reports, and these databases present an opportunity to study drug effects from patient population data. However, confounding factors such as concomitant medications, patient demographics, patient medical histories, and reasons for prescribing a drug often are uncharacterized in spontaneous reporting systems, and these omissions can limit the use of quantitative signal detection methods used in the analysis of such data. Here, we present an adaptive data-driven approach for correcting these factors in cases for which the covariates are unknown or unmeasured and combine this approach with existing methods to improve analyses of drug effects using three test data sets. We also present a comprehensive database of drug effects (Offsides) and a database of drug-drug interaction side effects (Twosides). To demonstrate the biological use of these new resources, we used them to identify drug targets, predict drug indications, and discover drug class interactions. We then corroborated 47 (P < 0.0001) of the drug class interactions using an independent analysis of electronic medical records. Our analysis suggests that combined treatment with selective serotonin reuptake inhibitors and thiazides is associated with significantly increased incidence of prolonged QT intervals. We conclude that confounding effects from covariates in observational clinical data can be controlled in data analyses and thus improve the detection and prediction of adverse drug effects and interactions.

## INTRODUCTION

Adverse drug events (ADEs) remain a significant source of mortality and morbidity around the world with costs estimated at several billion dollars each year (1, 2). Many ADEs are rare or occur only in a subset of the human population and not observed in relatively small clinical trials. To address this issue, the U.S. Food and Drug Administration (FDA), World Health Organization, and Health Canada (3) have created large adverse event reporting systems (AERSs) that collect data from clinicians, patients, and pharmaceutical companies. These resources present an opportunity to monitor drug safety in a large and diverse population of patients. Quantitative signal detection algorithms use these data to flag and prioritize drug-event signals for follow-up analysis via formal pharmacoepidemiological studies and to discover complex relationships that are difficult to identify manually [such as drug-drug interactions (DDIs)] (4, 5). Despite their power, these methods suffer from well-recognized limitations that result from sampling variance and reporting biases (4, 6).

Signal detection algorithms quantify the "unexpectedness" of an adverse event being reported for a drug through disproportionality analysis; the goal is to identify drugs that have a greater proportion of a particular event among their reported events compared to the proportion seen for other drugs. Signals are detected by comparing the observed reporting rates between a drug-event pair to an expected reporting rate derived from other drug-event pairs. Under the null hypothesis that the event occurred by chance, the observed and expected rates will be equivalent and their ratio equal to one. When this ratio is much larger than one, the null hypothesis is rejected.

Unfortunately, there are a number of extraneous causes of differential reporting that fall into two distinct classes: (i) sampling variance and (ii) selection biases. Sampling variance refers to reporting rates that vary widely across drugs and time and depends on many factors. One source of sampling variance comes from the underreporting of events by physicians, who may report only ADEs that they deem to be important or that result from a new or untrusted drug. On the other hand, some ADEs can be oversampled. For example, in 2006, more than 18,000 reports were submitted to the FDA that associated rofecoxib (Vioxx) and myocardial infarction—likely a result of the intense media attention that occurred during that time. Sampling variance has been effectively addressed in modern signal detection algorithms, such as the gamma Poisson shrinker (GPS) or Information Component (IC) (6, 7). These methods estimate confidence intervals (CIs) for the disproportionality statistics and then dampen drug-event signals that have little evidence to support them. However, these methods do not address the issue of reporting biases (6).

Selection biases result from the nonrandom selection of subjects exposed to the drug and experiencing adverse events. This selection may be driven by causative covariates other than the drug under analysis (for example, a patient's disease state or other medications). This faulty selection may cause the disproportionality analysis to associate the drug and the event when a causative covariate is not accounted for; we refer to this as a synthetic association. Indication bias is one of the most common examples of this and occurs when a drug is synthetically associated with an event that is more appropriately attributed to the underlying disease (4). For example, it is common for diabetes drugs to be reported with hyperglycemia, a symptom of diabetes and usually not an effect of treatment. Similarly, concomitant medications can also confound drug-effect associations. Drugs commonly co-prescribed with rofecoxib (Vioxx) were more likely to be associated with heart attack simply because these drugs were commonly taken together. These issues extend to other covariates as well, both common and uncommon. Patients reported to be taking a cholesterol-lowering agent are more likely to be older, and this may cause these drugs to be synthetically associated

[1]Biomedical Informatics Training Program, Stanford University, Stanford, CA 94305, USA. [2]Departments of Bioengineering and Genetics, Stanford University, Stanford, CA 94305, USA. [3]Department of Medicine, Stanford University, Stanford, CA 94305, USA.
*To whom correspondence should be addressed. E-mail: russ.altman@stanford.edu

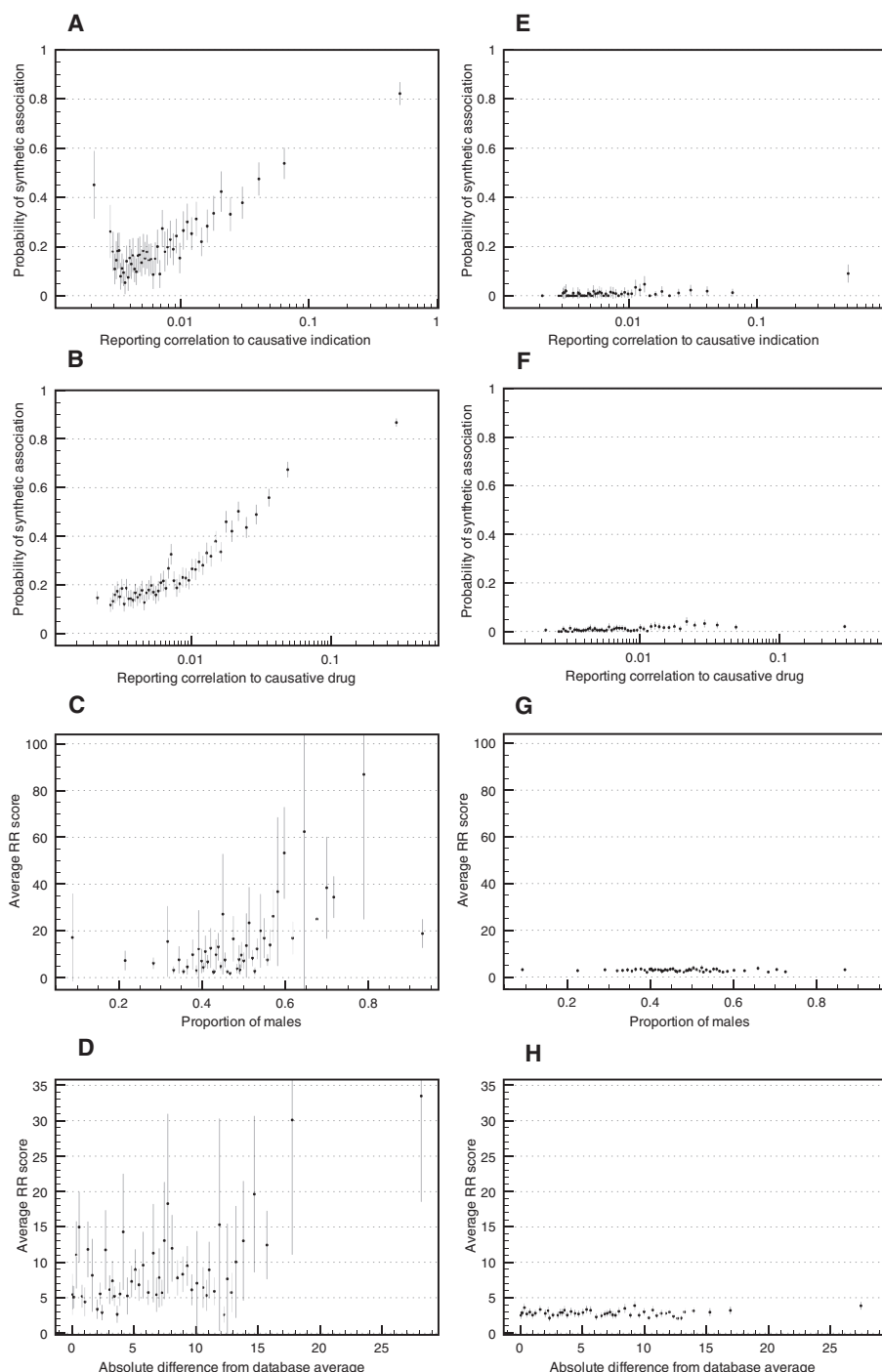**Fig. 1.** Synthetic associations in adverse event reports. (**A**) Disease indications are a significant source of synthetic associations. The more disproportionately a drug is reported with an indication (*x* axis), the more likely that drug will be synthetically associated with the indication's effects (*y* axis) (for example, it is common for hypoglycemic agents to be synthetically associated with hyperglycemia). (**B**) Concomitantly taken drugs are another significant source of synthetic associations. The more disproportionately two drugs are reported together (*x* axis), the more likely they will be associated with the other drug's effects (*y* axis). (**C**) Drugs that are preferentially reported with males are more likely to be synthetically associated with sex-related effects. (**D**) Similarly, drugs that are preferentially reported with relatively young or relatively old patients are more likely to be synthetically associated with age-related effects. (**E** to **H**) Application of SCRUB removes synthetic associations that result from disproportionate reporting with (E) disease indications, (F) concomitant drug use, (G) sex biases, and (H) age biases.

with age-related effects, such as hypertension or myocardial infarction (age bias). Patients who have recently had a renal transplant are often prescribed moxifloxacin, which has resulted in synthetic associations of the drug with renal impairment (prescribing bias).

The sources of bias are myriad and have not been directly addressed in modern signal detection algorithms. Indeed, stratification of the data on predefined covariates is the primary technique for removing these biases. However, stratification requires enumeration of all important covariates, which is computationally intractable to use for safety surveillance. Identifying a fixed set of likely covariates for all drugs can help the analysis, but reduces power when dividing the reports across strata that are not correlated with the outcome (*4*, *8*). In addition, stratification is impossible when the reporting systems do not contain reliable measures of the common covariates. These factors limit the benefit of applying stratification routinely (*8*).

Our new method (i) accomplishes the goals of stratification, dampening or removing the effect of covariates, without the need to divide drug-exposed reports into strata; (ii) is both adaptive (it removes different covariates for different drugs) and appropriate for systematic application and routine analysis; and (iii) is designed to complement modern signal detection approaches and thus extends the applicability and power of existing methods. Our model is inspired by the case-control approach to cohort selection in observational clinical studies. Each drug-exposed patient is matched to one (or more) nonexposed patients (controls). The nonexposed patients are selected on the basis of how well they match an exposed patient on a set of predefined covariates. Propensity score matching (PSM)—a statistical method designed to yield an unbiased estimate of treatment effects—has emerged as the preferred method of matching exposed and nonexposed patients in observational cohort studies and has yielded similar estimates of effects when compared to the results of randomized control trials (*9*–*11*). However, like other confounder controlling methods, PSM requires the covariates to be both known and measured; neither parameter is guaranteed to be present in spontaneous reporting systems. Instead, to match patients, we adapted PSM to use only the co-reported drugs and

co-reported indications. We hypothesize that many confounders correlate with these key variables and do not need to be modeled.

When we applied this data-driven approach to flag potentially significant drug-event associations in the AERS, we successfully removed many synthetic associations from indications, co-prescriptions, and hidden covariates. We call this new method the statistical correction of uncharacterized bias (SCRUB) and use it to construct comprehensive databases of off-label and DDI side effects (OFFSIDES and TWOSIDES, respectively). These databases contain information that is independent from the Side Effect Resource (SIDER), a database of drug effects mined from the package inserts (12). We demonstrate the biological use of these databases by showing improved performance (compared to SIDER) at predicting drug targets and drug indications (13, 14). Furthermore, we use the new methods to identify adverse drug class interactions and then corroborate 47 of the 395 predicted interactions with electronic medical records (EMRs) from Stanford University Hospital. Finally, we conclude our analysis with an association between adverse cardiovascular events and co-prescription of thiazides and selective serotonin reuptake inhibitors (SSRIs). Patients who take these drugs in combination are significantly more likely to have prolonged QT intervals than those who take thiazides or SSRIs alone.

## RESULTS

### Sources of synthetic associations: Disease indication, concomitant drug use, and characteristic biases

**Disease indication.** We manually constructed a set of 543 adverse events strongly associated with indications for which the indication and the adverse event have a known causative relationship. We call a drug-event association synthetic if it has a tight reporting correlation with the indication ($\rho \geq 0.1$) and a high relative reporting (RR) association score (RR $\geq$ 2). Drugs reported frequently with these indications were 80.0 (95% CI, 14.2 to 3132.8; $P < 0.0001$, Fisher's exact test) times as likely to have synthetic associations with indication events. We found a strong linear relationship between the correlation of reporting between a

drug and indication and the likelihood of a synthetic association ($\rho = 0.63$, $P < 0.0001$, Fig. 1A).

**Concomitant drugs.** We identified 1559 adverse events strongly associated with drugs and listed on the drug's package insert. These drug-event pairs represent a set of known strong positive associations. We call a drug-event association synthetic if it has a tight reporting correlation with the causative drug ($\rho \geq 0.1$) and a high association score (RR $\geq$ 2). Drugs co-reported with these drugs were 55.8 (95% CI, 29.3 to 122.1; $P < 0.0001$, Fisher's exact test) times as likely to have synthetic associations with the drug events. We found a strong linear relationship between the correlation of co-reporting between drugs and the likelihood of synthetic associations ($\rho = 0.93$, $P < 0.0001$, Fig. 1B).

**Characteristic biases (sex and age).** We identified 33 adverse events that, for physiological reasons, predominantly occur in males (for example, penile swelling and azoospermia; full list shown in table S1). We found that drugs that are disproportionately reported as causing adverse events in males were more likely to be synthetically associated with these events ($\rho = 0.57$, $P < 0.0001$, Fig. 1C). Similarly, we identified 48 adverse events that predominantly occur in either relatively young or relatively old patients (table S2). We found that drugs that are disproportionately reported to cause adverse events in relatively younger or relatively older patients (compared to the database average) were more likely to be synthetically associated with these events ($\rho = 0.60$, $P < 0.0001$, Fig. 1D).

### Correction of selection bias in adverse event report data

By identifying matched nonexposed reports to the exposed reports for each drug, SCRUB reduced the rate of synthetic associations that resulted from indications and concomitant drugs. We found that applying SCRUB removed 57% of the synthetic associations from indications and 49% of those that resulted from concomitant drugs. Further, we found that SCRUB preferentially dampened the signal of synthetic associations: odds ratio (OR) = 1.8 (95% CI, 1.4 to 2.4; $P < 0.0001$, Fisher's exact test) for associations from concomitant drugs and OR = 2.0 (95% CI, 1.8 to 2.3; $P < 0.0001$, Fisher's exact test) for those resulting from indications (Fig. 1, E and F). We evaluated the method against three
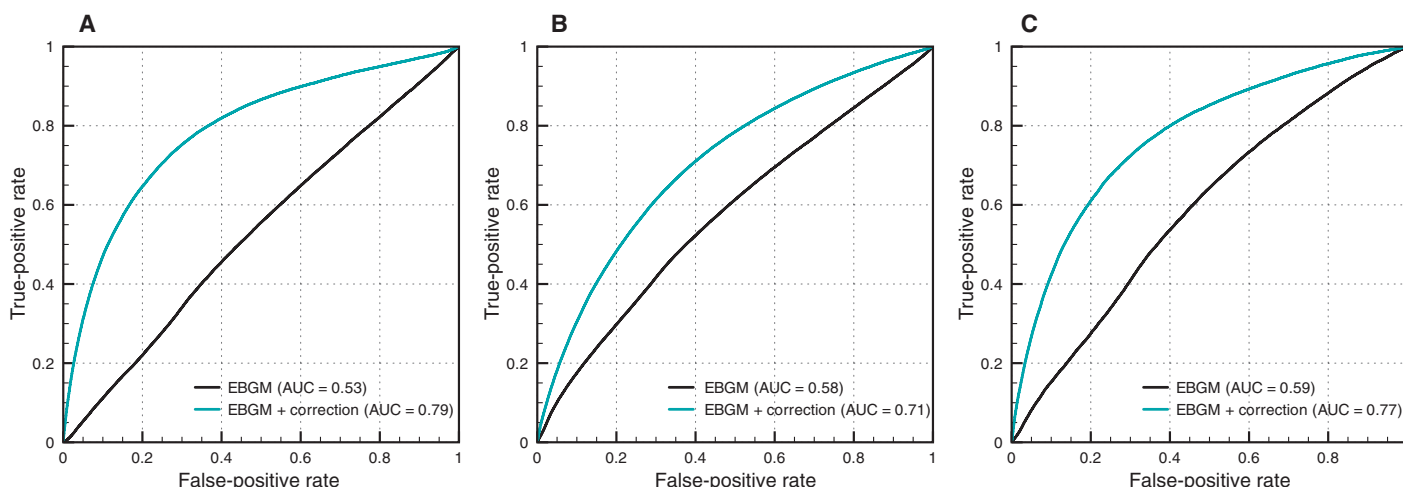
**Fig. 2.** Systematic evaluation against three independent silver standards of drug-effect associations. (**A**) Side effects mined from the package inserts. (**B**) Drug-effect pairs reported to the AERS after the original download date. (**C**) Drug-effect pairs reported to the Canadian system MedEffect. ROC curves for the empirical Bayes geometric mean (EBGM) (black) and a model combining EBGM and the correction factor derived from the SCRUB algorithm (aqua). In each case, including the correction term substantially improves the predictive power of the algorithm.

independent silver standards of ADEs: (i) side effects mined from the drug package inserts, (ii) adverse events reported to the FDA after our data extraction, and (iii) adverse event reports from Canada. We found that using SCRUB in combination with the GPS, a commonly used method for correcting sampling variance, significantly increased the predictive power in all three cases. We found that the area under the receiver operating characteristic curve (AUROC), a common method for evaluating the performance of predictive algorithms, increased from 0.53 to 0.79 ($\chi^2 = 19{,}130$, $P < 0.0001$), 0.58 to 0.71 ($\chi^2 = 13{,}963$, $P < 0.0001$), and 0.59 to 0.77 ($\chi^2 = 8598$, $P < 0.0001$) for each of the silver standards, respectively (Fig. 2). This finding is corroborated by several case studies reported in the Supplementary Materials (figs. S1 to S7).

## Correction of uncharacterized bias in adverse event report data

We hypothesized that the SCRUB algorithm would correct for synthetic associations caused by hidden, or unmeasured, covariates as well as those from indication and concomitant drug use. To test this notion, we hid age and sex data from the model and tested the ability of SCRUB to remove synthetic associations that resulted from these covariates. We found that the SCRUB-identified cohort matched more closely in age for 478 of 629 (76%) of drugs. We highlight the 20 most age-biased drugs and their cohort age differences in Fig. 3A. Similarly, we tested the ability of the algorithm to implicitly correct for sex differences. SCRUB-identified cohorts matched more closely in sex differences for 467 of 629 (74%) of drugs. Figure 3B highlights the top 20 drugs most biased in terms of sex of the patients and our corrections.

## Off-label and polypharmacy side-effect databases

We use the term "off-label" to refer to any drug effect not already listed on the drug's package insert. We constructed a database of 438,801 off-label side effects for 1332 drugs and 10,097 adverse events. The average drug label lists 69 "on-label" adverse events. We list an average of 329 high-confidence off-label adverse events for each drug. For comparison, the SIDER database, extracted from drug package inserts, lists 48,577 drug-event associations for 620 drugs and 1092 adverse events that are also covered by the data mining. OFFSIDES recovers 38.8% (18,842 drug-event associations) of SIDER associations from the adverse event reports. Thus, OFFSIDES finds different associations from those reported during clinical trials before drug approval. We found that the drug-event associations reported in OFFSIDES were predictive of known class-wide drug effects, such as the adverse events of the nervous systems associated with antiparasitics and insecticides (fig. S8).

In addition, we constructed a database of polypharmacy side effects for pairs of drugs (TWOSIDES). This database contains 868,221 significant associations between 59,220 pairs of drugs and 1301 adverse events. These associations are limited to only those that cannot be clearly attributed to either drug alone (that is, those associations covered in OFFSIDES). The database contains an additional 3,782,910 significant associations for which the drug pair has a higher side-effect association score, determined using the proportional reporting ratio (PRR), than those of the individual drugs alone. We found that the TWOSIDES database is enriched for pairs of drugs with known interactions ($t = 6.6$, $P = 4.9 \times 10^{-11}$).

## Use of drug-effect associations to predict protein targets and drug indications

Using a simplified version of the analysis by Campillos *et al.* (13), we computed pairwise similarity metrics between all drugs in the OFFSIDES
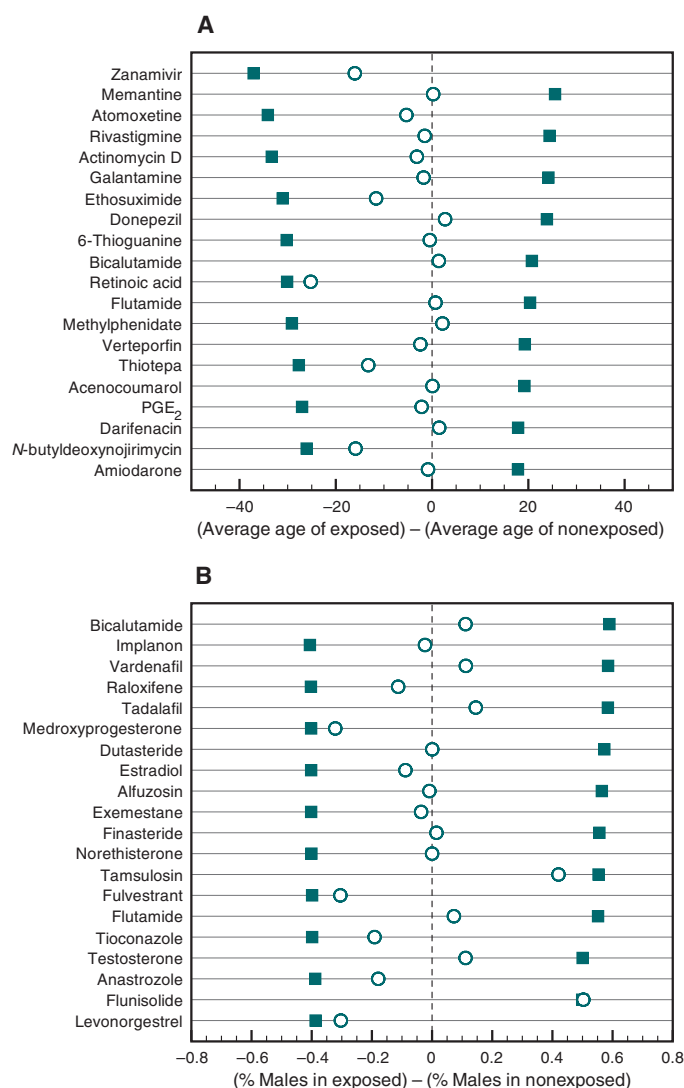


**Fig. 3.** Implicit matching of covariates. (**A**) Biases from age. Average age differences between cohorts of reports for those patients exposed to the drug and those who were not exposed (controls). The average difference for the uncorrected (solid squares) and corrected (open circles) nonexposed control reports is shown. Ideally, the difference between the two cohorts of reports is zero. (**B**) Biases from sex. Difference in the proportion of males reported to be exposed to the query drug versus those who were not exposed (controls). The difference for the uncorrected (solid squares) and corrected (open circles) nonexposed reports is shown. Ideally, this difference is zero.

and SIDER databases. We found that side-effect similarities derived from OFFSIDES were predictive of the proportion of shared targets between drugs (Fig. 4A, $\rho = 0.92$, $P < 0.0001$). For example, diazepam (Valium) and zolpidem (Ambien) share seven of the same protein targets, and although they have different chemical structures and are used for different indications, the two drugs have similar side-effect profiles. This similarity in side-effect profiles can be used to identify pairs of drugs likely to share targets and, when those associations are not yet known, to find new targets for existing drugs. Furthermore, similarities

derived from OFFSIDES provided information that was independent of similarities derived from SIDER (Fig. 4B) with respect to predicting shared targets ($\chi^2 = 177.2$, $P < 0.0001$), as determined by an analysis of variance (ANOVA). The model that was most predictive of shared targets was one that included information from both databases (AUROC = 0.75), followed by SIDER alone (AUROC = 0.71), and then OFFSIDES alone (AUROC = 0.70) (Fig. 4C).

Similarly, we found a linear relationship between the proportion of shared indications between a pair of drugs and the similarity of their side-effect profiles in OFFSIDES (Fig. 5A). This opens up the possibility of using side-effect profiles to suggest new uses for old drugs. Again, we found that OFFSIDES provided information independent of that provided by SIDER (Fig. 5B) with respect to predicting shared indications ($\chi^2 = 874.5$, $P < 0.0001$), as determined by an ANOVA. We found that a combination of the two databases performed best in terms of predicting existing therapeutic indications of known drugs (AUROC = 0.83), followed by SIDER alone (AUROC = 0.78), and then OFFSIDES (AUROC = 0.75) (Fig. 5C).

## Corroboration of class-wide interaction effects with EMRs

We used the TWOSIDES database to identify DDIs shared by an entire drug class (see Materials and Methods). Our class-class interaction analysis produced 1732 putative drug class interactions. We then identified laboratory reports commonly recorded in EMRs—for patients who visited Stanford University Hospital—that may be used as markers of these class-specific DDIs. We tested 596 of these interactions and found significant changes in the long-term ($\geq 1$ year) laboratory markers for 395 (66%; $P < 0.0001$, Fisher's exact test) of the interactions using a Cox proportional hazards model. Further, we found additional evidence of drug effects for 47 of 395 interactions when looking for short-term ($\leq 36$ days) changes in laboratory markers after the start of treatment (Table 1).

To test whether these discoveries were spurious fluctuations of the laboratory markers, we constructed a random set of interaction predictions for comparison. We found that our class-class interaction predictions were significantly enriched for interactions for which there was evidence in the EMRs (OR, 36.8; 95% CI, 6.2 to 1481.9; $P < 0.0001$,
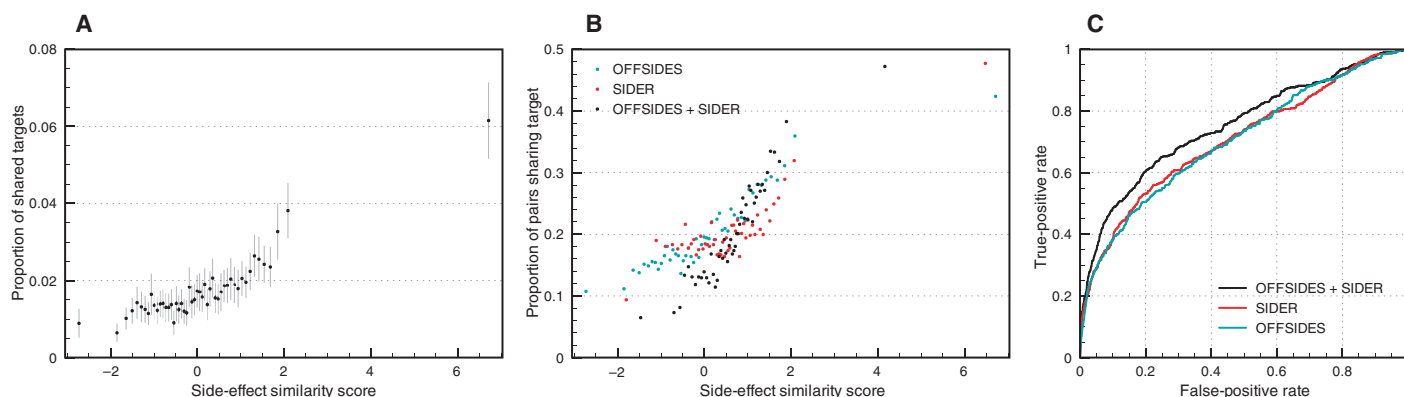
**Fig. 4.** Predicting shared protein targets using drug-effect similarities. (**A**) The side-effect similarity score between two drugs is linearly related to the number of targets that those drugs share. (**B**) A scatter plot showing the relationship between the side-effect similarity score and the number of shared targets for side effects derived from OFFSIDES (blue), SIDER (red), and both combined (black). (**C**) ROC curve representing the ability of the side-effect similarity scores to predict which pairs of drugs share targets. The best performance is reached by combining both data sets.
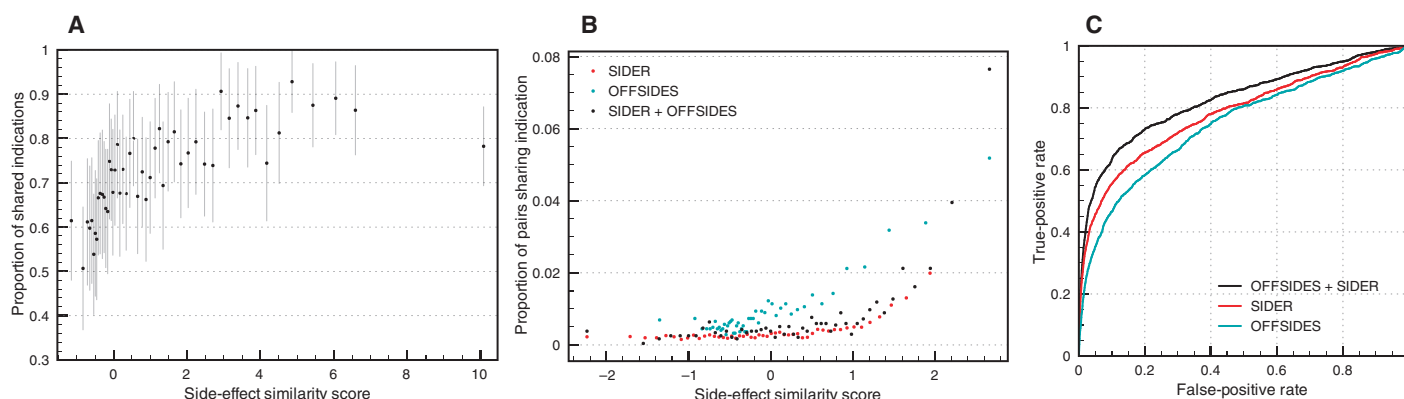


**Fig. 5.** Drug repurposing using drug-effect similarities in OFFSIDES. (**A**) The side-effect similarity score between two drugs is linearly related to the number of indications those drugs share. (**B**) Scatter plot showing the relationship between the side-effect similarity score and the number of shared targets for side effects derived from OFFSIDES (blue), SIDER (red), and both combined (black). (**C**) ROC curve representing the ability of the side-effect similarity scores to predict which pairs of drugs share indications. The best performance is reached by combining both data sets.

**Table 1.** Class-class interactions identified in Twosides and corroborated by EMRs. Forty-seven class-class interactions were corroborated by two retrospective analyses of EMRs: (i) short-term combination effects observed in the 36 days after combination treatment and (ii) long-term incidence of abnormal laboratory values associated with combination treatment. The posttreatment lab changes and long-term risk values shown are all statistically significant ($P < 0.05$) as determined by a paired $t$ test and Cox proportional hazards model, respectively (see the Supplementary Methods). CA, calcium; MCH, mean corpuscular hemoglobin; MCHC, MCH concentration; PLT, platelets; RDW, red cell distribution width; K, potassium; Mg, magnesium; QTC, corrected QT interval; Na, sodium; PCO2A, arterial $CO_2$; PTT, partial thromboplastin time; ALB, albumin; CR, creatine; GLOB, globulin; RBC, red blood cells; TP, total protein; AST, aspartate aminotransferase; TBIL, total bilirubin; BUN, blood urea nitrogen; ALKP, alkaline phosphatase.

| Drug class A (#Rx in millions) | Drugs (A) | Drug class B (#Rx in millions) | Drugs (B) | Effect category | Lab marker | Posttreatment lab change ± SEM | Long-term risk versus A (95% CI) | Long-term risk versus B (95% CI) |
|---|---|---|---|---|---|---|---|---|
| A01AB (13) | Doxycycline, minocycline, metronidazole tabs | G01AF (6) | Ketoconazole topical, metronidazole tabs | Cardiovascular | CA | −0.227 ± 0.013 mg/dl | 1.570 (1.658–1.486) | 1.107 (1.211–1.012) |
| | | | | Pulmonary | $CO_2$ | −0.200 ± 0.056 mM | 1.837 (1.981–1.703) | 1.359 (1.541–1.198) |
| A01AB (13) | Doxycycline, minocycline, metronidazole tabs | N05CF (30) | Zolpidem tartrate | Hemorrhage | MCH | −0.015 ± 0.017 g/dl | 1.671 (1.781–1.567) | 2.129 (2.257–2.008) |
| | | | | | MCHC | −0.001 ± 0.010 g/dl | 2.178 (2.396–1.980) | 2.523 (2.745–2.319) |
| | | | | | PLT | 14.113 ± 2.036 k/μl | 2.362 (2.563–2.178) | 3.179 (3.419–2.956) |
| A02BC (38) | Omeprazole, pantoprazole | N05CD (9) | Temazepam | Metabolic | K | −0.029 ± 0.007 mEq/liter | 1.434 (1.511–1.361) | 3.037 (3.303–2.793) |
| | | | | | RDW | 0.444 ± 0.024% | 1.317 (1.376–1.261) | 2.701 (2.889–2.526) |
| A07AA (2) | Nystatin topical, nystatin systemic | N02AA (27) | Morphine sulfate ER, oxycodone with APAP, hydromorphone HCl, oxycodone | Cardiovascular | CA | −0.234 ± 0.010 mg/dl | 1.131 (1.195–1.072) | 1.718 (1.782–1.656) |
| | | | | | Mg | −0.002 ± 0.006 mg/dl | 1.784 (1.991–1.598) | 7.407 (8.065–6.803) |
| A07AA (2) | Nystatin topical, nystatin systemic | N06AB (30) | Citalopram HBR, paroxetine, sertraline, fluoxetine | Cardiovascular | CA | −0.162 ± 0.020 mg/dl | 1.149 (1.238–1.066) | 1.506 (1.623–1.398) |
| | | | | | QTC | 3.853 ± 2.661 ms | 1.476 (1.676–1.299) | 3.009 (3.446–2.628) |
| A07EA (26) | Dexamethasone oral, hydrocortisone top. Rx, prednisone oral | S03AA (19) | Ciprofloxacin HCl | Cardiovascular | CA | −0.086 ± 0.016 mg/dl | 1.081 (1.138–1.028) | 1.417 (1.497–1.341) |
| | | | | | Mg | −0.006 ± 0.008 mg/dl | 4.708 (5.131–4.322) | 6.653 (7.413–5.970) |
| A10BB (9) | Glipizide ER, glipizide, glimepiride, glyburide | S03AA (19) | Ciprofloxacin HCl | Metabolic | Na | −0.375 ± 0.161 mEq/liter | 1.578 (1.761–1.413) | 1.560 (1.725–1.411) |
| B05XA (10) | Ferrous sulfate, potassium chloride | N05CF (30) | Zolpidem tartrate | Cardiovascular | CA | −0.265 ± 0.013 mg/dl | 1.299 (1.377–1.226) | 1.155 (1.215–1.097) |
| C03AA (46) | Hydrochlorothiazide | H02AB (26) | Triamcinolone acet. top, hydrocortisone top. Rx, prednisone oral, methylprednisolone tabs, dexamethasone oral | Pulmonary | PCO2A | 2.194 ± 0.698 mmHg | 3.077 (3.776–2.506) | 1.955 (2.351–1.626) |
| C03AA (46) | Hydrochlorothiazide | N06AB (30) | Citalopram HBR, paroxetine, sertraline, fluoxetine | Cardiovascular | QTC | 15.349 ± 5.675 ms | 1.609 (1.883–1.375) | 1.605 (1.878–1.372) |
| C09AA (74) | Lisinopril, ramipril, benazepril, enalapril, quinapril, fosinopril sodium | J02AC (13) | Fluconazole | Cardiovascular | Mg | −0.005 ± 0.019 mg/dl | 6.803 (8.143–5.685) | 1.630 (1.967–1.351) |
| C10AA (72) | Simvastatin, Pravastatin, Lovastatin | S02AA (19) | Ciprofloxacin HCl | Cardiovascular | K | −0.022 ± 0.009 mEq/liter | 1.838 (1.966–1.719) | 1.087 (1.167–1.013) |

| Drug class A (#Rx in millions) | Drugs (A) | Drug class B (#Rx in millions) | Drugs (B) | Effect category | Lab marker | Posttreatment lab change ± SEM | Long-term risk versus A (95% CI) | Long-term risk versus B (95% CI) |
|---|---|---|---|---|---|---|---|---|
| D01AC (13) | Fluconazole, ketoconazole topical | A02BC (38) | Omeprazole, pantoprazole | Metabolic disorders | CA | $-0.038 \pm 0.017$ mg/dl | 1.144 (1.226–1.068) | 1.471 (1.553–1.393) |
| D06AX (4) | Mupirocin | N06AB (30) | Citalopram HBR, paroxetine, sertraline, fluoxetine | Cardiovascular | K | $-0.079 \pm 0.037$ mEq/liter | 1.652 (2.036–1.340) | 2.271 (2.775–1.859) |
| | | | | | Mg | $-0.078 \pm 0.029$ mg/dl | 2.195 (2.912–1.655) | 3.975 (5.203–3.035) |
| D07AB (8) | Triamcinolone acet. top, dexamethasone oral | R03BA (18) | Triamcinolone acet. top, dexamethasone oral, fluticasone nasal | Cardiovascular | CA | $-0.390 \pm 0.011$ mg/dl | 1.262 (1.433–1.111) | 1.158 (1.200–1.117) |
| D07AC (18) | Fluocinonide, dexamethasone oral, fluticasone nasal | S03AA (19) | Ciprofloxacin HCl | Cardiovascular | K | $-0.027 \pm 0.010$ mEq/liter | 1.982 (2.160–1.818) | 1.773 (1.936–1.624) |
| D10AA (12) | Dexamethasone oral, methylprednisolone tabs | N06AX (17) | Mirtazapine, bupropion SR, bupropion ER, bupropion XL, venlafaxine, trazodone HCl | Cardiovascular | PTT | $0.158 \pm 0.897$ s | 1.210 (1.380–1.061) | 1.828 (2.092–1.598) |
| D10AF (7) | Clindamycin topical, clindamycin systemic | N06AB (30) | Citalopram HBR, paroxetine, sertraline, fluoxetine | Cardiovascular | ALB | $-0.353 \pm 0.048$ g/dl | 1.217 (1.427–1.039) | 1.647 (1.919–1.413) |
| G01AA (7) | Nystatin topical, clindamycin topical, nystatin systemic, clindamycin systemic | N03AX (23) | Gabapentin, lamotrigine, levetiracetam, topiramate | Pulmonary | $CO_2$ | $0.672 \pm 0.110$ mM | 1.311 (1.463–1.175) | 1.764 (1.965–1.583) |
| G01AF (6) | Metronidazole tabs, ketoconazole topical | N05CF (30) | Zolpidem tartrate | Gastrointestinal | ALB | $-0.272 \pm 0.015$ g/dl | 2.105 (2.256–1.964) | 3.059 (3.253–2.877) |
| G02CC (25) | Ibuprofen, naproxen | C08CA (10) | Nifedipine ER, felodipine ER, amlodipine besyl. benz. | Gastrointestinal | ALB | $-0.308 \pm 0.034$ g/dl | 1.349 (1.498–1.214) | 1.237 (1.374–1.115) |
| H03AA (63) | Levothyroxine | A02BC (38) | Omeprazole, pantoprazole | Gastrointestinal | PTT | $3.867 \pm 0.573$ s | 2.538 (2.770–2.325) | 1.085 (1.157–1.018) |
| J01DB (21) | Cephalexin | C09AA (74) | Benazepril, enalapril, lisinopril, ramipril, quinapril, fosinopril sodium | Metabolic | Na | $-0.624 \pm 0.075$ mEq/liter | 1.127 (1.217–1.044) | 1.162 (1.253–1.078) |
| J01DB (21) | Cephalexin | N05CD (9) | Temazepam | Cardiovascular | K | $0.015 \pm 0.010$ mEq/liter | 1.280 (1.407–1.165) | 1.215 (1.336–1.105) |
| | | | | | Mg | $-0.009 \pm 0.015$ mg/dl | 1.575 (1.822–1.360) | 1.640 (1.903–1.413) |
| | | | | Gastrointestinal | PTT | $1.792 \pm 0.584$ s | 1.353 (1.485–1.233) | 1.419 (1.559–1.290) |
| J01DD (8) | Cefdinir | R03BA (18) | Triamcinolone acet. top, dexamethasone oral, fluticasone nasal | Gastrointestinal | PTT | $-0.857 \pm 0.316$ s | 1.410 (1.625–1.224) | 2.353 (2.666–2.077) |
| J01FF (7) | Clindamycin systemic, clindamycin topical | C05AA (8) | Triamcinolone acet. top, dexamethasone oral, hydrocortisone top. Rx, fluocinonide | Pulmonary | $CO_2$ | $0.291 \pm 0.125$ mM | 1.211 (1.389–1.055) | 1.137 (1.286–1.005) |
| | | | | Cardiovascular | K | $-0.036 \pm 0.017$ mEq/liter | 1.817 (2.093–1.578) | 1.558 (1.761–1.378) |

| Drug class A (#Rx in millions) | Drugs (A) | Drug class B (#Rx in millions) | Drugs (B) | Effect category | Lab marker | Posttreatment lab change ± SEM | Long-term risk versus A (95% CI) | Long-term risk versus B (95% CI) |
|---|---|---|---|---|---|---|---|---|
| J01MA (19) | Ciprofloxacin HCl | A07AA (2) | Nystatin systemic, nystatin topical | Metabolic | CR | −0.001 ± 0.010 mg/dl | 1.565 (1.639–1.495) | 1.853 (1.969–1.745) |
| | | | | | GLOB | −0.020 ± 0.015 g/dl | 1.573 (1.659–1.491) | 1.853 (1.988–1.726) |
| | | | | | MCH | −0.059 ± 0.014 g/dl | 1.727 (1.818–1.642) | 1.984 (2.121–1.855) |
| | | | | | MCHC | −0.012 ± 0.009 g/dl | 2.211 (2.389–2.047) | 2.386 (2.643–2.153) |
| | | | | | RBC | −0.183 ± 0.007 ×$10^6$/μl | 2.287 (2.411–2.170) | 2.118 (2.269–1.977) |
| | | | | | RDW | 0.347 ± 0.028% | 2.144 (2.251–2.042) | 2.180 (2.324–2.045) |
| | | | | | TP | −0.340 ± 0.017 g/dl | 2.039 (2.165–1.920) | 2.275 (2.466–2.098) |
| J01MA (19) | Ciprofloxacin HCl | S03AA (19) | Ciprofloxacin HCl | Gastrointestinal | AST | 6.204 ± 7.721 U/liter | 1.690 (1.816–1.573) | 2.841 (3.273–2.466) |
| | | | | | TBIL | 0.139 ± 0.057 mg/dl | 1.418 (1.499–1.342) | 2.549 (2.843–2.286) |
| J02AB (3) | Ketoconazole topical | S03AA (19) | Ciprofloxacin HCl | Cardiovascular | K | −0.026 ± 0.013 mEq/liter | 2.035 (2.284–1.813) | 2.768 (3.002–2.552) |
| J05AB (5) | Acyclovir | B05XA (10) | Ferrous sulfate, potassium chloride | Metabolic | BUN | 1.000 ± 0.602 mg/dl | 1.246 (1.352–1.149) | 1.949 (2.108–1.802) |
| | | | | | Na | −0.388 ± 0.132 mEq/liter | 4.002 (4.348–3.682) | 4.717 (5.084–4.378) |
| M01AB (5) | Diclofenac sodium SR, etodolac, indomethacin | B05XA (10) | Ferrous sulfate, potassium chloride | Cardiovascular | CA | −0.227 ± 0.020 mg/dl | 1.230 (1.351–1.120) | 1.215 (1.334–1.106) |
| M03BX (21) | Cyclobenzaprine, baclofen, tizanidine HCl | G01AA (7) | Clindamycin systemic, nystatin topical, clindamycin topical, nystatin systemic | Pulmonary | $CO_2$ | 0.656 ± 0.143 mM | 1.501 (1.745–1.290) | 1.439 (1.671–1.239) |
| N02AB (4) | Fentanyl transdermal | B05XA (10) | Ferrous sulfate, potassium chloride | Cardiovascular | CA | −0.419 ± 0.012 mg/dl | 1.097 (1.151–1.045) | 1.116 (1.185–1.050) |
| | | | | | K | −0.001 ± 0.007 mEq/liter | 1.841 (1.964–1.726) | 1.303 (1.409–1.204) |
| N02AB (4) | Fentanyl transdermal | C05AA (8) | Triamcinolone acet. top, dexamethasone oral, hydrocortisone top. Rx, fluocinonide | Cardiovascular | K | −0.012 ± 0.008 mEq/liter | 1.949 (2.078–1.827) | 1.590 (1.701–1.486) |
| N02AB (4) | Fentanyl transdermal | G04BD (2) | Oxybutynin chloride ER, oxybutynin chloride | Cardiovascular | Mg | −0.061 ± 0.031 mg/dl | 1.389 (1.882–1.025) | 2.717 (3.942–1.872) |
| N05BA (44) | Diazepam, lorazepam, alprazolam | N05AB (2) | Prochlorperazine maleate | Cardiovascular | CA | −0.202 ± 0.013 mg/dl | 2.024 (2.112–1.940) | 2.583 (2.768–2.411) |
| N06AB (30) | Citalopram HBR, paroxetine, sertraline, fluoxetine | A07AA (2) | Nystatin systemic, nystatin topical | Metabolic | CA | −0.162 ± 0.020 mg/dl | 1.427 (1.537–1.325) | 1.187 (1.280–1.101) |
| | | | | | $CO_2$ | 0.139 ± 0.102 mM | 1.866 (2.046–1.702) | 1.304 (1.430–1.189) |
| | | | | | GLOB | 0.016 ± 0.029 g/dl | 2.005 (2.194–1.833) | 1.209 (1.322–1.106) |
| | | | | | MCHC | −0.015 ± 0.017 g/dl | 2.729 (3.116–2.390) | 1.223 (1.390–1.076) |
| | | | | | QTC | 3.853 ± 2.661 ms | 2.608 (2.976–2.284) | 1.513 (1.720–1.330) |
| | | | | | RDW | 0.383 ± 0.048% | 2.924 (3.173–2.695) | 1.313 (1.421–1.213) |

| Drug class A (#Rx in millions) | Drugs (A) | Drug class B (#Rx in millions) | Drugs (B) | Effect category | Lab marker | Posttreatment lab change ± SEM | Long-term risk versus A (95% CI) | Long-term risk versus B (95% CI) |
|---|---|---|---|---|---|---|---|---|
| N06AB (30) | Citalopram HBR, paroxetine, sertraline, fluoxetine | S02AA (19) | Ciprofloxacin HCl | Metabolic | GLOB | −0.028 ± 0.026 g/dl | 1.665 (1.810–1.531) | 1.144 (1.239–1.056) |
| | | | | | RDW | 0.356 ± 0.043% | 2.376 (2.566–2.201) | 1.368 (1.469–1.274) |
| | | | | | MCHC | 0.004 ± 0.018 g/dl | 2.076 (2.396–1.799) | 1.208 (1.387–1.052) |
| | | | | | RDW | 0.395 ± 0.049% | 2.157 (2.351–1.978) | 1.320 (1.436–1.214) |
| N06AX (17) | Mirtazapine, bupropion SR, bupropion ER, bupropion XL, venlafaxine, trazodone HCl | A07AA (2) | Nystatin systemic, nystatin topical | Cardiovascular | CA | −0.101 ± 0.019 mg/dl | 1.401 (1.509–1.301) | 1.249 (1.345–1.160) |
| N06AX (17) | Mirtazapine, bupropion SR, bupropion ER, bupropion XL, venlafaxine, trazodone HCl | S01AX (19) | Ciprofloxacin HCl | Cardiovascular | CA | −0.067 ± 0.016 mg/dl | 1.308 (1.399–1.222) | 1.198 (1.277–1.123) |
| N06AX (17) | Mirtazapine, bupropion SR, bupropion ER, bupropion XL, venlafaxine, trazodone HCl | S01AX (19) | Ciprofloxacin HCl | Cardiovascular | CA | −0.064 ± 0.016 mg/dl | 1.305 (1.397–1.220) | 1.181 (1.260–1.107) |
| R03CC (9) | Albuterol nebulizer sol. | C10AA (72) | Simvastatin, pravastatin, lovastatin | Cardiovascular | Mg | 0.063 ± 0.017 mg/dl | 1.754 (2.073–1.484) | 1.850 (2.135–1.603) |
| R06AE (5) | Meclizine HCl | S02AA (19) | Ciprofloxacin HCl | Metabolic | MCH | −0.128 ± 0.071 g/dl | 1.753 (2.146–1.432) | 1.255 (1.525–1.033) |
| S01AD (5) | Acyclovir | B05XA (10) | Ferrous sulfate, potassium chloride | Cardiovascular | Mg | −0.057 ± 0.011 mg/dl | 7.087 (8.224–6.109) | 7.669 (8.584–6.854) |
| S01AD (5) | Acyclovir | N03AX (23) | Gabapentin, lamotrigine, levetiracetam, topiramate | Cardiovascular | K | −0.064 ± 0.022 mEq/liter | 1.450 (1.689–1.245) | 2.611 (3.020–2.257) |
| S01AX (19) | Ciprofloxacin HCl | A07AA (2) | Nystatin systemic, nystatin topical | Metabolic | ALB | −0.364 ± 0.010 g/dl | 1.744 (1.827–1.664) | 1.912 (2.042–1.789) |
| | | | | | CR | 0.009 ± 0.010 mg/dl | 1.499 (1.567–1.433) | 1.685 (1.791–1.584) |
| | | | | | MCH | −0.045 ± 0.013 g/dl | 1.653 (1.736–1.573) | 1.818 (1.947–1.698) |
| | | | | | MCHC | −0.002 ± 0.008 g/dl | 2.111 (2.278–1.957) | 2.092 (2.325–1.883) |
| | | | | | RDW | 0.394 ± 0.026% | 2.045 (2.144–1.952) | 2.018 (2.155–1.889) |
| | | | | | TP | −0.417 ± 0.017 g/dl | 1.960 (2.079–1.849) | 2.094 (2.276–1.926) |
| S01AX (19) | Ciprofloxacin HCl | D06AX (4) | Mupirocin | Gastrointestinal | ALB | −0.135 ± 0.025 g/dl | 1.377 (1.514–1.253) | 2.399 (2.743–2.098) |
| | | | | | TBIL | 1.017 ± 0.222 mg/dl | 1.425 (1.576–1.289) | 2.106 (2.420–1.833) |
| S01BA (8) | Dexamethasone oral, triamcinolone acet. top, hydrocortisone top. Rx | N01AH (4) | Fentanyl transdermal | Cardiovascular | Mg | 0.031 ± 0.010 mg/dl | 2.796 (3.125–2.501) | 3.403 (3.923–2.951) |

*continued on next page*

| Drug class A (#Rx in millions) | Drugs (A) | Drug class B (#Rx in millions) | Drugs (B) | Effect category | Lab marker | Posttreatment lab change ± SEM | Long-term risk versus A (95% CI) | Long-term risk versus B (95% CI) |
|---|---|---|---|---|---|---|---|---|
| S01BC (5) | Diclofenac sodium SR, indomethacin | B05XA (10) | Ferrous sulfate, potassium chloride | Cardiovascular | CA | −0.222 ± 0.020 mg/dl | 1.226 (1.347–1.115) | 1.195 (1.313–1.088) |
| S01CB (2) | Hydrocortisone top. Rx, dexamethasone oral | N01AH (4) | Fentanyl transdermal | Cardiovascular | Mg | 0.034 ± 0.010 mg/dl | 2.372 (2.657–2.118) | 3.360 (3.871–2.917) |
| S01EA (11) | Clonidine | A04AA (3) | Ondansetron | Cardiovascular | K | −0.055 ± 0.011 mEq/liter | 2.301 (2.595–2.040) | 2.054 (2.192–1.925) |
| S01XA (4) | Ferrous sulfate | N05CF (30) | Zolpidem tartrate | Cardiovascular | CA | −0.263 ± 0.009 mg/dl | 1.211 (1.258–1.165) | 1.419 (1.478–1.363) |
| S02BA (2) | Hydrocortisone top. Rx, dexamethasone oral | N01AH (4) | Fentanyl transdermal | Cardiovascular | Mg | 0.034 ± 0.010 mg/dl | 2.356 (2.639–2.103) | 3.362 (3.873–2.919) |
| V03AB (4) | Ferrous sulfate | S01AX (19) | Ciprofloxacin HCl | Gastrointestinal | ALB | −0.365 ± 0.010 g/dl | 1.719 (1.800–1.641) | 1.668 (1.753–1.586) |
| | | | | | ALKP | 7.739 ± 1.842 U/liter | 1.516 (1.583–1.452) | 1.316 (1.378–1.257) |

Fisher's exact test). Figure 6 summarizes our drug effect and interaction findings for cardiovascular adverse events.

## Association of co-prescription of thiazides and SSRIs with a prolonged QT interval

The DDI with the largest effect size was an association between co-prescriptions of thiazides and serotonin reuptake inhibitors (SSRIs) with prolonged QT intervals (QTc > 440 ms) (Supplementary Materials). Prolonged QT intervals on an electrocardiogram have been associated with increased risk of spontaneous arrhythmias and sudden death. For EMR analysis, we removed patients who had a previous history of prolonged QT. Of 932 patients who were co-prescribed a thiazide and an SSRI, 87 (9.3%) displayed prolonged QT, whereas 588 of 9008 (6.5%) patients who were prescribed thiazides alone had prolonged QT, and 684 of 14,218 (4.8%) patients who were prescribed SSRIs alone had prolonged QT. Using Cox proportional hazards regression with covariates, we performed a time-to-event analysis on the association between co-prescription of thiazides and SSRIs and the incidence of prolonged QT. This analysis showed that patients who were co-prescribed a thiazide and an SSRI were 1.5 (95% CI, 1.2 to 1.9) times as likely ($P = 4.46 \times 10^{-4}$) to record a prolonged QT interval when compared to patients prescribed a thiazide alone and 1.4 (95% CI, 1.2 to 1.8) times as likely ($P = 0.0013$) as those prescribed an SSRI alone (Fig. 7). We ruled out the possibility that other co-prescribed medications may have been associated with QT prolongation by testing 38 commonly co-prescribed drugs in independent regression models and, in each case, found a significant effect from the combined use of thiazides and SSRIs (table S3). We corrected for the effects of co-prescriptions, age, race, and sex (table S4). Together, these results suggest that further study of the potential interaction between thiazides and SSRIs may be warranted.

## DISCUSSION

The methods we present here build upon the foundation of signal detection algorithms developed for drug safety surveillance. The use of spontaneous reporting systems for identifying ADEs faces challenges as a result of sampling variance and reporting biases (4, 6). Modern signal detection algorithms address the issue of sampling variance by using shrinkage to down-weight drug-event associations with little evidence to support them (6, 7). Stratification is designed to address reporting biases by dividing the data across covariate-defined strata. However, systematic application of stratification using a fixed set of covariates reduces power by dividing up the available data across unimportant strata (4, 15). Our approach does not divide data across strata and can correct for the effects of confounders even if those variables are unknown or unmeasured. The key insight is that, at least for drugs, the indications of use and other drugs used capture most of many important covariates. Although our approach is inspired by those used in observational cohort analysis, it does not enable causative inference. Like other signal detection techniques, the goals are to generate quality hypotheses for follow-up analysis. Our method has a comparable running time to current techniques, making it suitable for systematic drug surveillance.

The successful prediction of side effects before a drug enters clinical trials remains a tantalizing goal. Chemical informatics techniques can predict drug side effects by comparing the structural similarity of drugs (16, 17). In an analogous manner, protein structural similarity can explain and predict drug side effects (18). More recently, network and chemical properties have been combined together into predictive models of drug effects (19); these approaches all rely on a comprehensive database of known drug effects. Package inserts list drug side effects and could serve as a primary source of known side effects, but these data are limited. First, because clinical trials are conducted on relatively small patient populations, only common effects can be detected with sufficient confidence to be listed on a drug's package insert. Second, effects observed during the clinical trials may be incidental and not actually caused by the drug. Nonetheless, recent work in chemical biology has used the SIDER (a text-mined database of drug package inserts) to good effect (12, 13, 20). Our OFFSIDES database contains information complementary to that found in SIDER and improves the prediction of protein targets and drug indications. As a complement to OFFSIDES, our TWOSIDES database of
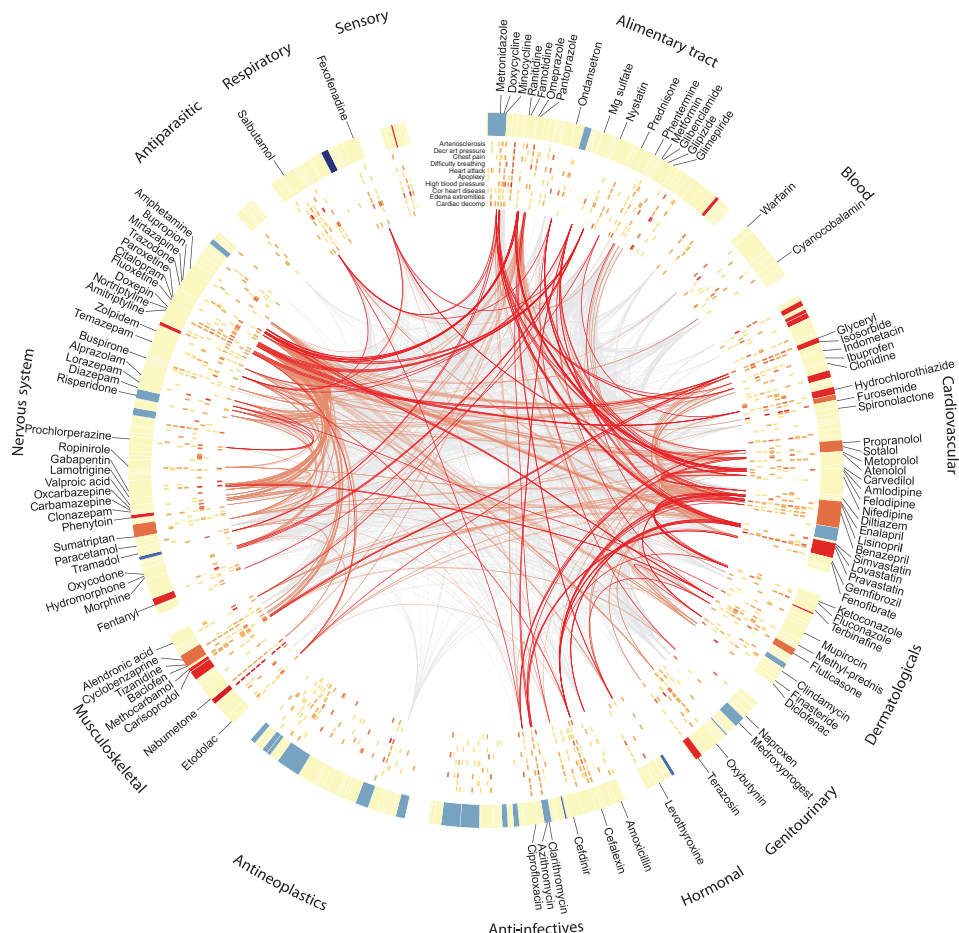
**Fig. 6.** Interaction diagram depicting single-drug effects, drug-class effects, DDIs, and class-class interactions for cardiovascular adverse events. Drugs are sorted clockwise around the ring by the physiological system they treat. Drugs labeled by name are members of data-mined DDIs. Within each physiological system, drugs are grouped into lower-order drug classes according to structural similarity or treatment indication. These lower-order classes are colored by their class-wide association with adverse cardiovascular effects (red for most severe to blue for least severe). Each arc across the center represents one DDI according to the data mining. The arc is colored red if the drug interaction is corroborated with evidence from the EMRs and brown if the drugs are members of class-class interactions. The heat map around the interior of the ring indicates the individual drug effects with the top 10 cardiovascular adverse events (arteriosclerosis, decreased arteriole pressure, chest pain, difficulty breathing, heart attack, apoplexy, high blood pressure, coronary heart disease, edema in extremities, cardiac decompression) (dark red for strong associations to white for weak or no association).

tion of these methods with TwoSIDES may lead to further understanding of the molecular etiology of these effects (figs. S9 and S10). We highlight one potentially clinically significant association between co-prescription of thiazides and SSRIs and QT interval prolongation. Prolonged QT is not a known interaction effect of thiazides and SSRIs. However, each drug class is individually implicated in causing hyponatremia (32–34), and the mechanisms that cause this side effect may interact synergistically. The EMR analyses we report are not full epidemiological studies. EMR records are incomplete and may be missing data on medical history and prescription orders. In addition, patients who take multiple drugs may have a higher rate of adverse events than less-medicated patients. Further analysis is needed to evaluate these potentially important drug interactions.

Evaluation of signal detection algorithms and side-effect prediction algorithms, in general, is not straightforward; no gold standard of known ADEs exists. In lieu of a standard, we evaluated our proposed methodology against three "silver" standards: (i) effects listed on the drug's package inserts, (ii) ADEs reported after the original download date of September 2009, and (iii) ADEs reported to the Canadian spontaneous reporting system. We found that when used in combination with modern signal detection algorithms, our method significantly improved performance. These standards, however, are biased toward more common effects, and so the performance of our method with respect to detecting rare events may be less reliable. A publicly available resource of drug effects would enhance the evaluation of this and other predictive algorithms.

In summary, we present a new methodology for correcting for the effects of confounding variables in large clinical observational databases when those variables are unknown, unmeasured, or sparsely collected. The goals of this work parallel those of patient stratification; however, our presented methodology adapts to specific drug-event pairs, does not require data to be split across strata, and can implicitly correct for unmeasured covariates. The key assumption of the method is that many patient covariates will be represented by the concomitant drugs the patient is taking and indications for which the patient is being treated. The method improves the performance of modern signal detection techniques and is suitable for systematic and routine safety surveillance. Finally, we present two new resources of adverse drug effects and drug interactions for use in drug discovery, repositioning, chemical biology, and pharmacoepidemiology studies.
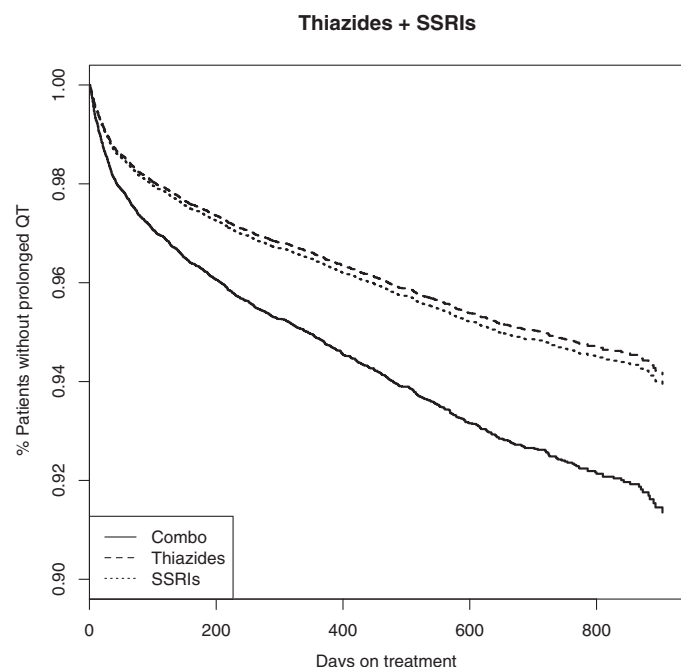
mined putative DDIs also lists predicted adverse events. These databases will serve as valuable resources for chemical biology, drug discovery, and pharmacoepidemiology studies. These databases are made available in the Supplementary Materials and at the http://PharmGKB.org Web site.

Identification and prediction of DDIs is a critical activity for improved patient care (21–26). Clinical trials do not routinely investigate DDIs because they are focused on establishing safety and efficacy of single-agent therapeutics. A wide range of methods, from text mining (27, 28) to network modeling (29, 30), can detect, explain, and predict DDIs. Recently, a systems pharmacology approach was presented to identify genes associated with adverse cardiovascular drug effects (31). Integra-

**Thiazides + SSRIs**



**Fig. 7.** Kaplan-Meier curves showing the proportion of patients that had prolonged QT corrected values after the start of drug therapy. The solid line represents patients who received both thiazides and SSRIs, the dashed line represents patients who received only thiazides, and the dotted line represents patients who received only SSRIs.

## MATERIALS AND METHODS

**Data source** large adverse event reporting systems (AERSs)

We downloaded the following: (i) 1,851,171 adverse event reports in the AERS from the FDA's Web site from the first quarter of 2004 to the first quarter of 2009; (ii) the SIDER, a database of the drugs, adverse events, and indications mined from the FDA drug labels (*12*) and Canada's MedEffect resource, the sister database to the AERS containing about 300,000 adverse event reports (downloaded September 2009); (iii) the drug target information from the DrugBank, Matador, and Psychoactive Drug Screening Program (PDSP) chemical databases referenced by Campillos and colleagues for use in correlating side-effect similarity to shared drug targets (*13*); and (iv) an independent database of the adverse event reports in AERS for the third quarter of 2009 to the fourth quarter of 2010 for validation purposes.

### Statistical model and assumptions

Not all ADEs that occur are captured by spontaneous reporting systems. The drug effects need to be observed, recognized, attributed to a drug, and then reported. Therefore, differential reporting and covariate biases prohibit a straightforward interpretation of the reports. Disproportionality analysis addresses some of these issues by looking for drugs that are disproportionately reported with a particular event compared to that same proportion for other drugs. They do so by comparing the observed number of reports to an expected number estimated from the proportions of other drugs. In our model, under the null hypothesis, we view the observed ($O_{xy}$) and expected ($E_{xy}$) as biased estimators for the incidence ($I_{xy}$), where the incidence is the rate at which the event would be reported absent of any confounding variables. We can then write the observed-to-expected ratio as follows:

$$\frac{I_{xy} + \varepsilon}{I_{xy} + \beta}$$

where $\beta$ is the bias of $E_{xy}$ and $\varepsilon$ is the bias in $O_{xy}$. It is clear that synthetic associations occur when the bias in the observed is greater than the bias in the expected, and in general, any time the biases are not equal, errors may occur. Because we do not have complete knowledge of the patients who are prescribed any given drug and the events that occur, we cannot compute the bias terms directly. Instead, we adapted the tools of cohort selection in observational studies to match each exposed case report to a nonexposed control report. PSM models the probability that a patient (also known as a report in our case) is selected into the exposed group versus the nonexposed group as a function of the available covariates using logistic regression (*10*). Each exposed patient (that is, report) is matched to a nonexposed patient with a similar probability according to the PSM model, thereby mitigating the effects of confounders. In general, AERSs do not collect data on all of the covariates necessary to implement a PSM model. However, our hypothesis is that many of the important covariates for a patient will be captured by the concomitant medications that patient is taking and the indications they are being treated for. These are data that are collected by spontaneous reporting systems, and these are the variates we use in our PSM model.

### Identification of nonexposed control reports to estimate expected values

For each drug, *x*, we used PSM to model the probability that a given report lists *x* as a concomitant medication. That is, the dependent variable for the logistic regression model was an indicator variable for the listing of drug *x* on the adverse event report. The independent variables were other drugs and indications. However, rather than using all drugs and indications (which would lead to a model with thousands of features), we used only drugs and indications found to be preferentially reported with the given drug, *x*. In addition, we limited the total number of possible features to the top 200 covariates (sorted by their Spearman correlation coefficient, $\rho$). Finally, we removed any reports that listed none of the chosen features (that is, their feature vector would be all zeroes). In this way, we built a PSM for each of the 632 drugs and used the model to score each report for each drug to compute its PSM score. Now, for each drug, we have generated PSM scores for each of the exposed reports and a subset of the nonexposed reports. We divide the exposed reports into 20 equally spaced bins on the basis of their PSM scores, and for each bin, we sample, with replacement, nonexposed control reports with PSM scores within the bin range until we have 10 times as many nonexposed reports as exposed reports. We discard any bins that have no matching controls available. The result is two sets of reports for each drug (exposed and nonexposed) from which we compute the observed and expected ratios and the resulting disproportionality statistics. To identify DDIs, we used a slightly modified approach that was computationally more efficient (table S5).

### Construction of the silver standard sets of drug-event associations

To evaluate the presented method, we tested its predictive performance against three standards of drug effects and compared this performance

to the GPS. No unbiased gold standard for ADEs exists. The drug-event associations from the FDA drug labels are the most obvious option for comparison. However, it is important to note that the labels are biased toward the more common adverse events that are observed and reported in premarketing clinical trials. This bias will limit the applicability of the drug labels because the goal of signal detection is to identify rare and unexpected side effects of drugs. An independent adverse event database, such as Canada's MedEffect database, can also be used for evaluation. However, because such a database will suffer from the same types of errors, it is necessary to take only a subset of high-confidence associations. We extracted only those associations where there was only one drug listed on the report (according to the publicly available structured data) under the assumption that if only one drug is listed, then it is the causative agent. Similarly, to asses reproducibility, we used a subset of AERS (quarter 3 of 2009 to quarter 4 of 2010) that was not used in the original analysis as a third silver standard. Again, to mitigate confounding effects, we use only those reports that list exactly one drug and flag that drug as the primary suspect.

### Using drug side-effect similarities to predict drug targets and indications

Previous work has shown that a drug's side effects can be used to predict protein targets. Specifically, Campillos and colleagues have shown that if two drugs are similar in the side effects they elicit, then they are more likely to share a common drug target (13). As validation of the biological relevance of the methods we present, we replicated this result in our mined associations. We calculated the similarity between two drugs by computing the Tanimoto coefficient between the drug's adverse event bit vectors (in these adverse event bit vectors, each bit represents one adverse event and is set if the drug has a significant association with the adverse event). Some drugs have higher similarity scores on average using this metric, so we perform a z-score normalization by drug. We calculated these "z similarities" for both the SIDER data set (the side effects extracted from the drug's package inserts) and the OFFSIDES data set. We tested the similarity score's ability to predict the number of targets two drugs share using a multivariate linear regression (modeling proportion of shard targets) and logistic regression (modeling shared targets), and tested for independence between SIDER and OFFSIDES using an ANOVA.

Recent studies have also shown that drug indications can be predicted using side-effect similarities (14). We use the same side-effect similarity scores computed for target identification to predict drug indications. We tested the similarity score's ability to predict the number of targets two drugs share using a multivariate linear regression (modeling proportion of shard targets) and logistic regression (modeling shared targets), and tested for independence between SIDER and OFFSIDES using an analysis of covariance (ANCOVA).

### Construction of the off-label and polypharmacy side-effect databases

The side-effect resources we present and made publicly available are a subset of the associations analyzed in the analysis. In the OFFSIDES database, we removed any associations that are not nominally significant (uncorrected $P > 0.05$); the remaining associations were included. In the TWOSIDES database, we also removed associations that were not significant. In addition, we removed any associations where there is evidence, according to OFFSIDES, that one drug of the pair is likely responsible for the adverse event. This step improves the chances that

the reported drug interaction effects are due to synergistic interactions and not recapitulations of known effects.

### Associating drug classes to adverse event categories from OFFSIDES

We associated Anatomic Therapeutic Chemical (ATC) drug classes (levels 1 and 4) with adverse event categories by linear modeling. To facilitate the statistical modeling, we constructed a table where the rows are all drug-event pairs determined to have significant associations by SCRUB. The model contained two features: (i) indicator variable of the membership of the drug in the ATC drug class and (ii) indicator variable of the membership of the event in the category. The dependent variable is the reporting frequency for the drug-event pair observed in the AERS. For each class-category pair, we then modeled the reporting frequency between all drugs and adverse events as a function of the two indicator variables and an interaction term between the two indicator variables. We filtered for those class-category associations where the interaction term was significant, after multiple hypothesis correction.

Methodological detail covering the statistical analysis of the drug and indication case studies, computing the drug-effect association statistics, identifying DDI effects and the analytical methods for validating acute, and long-term effects using EMR data can be found in the Supplementary Materials and tables S4, S6, and S7. Institutional review board approval was obtained for the EMR studies.

## SUPPLEMENTARY MATERIALS

www.sciencetranslationalmedicine.org/cgi/content/full/4/125/125ra31/DC1
Materials and Methods
Fig. S1. Case study: abacavir and rash.
Fig. S2. Case study: isoniazid and hepatic failure.
Fig. S3. Case study: pergolide and heart valve damage.
Fig. S4. Case study: rofecoxib and myocardial infarction.
Fig. S5. Case study: arrhythmias.
Fig. S6. Case study: hypercholesterolemia.
Fig. S7. Case study: hyperglycemia.
Fig. S8. Heat map of the interaction coefficients between adverse event categories and drug classes.
Fig. S9. Performance of disproportionality statistics for drug-drug interactions.
Fig. S10. Trends in disproportionality statistics for drug-drug interactions.
Table S1. Adverse events primarily reported with males.
Table S2. Age-related adverse events.
Table S3. Effect of commonly co-prescribed drugs on association with thiazides and SSRIs.
Table S4. Covariates in Cox regression model of thiazides and SSRIs.
Table S5. Drug class–adverse event category associations in OFFSIDES.
Table S6. Labs used as markers for adverse event categories.
Table S7. Normal lab value ranges.
References
Tab delimited database file: offsides.tsv
Tab delimited database file: twosides.tsv

## REFERENCES AND NOTES

1. D. W. Bates, D. J. Cullen, N. Laird, L. A. Petersen, S. D. Small, D. Servi, G. Laffel, B. J. Sweitzer, B. F. Shea, R. Hallisey, M. V. Vliet, R. Nemeskal, L. L. Leape, ADE Prevention Study Group; L. L. Leape, D. Servi, N. Laird, D. Bates, P. Hojnowski-Diaz, L. A. Petersen, S. Petrycki, M. V. Vliet, M. Cotugno, H. Patterson, B. F. Shea, M. Hickey, S. Kleefield, J. Cooper, D. J. Cullen, E. Kinnealy, R. Nemeskal, B. J. Sweitzer, S. D. Small, H. J. Demonaco, M. D. Clapp, R. Hallisey, T. Gallivan, J. Ives, K. Porter, B. T. Thompson, G. Laffel, J. R. Hackman, A. Edmondson, Incidence of adverse drug events and potential adverse drug events. Implications for prevention. *JAMA* **274**, 29–34 (1995).
2. D. C. Classen, S. L. Pestotnik, R. S. Evans, J. F. Lloyd, J. P. Burke, Adverse drug events in hospitalized patients. Excess length of stay, extra costs, and attributable mortality. *JAMA* **277**, 301–306 (1997).

3. A. Szarfman, J. M. Tonning, P. M. Doraiswamy, Pharmacovigilance in the 21st century: New systematic tools for an old problem. *Pharmacotherapy* **24**, 1099–1104 (2004).

4. A. Bate, S. J. Evans, Quantitative signal detection using spontaneous ADR reporting. *Pharmacoepidemiol. Drug Saf.* **18**, 427–436 (2009).

5. A. Szarfman, S. G. Machado, R. T. O'Neill, Use of screening algorithms and computer systems to efficiently signal higher-than-expected combinations of drugs and events in the US FDA's spontaneous reports database. *Drug Saf.* **25**, 381–392 (2002).

6. W. DuMouchel, Bayesian data mining in large frequency tables, with an application to the FDA spontaneous reporting system. *Am. Stat.* **53**, 177–190 (1999).

7. G. N. Norén, A. Bate, R. Orre, I. R. Edwards, Extending the methods used to screen the WHO drug safety database towards analysis of complex associations and improved accuracy for rare events. *Stat. Med.* **25**, 3740–3757 (2006).

8. J. Hopstadius, G. N. Norén, A. Bate, I. R. Edwards, Impact of stratification on adverse drug reaction surveillance. *Drug Saf.* **31**, 1035–1048 (2008).

9. E. F. Cook, L. Goldman, Performance of tests of significance based on stratification by a multivariate confounder score or by a propensity score. *J. Clin. Epidemiol.* **42**, 317–324 (1989).

10. M. S. Cepeda, R. Boston, J. T. Farrar, B. L. Strom, Comparison of logistic regression versus propensity score when the number of events is low and there are multiple confounders. *Am. J. Epidemiol.* **158**, 280–287 (2003).

11. O. Kuss, T. Legler, J. Börgermann, Treatments effects from randomized trials and propensity score analyses were similar in similar populations in an example from cardiac surgery. *J. Clin. Epidemiol.* **64**, 1076–1084 (2011).

12. M. Kuhn, M. Campillos, I. Letunic, L. J. Jensen, P. Bork, A side effect resource to capture phenotypic effects of drugs. *Mol. Syst. Biol.* **6**, 343 (2010).

13. M. Campillos, M. Kuhn, A.-C. Gavin, L. J. Jensen, P. Bork, Drug target identification using side-effect similarity. *Science* **321**, 263–266 (2008).

14. L. Yang, P. Agarwal, Systematic drug repositioning based on clinical side-effects. *PLoS One* **6**, e28025 (2011).

15. J. Hopstadius, G. N. Norén, A. Bate, I. R. Edwards, Stratification for spontaneous report databases. *Drug Saf.* **31**, 1145–1147 (2008).

16. J. Scheiber, J. L. Jenkins, S. C. Sukuru, A. Bender, D. Mikhailov, M. Milik, K. Azzaoui, S. Whitebread, J. Hamon, L. Urban, M. Glick, J. W. Davies, Mapping adverse drug reactions in chemical space. *J. Med. Chem.* **52**, 3103–3107 (2009).

17. J. Scheiber, B. Chen, M. Milik, S. C. Sukuru, A. Bender, D. Mikhailov, S. Whitebread, J. Hamon, K. Azzaoui, L. Urban, M. Glick, J. W. Davies, J. L. Jenkins, Gaining insight into off-target mediated effects of drug candidates with a comprehensive systems chemical biology analysis. *J. Chem. Inf. Model.* **49**, 308–317 (2009).

18. L. Xie, J. Li, L. Xie, P. E. Bourne, Drug discovery using chemical systems biology: Identification of the protein-ligand binding network to explain the side effects of CETP inhibitors. *PLoS Comput. Biol.* **5**, e1000387 (2009).

19. A. Cami, A. Arnold, S. Manzi, B. Reis, Predicting adverse drug events using pharmacological network models. *Sci. Transl. Med.* **3**, 114ra127 (2011).

20. S. N. Deftereos, C. Andronis, E. J. Friedla, A. Persidis, A. Persidis, Drug repurposing and adverse event prediction using high-throughput literature analysis. *Wiley Interdiscip. Rev. Syst. Biol. Med.* **3**, 323–334 (2011).

21. C. V. Fletcher, Drug interactions should be evaluated in patients. *Clin. Pharmacol. Ther.* **88**, 585–587 (2010).

22. M. Pirmohamed, Drug-drug interactions and adverse drug reactions: Separating the wheat from the chaff. *Wien. Klin. Wochenschr.* **122**, 62–64 (2010).

23. F. Montastruc, A. Sommet, E. Bondon-Guitton, G. Durrieu, E. Bui, H. Bagheri, M. Lapeyre-Mestre, L. Schmitt, J. L. Montastruc, The importance of drug–drug interactions as a cause of adverse drug reactions: A pharmacovigilance study of serotoninergic reuptake inhibitors in France. *Eur. J. Clin. Pharmacol.*, 10.1007/s00228-011-1156-7 (2011).

24. R. Leone, L. Magro, U. Moretti, P. Cutroneo, M. Moschini, D. Motola, M. Tuccori, A. Conforti, Identifying adverse drug reactions associated with drug-drug interactions: Data mining of a spontaneous reporting database in Italy. *Drug Saf.* **33**, 667–675 (2010).

25. G. N. Norén, R. Sundberg, A. Bate, I. R. Edwards, A statistical methodology for drug–drug interaction surveillance. *Stat. Med.* **27**, 3057–3070 (2008).

26. R. Harpaz, H. S. Chase, C. Friedman, Mining multi-item drug adverse effect associations in spontaneous reporting systems. *BMC Bioinformatics* **11** (Suppl. 9), S7 (2010).

27. L. Tari, S. Anwar, S. Liang, J. Cai, C. Baral, Discovering drug–drug interactions: A text-mining and reasoning approach based on properties of drug metabolism. *Bioinformatics* **26**, i547–i553 (2010).

28. B. Percha, Y. Garten, R. B. Altman, Discovery and explanation of drug-drug interactions via text mining. *Pac. Symp. Biocomput.* 410–421 (2012).

29. M. Takarabe, D. Shigemizu, M. Kotera, S. Goto, M. Kanehisa, Network-based analysis and characterization of adverse drug-drug interactions. *J. Chem. Inf. Model.* **51**, 2977–2985 (2011).

30. S.-F. Lin, K.-T. Xiao, Y.-T. Huang, C.-C. Chiu, V.-W. Soo, Analysis of adverse drug reactions using drug and drug target interactions and graph-based methods. *Artif. Intell. Med.* **48**, 161–166 (2010).

31. S. I. Berger, R. Iyengar, Network analyses in systems pharmacology. *Bioinformatics* **25**, 2466–2472 (2009).

32. M. P. Fichman, H. Vorherr, C. R. Kleeman, N. Telfer, Diuretic-induced hyponatremia. *Ann. Intern. Med.* **75**, 853–863 (1971).

33. E. Friedman, M. Shadel, H. Halkin, Z. Farfel, Thiazide-induced hyponatremia. Reproducibility by single dose rechallenge and an analysis of pathogenesis. *Ann. Intern. Med.* **110**, 24–30 (1989).

34. W. L. ten Holt, C. E. van Iperen, G. Schrijver, A. K. Bartelink, Severe hyponatremia during therapy with fluoxetine. *Arch. Intern. Med.* **156**, 681–682 (1996).

# Science Translational Medicine

## Data-Driven Prediction of Drug Effects and Interactions

Nicholas P. Tatonetti, Patrick P. Ye, Roxana Daneshjou, and Russ B. Altman

### Avoiding Adversity

For some disease-therapy pairs, Francis Bacon was right: "The remedy is worse than the disease." And when several drugs collide in an individual, the troubling effects can multiply. One goal of the developing field of pharmacogenomics is to make use of clinical data to predict adverse drug events so that future patients can be protected from the sometimes serious consequences. Now, Tatonetti *et al.* describe two new databases—one of drug effects and a second of drug-drug interaction side effects—that permit the identification of drug targets, prediction of drug indications, and discovery of drug-class interactions.

Adverse drug events aren't merely a nuisance; these toxic interactions can cause debilitating illness and death. The nature of clinical trials doesn't allow the detection of all serious side effects and drug interactions before approval of the therapy, by regulatory agencies, for use in patients. But these agencies along with pharmaceutical companies, hospitals, and other institutions collect adverse event reports after the drugs are in use in the clinic. When delineated in databases and coupled with the impressive computing power now available, these reports have the potential to permit characterization of drug effects at the population level. However, even with the recent move toward electronic health records, adverse event data often lack crucial information about co-prescribed medications, patient demographics and medical histories, and the reasons that a given drug was prescribed in the first place. One can easily see how that lack of such information thwarts the ability to obtain meaningful analyses of drug side effects and interactions.

To address this problem of omission and improve the ability to analyze drug effects, Tatonetti *et al.* use an adaptive data-driven approach to correct for the lack of such information—the so-called unknown "covariates." Using this information, the authors developed two comprehensive databases—one of drug effects (Offsides) and another of drug-drug interaction side effects (Twosides)—and then used their new databases to pinpoint drug targets and discover drug-class interactions. Finally, the authors validated 47 of the drug-class interactions in an independent analysis of patient electronic health records. When prescribed together, widely used antidepressant drugs (selective serotonin reuptake inhibitors) and thiazide diuretics were associated with an increase in the incidence of prolonged QT, which indicates a delayed repolarization of the heart after a heartbeat. Prolonged QT can increase a patient's risk of palpitations, fainting, and even sudden death resulting from ventricular fibrillation. Better than tarot cards or crystal balls, the authors show that intricate analyses of observational clinical data can improve physicians' ability to predict the future—at least with respect to as yet uncharacterized adverse drug effects and interactions.

**View the article online**
https://www.science.org/doi/10.1126/scitranslmed.3003377
**Permissions**
https://www.science.org/help/reprints-and-permissions