www.sciencetranslationalmedicine.org/cgi/content/full/4/125/125ra31/DC1

Supplementary Materials for

Data-Driven Prediction of Drug Effects and Interactions

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

*To whom correspondence should be addressed. E-mail: russ.altman@stanford.edu

Published 14 March 2012, *Sci. Transl. Med.* **4**, 125ra31 (2012)
DOI: 10.1126/scitranslmed.3003377

This PDF file includes:

Other Supplementary Material for this manuscript includes the following:

Tab delimited database file: offsides.tsv
(available at: http://helix-web.stanford.edu/download/tatonetti/3003377s-offsides.zip)
Tab delimited database file: twosides.tsv
(available at: http://helix-web.stanford.edu/download/tatonetti/3003377s-twosides.zip)

A. Materials and Methods

1. Supplemental Note: Signal detection for common adverse events.

Drug-drug interactions (DDIs) are a growing public health concern account for up to 40% of adverse drug events(S*1*). DDIs can occur any time a patient is taking more than one drug concurrently and result from shared metabolism pathways or interacting mechanisms of action. Some DDIs can be predicted through careful analysis of these pathways, but many are idiosyncratic and difficult to predict. Drug surveillance databases, such as the Food and Drug Administration's Adverse Event Reporting System (AERS), represent an opportunity to identify DDIs from post-marketing data. Signal detection using disproportionally analysis is the prevailing technique for analyzing spontaneous reporting data(S*2*). The goals of signal detection algorithms are to flag drug-event associations that are reported a higher than expected rates for follow up analysis(S*2*). They focus on identifying rare events that will only manifest once the drug is exposed to a large and diverse population of patients. The common effects of drugs are primarily captured by the pre-marketing clinical trials. In order to detect potentially dangerous adverse drug effects early, these algorithms flag associations even when there are only a few reports of supporting evidence. Extensions on this approach for detecting DDIs (or more generally multi-item associations) have previously been presented(S*3*, S*4*). These statistics, however, have a bias towards associations with relatively few supporting reports. This is a desirable trait of the statistic when the goal of the algorithm is to find rare effects early. However, unlike for single-drug treatment, there are no clinical trials of drug-drug interactions. For the most part, clinical trials exclude patients taking multiple drugs in order to mitigate confounding effects. Because of this, even common effects of drug-drug interactions may be unknown. We propose an alternative signal detection statistic based on differences of report rates (as opposed to ratios) for the detection of common effects with specific application to drug-drug interactions.

We present a signal detection statistic based on the differences of reporting rates. Given an exposed set of a reports and a non-exposed set of control reports we use sub-sampling to estimate the mean and variance of the reporting rate. To do this we sample, without replacement, 10% of the reports and compute the reporting frequency. We then repeat this sampling 1,000 times, with replacement, for both the exposed set of reports and the non-exposed. From these samples we can compute the average reporting rate as well as the variance for both groups. By the central limit theorem these distributions of means will be normally distributed. We then use a Student’s T test to evaluate the difference between the exposed and non-exposed groups. We found that the T statistic has two important properties. First, when evaluated against two lists of drug interactions, (S1) a list of critical drug interactions maintained by the VA(S*5*) and (S2) adverse event reports for pairs of drugs in Canada’s spontaneous reporting system, we found that the difference statistic performed significantly better than the a ratio statistic recently proposed for DDis(S*3*) (fig. S9). Second, we found that unlike the disproportionality statistics, the difference statistic increased as the number of reports supporting the association increased (fig. S10). We believe that these features make the difference statistic more suitable in two situations: (S1) identifying common effects

of drug interactions in spontaneous reporting systems, and (S2) characterizing the overall side-effect profile of drugs as reported in spontaneous reporting systems.

2. Drug-effect databases are predictive of known class-wide effects

We used the high confidence associations from OFFSIDES to establish class-wide adverse event associations. We identified 67 significant interactions between ATC drug classes and top level adverse event (COSTART) terms (fig. S8). Twenty-two of these associations are significant after multiple hypothesis correction (table s5). These results recapitulate known effects of drugs. For example, there is a significant positive association between antiparasitics/insecticides and nervous systems adverse events. Diethylcarbamazine, used in the treatment of toxocariasis and other parasitic diseases, has been associated with cases of paresthesia(S*6*). More famously, organophosphates, which are used as insecticides, mediate neurological adverse events by inhibiting acetylcholinesterase, and have also been implicated as an environmental risk for dementia and Alzheimer's disease(S*7*). Additionally, anti-neoplastic and immune-modulating drugs were significantly associated with maternal-fetal disorders. There are documented examples of this association: for instance, Mycophenolate mofetil (MMF), a popular immunosuppressant agent used in transplant patients, is associated with teratogenesis in humans(S*8*-S*10*). Methotrexate, which is both an immune modulator and an anti-neoplastic drug, has also been well-documented to cause birth defects(S*11*). Another corroborated connection is that between genito-urinary/sex hormones and the reticuloendothelial system. Hormonal contraceptive therapy has been shown to effect immune modulation, and research has shown that hormonal contraceptive therapy may be tied to increased risk of sexually transmitted infections (S*12*, S*13*).

3. Case studies in the statistical correction of uncharacterized bias

To provide intuition for the SCRUB algorithm, we examined seven case studies that exemplify the errors found in the AERS database. In each case, we identified a potential source of synthetic associations (either a concomitant medication or drug indication) and visualized how SCRUB corrects for these associations while retaining known true signals. We use the drug's package inserts (SIDER) as the source of known true associations and it is important to note that not all events caused by drugs are listed on their inserts nor are all events listed caused by the drug.

Abacavir is known to cause severe rashes. Drugs that are commonly co-reported with abacavir (and thus presumed to be often co-prescribed) are commonly synthetically associated with severe rash. Fig. S1A shows the un-corrected association score, "Original PRR" as well as the corrected association score "Corrected PRR." We found that SCRUB dampens the association scores of synthetic associations. Fig. S1B shows that at the same time that SCRUB is dampening the signal of synthetic associations positives, "Synthetic," it is not dampening the signal of other drugs, "Known True," which are known to also cause severe rash, according to the package insert. This demonstrates the selectivity of the algorithm for dampening only those signals which are contributing noise to the association analysis. The astute reader will note that some of the drugs listed to have synthetic associations (fig. S1A) actually have rash listed on their package inserts. However, it is well known that not all the associations listed on the package insert are actually caused by the drug (and vice versa). 5 of the 10 drugs list

rash, 3 of which include 0 in their confidence interval for the incidence (according to the clinical trials(S*14*)). Of the remaining 2, our algorithm retains the signal for amprenavir (PRR = 3.4). If we consider those two to be true associations and the remaining synthetic then the algorithm reduced the false positive rate from 80% (8/10) to 10% (1/10), while the true positive rate increased from 20% (2/10) to 50% (1/2).

We repeated this analysis for isoniazid, rofecoxib, and pergolide, drugs that are associated with hepatic failure, myocardial infarction, and heart valve damage, respectively. In each case, SCRUB significantly dampens the signals of synthetic associations while maintaining the true signals (fig. S1-S4).

Drugs that treat arrhythmias are often synthetically associated with arrhythmia-related adverse events. This is due to the higher incidence of arrhythmias in patients with a history of arrhythmia. However, this is a special case as some anti-arrhythmia agents are known to have pro-arrhythmic effects. We found that application of SCRUB dampens the signal of the synthetic signals (fig. S5A), while maintaining the signal of the known true effects (fig. S5B). In this example, three drugs remain above the traditional score threshold of 2, propafenone, sotalol, and dofetilide. In fact, each of these drugs is known to have a significant proarrhythmic effects(S*15*-S*18*). In addition to retaining these signals SCRUB also significantly retained the signals of other drugs (co-reported anti-arrhythmic or not) that are known to cause arrhythmias. This same trend (dampening of synthetic signals and retention of true signals) is true for the other two examples as well: 1) cholesterol lowering drugs and hypercholestermia, and 2) drugs reported with diabetes and their association with hyperglycemia (fig. S6-S7).

Case studies statistical analysis
To provide intuition for the methodology we manually selected four drugs and three indications that likely sources of synthetic signals. We chose drugs which are infamously associated with severe side effects. These drugs-side effect pairs were refecoxib and heart attacks, isoniazid and hepatic failure, pergolide and heart valve damage, and abacavir and rash. Other drugs that are concomitantly taken with these four drugs are at risk of being synthetically associated with these side effects simply through correlation in reporting. We then used the drug's package inserts to define true associations between drugs and the four adverse events. If the method works as described then the signal for synthetic associations will be dampened to a greater degree than that of known true signals. We then test for a difference in the log ratio of the original to corrected association scores. We use the log ratio so that the data conforms to the assumptions of the statistical test (Student's t-test), and is uncorrelated with the magnitude of the association score. We repeat this analysis for three indications. Drugs that are significantly reported with these indications are more likely to be synthetically associated with the indications effects. Again we define a set of known true associations from the drug's package inserts and repeat the statistical analysis.

4. Identification of drug-drug interactions and EMR analysis of putative class-class interactions

Statistical analysis of class-class interactions from drug-drug-event associations in TWOSIDES

We identified class-class interactions for follow-up validation analysis by linear modeling of the adverse event reporting frequencies. We mapped each adverse event to one of 19 high level UMLS identifiers. For each of these adverse event categories we performed the following statistical analysis to identify putative class-class interactions. For each pair of level 4 ATC drug classes, we first constructed a table where the rows were drug-drug-event triplets that were found to have significant associations by SCRUB (i.e. associations reported in the TWOSIDES database). These triplets were restricted to those events that were contained in the high level adverse event category of interest. The model contained two independent variables: (1) an indicator variable signifying membership of the first drug class of the pair, (2) an indicator variable signifying membership of the second drug class of the pair, and the dependent variable: the drug-drug-event reporting frequency observed in the AERS. We modeled the reporting frequency linearly as a function of the two indicator variables including an interaction term. We identified any statistical model that included a significant interaction term, after multiple hypothesis correction, for follow up analysis using the electronic medical records.

IRB approval was obtained to perform the EMR analysis.

Acute effects analysis using baseline and treatment laboratory values

We performed a retrospective analysis of the electronic medical records to corroborate the putative class-class interactions. We used an analysis of covariance (ANCOVA) model to compare laboratory values taken shortly before a prescription was ordered (baseline) to those taken shortly after (treatment). This approach was used previously to investigate a drug interaction affecting glucose homeostasis (S*19*). A list of the labs we used for each adverse event category is given in table s6. In this approach a common clinical lab is chosen as a marker for the predicted adverse event. For example, for hyperglycemia we may choose measures of blood glucose levels. For each putative interaction we manually identified labs that may be used as markers of the predicted event (table s6). The next step is to identify three cohorts of patients. The first cohort consists of patients that have a prescription for a drug from first class and at least one laboratory result in the preceding 36 days before treatment (baseline) and at least one in the following 36 days (treatment). The second cohort mirrors the first except that the prescription is for a drug from the other class. The third cohort of patients have been prescribed both drugs during the same visit. In this case the start of treatment is the day the second drug is prescribed. All patients in all cohorts must have at least one baseline and one treatment result measured. For the individual drug cohorts, we use the first prescription recorded in the database as the start date of treatment. Any prescriptions for the same drug after that first date are ignored. Note that since Stanford is a tertiary hospital system there is the possibility that this was not the actual treatment start date and the first prescription was made by a primary care physician.

We used analysis of variance with covariates (ANCOVA) to model the change from baseline to treatment as a function of the treatment group (combined, drug-class1, or drug-class2) including a common set of covariates (age, sex, and race). The ANCOVA can establish that the cohorts of patients (1, 2, and 3) have different outcomes, however, this result does not necessarily validate the presence of an interaction. To filter for those models that may indicate the presence of an interaction, we performed Tukey's post-hoc testing and filtered for those interactions where co-prescription exposure had a significant effect when compared directly the single exposure cohorts. In addition, we also required the co-prescription cohort effect when compared to the single prescriptions group to be in the same direction (i.e., the co-prescription effect could not be "between" the effects of the single treatment groups). We corrected the derived *P* values using Bonferonni multiple hypothesis correction.

Long-term effects through Cox proportional hazards analysis of abnormal lab values

The putative class-class interactions may have longer term effects that the acute analysis method may not be able to identify. We used Cox proportional hazards regression, with covariates (age, sex, and race), to identify these effects in patients who were co-prescribed pairs of drugs where each drug of the pair was a member of one of the drug classes in the putative class-class interaction. As we did in the acute analysis we associated common laboratory results to adverse event categories and predicted that true class-class interactions would cause a high incidence of abnormal laboratory results over time (table s7 lists the normal ranges used for the laboratory values). The patient cohorts were defined in the same way as the acute analysis (previous section), however, without the restriction of having a measured lab value in the patients clinical record. As in the acute analysis we filtered for those interactions where the treatment group is significant, after multiple hypothesis correction (Bonferonni), and the risk associated with the co-prescription of the drugs was greater than the risk of the drugs separately. As a final pruning step we filtered for only those class-class interactions that validated in both the acute and long-term analyses (table 1). We then computed an empirical estimate of the significance of the analytical pipeline by randomization.

5. Generalization of the method for multi-item associations (drug-drug interactions)

We repeated the analysis used for single drug-event associations for drug-drug-event associations. Due to the increased computational complexity that results from the pair-wise analysis we adapted the method slightly for drug-drug interactions. Rather than fit a logistic model (PSM) for each pair of drugs we identified those concomitant medications and indications that were significantly co-reported with either drug of the pair. We ordered these covariates by their reporting correlation. We then iterated through this list of covariates. We extracted the reports for each covariate, in turn, and added them into our control cohort of non-exposed reports. We repeated this until the number of desired control reports was obtained (10X the number of exposed reports). We evaluated these putative drug interactions against a list of critical or significant interactions maintained by the VA(S*5*).

B. Supplemental Figures

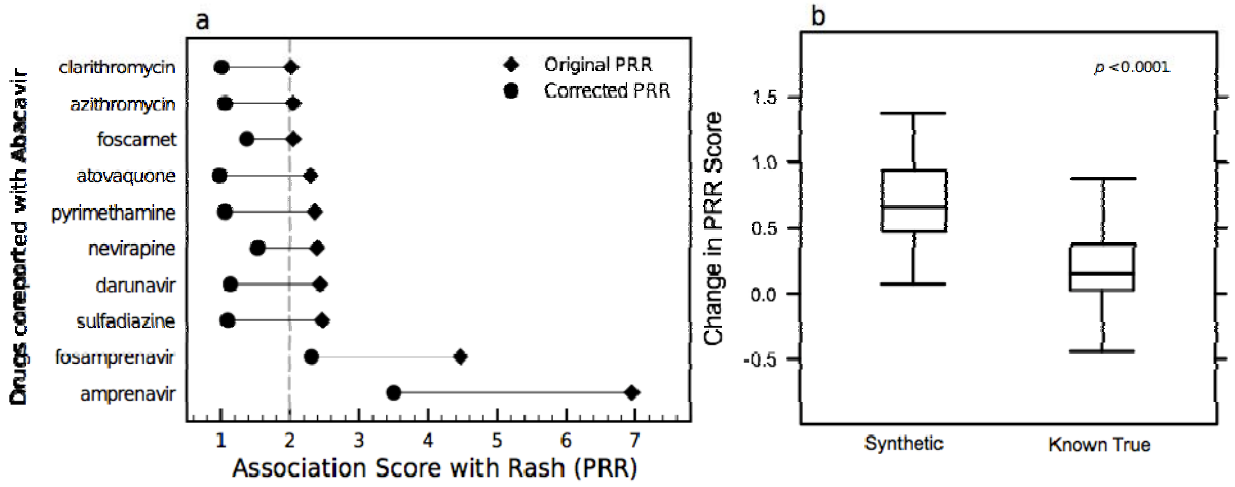

Fig. S1. Case study of the correction of bias. Drugs co-reported with abacavir commonly have high association scores with severe rash (one of abacavir's known side effects). The SCRUB method dampens the original scores (shown as diamonds) and corrects for this co-reporting bias to produce new association scores (shown as circles). SCRUB does not dampen signal overall, but preferentially dampens the likely false positive associations (those for drugs that are commonly co-reported with abacavir) while maintaining known true associations (those for drugs known to also cause severe rash). The dashed line is the traditional significance cutoff.

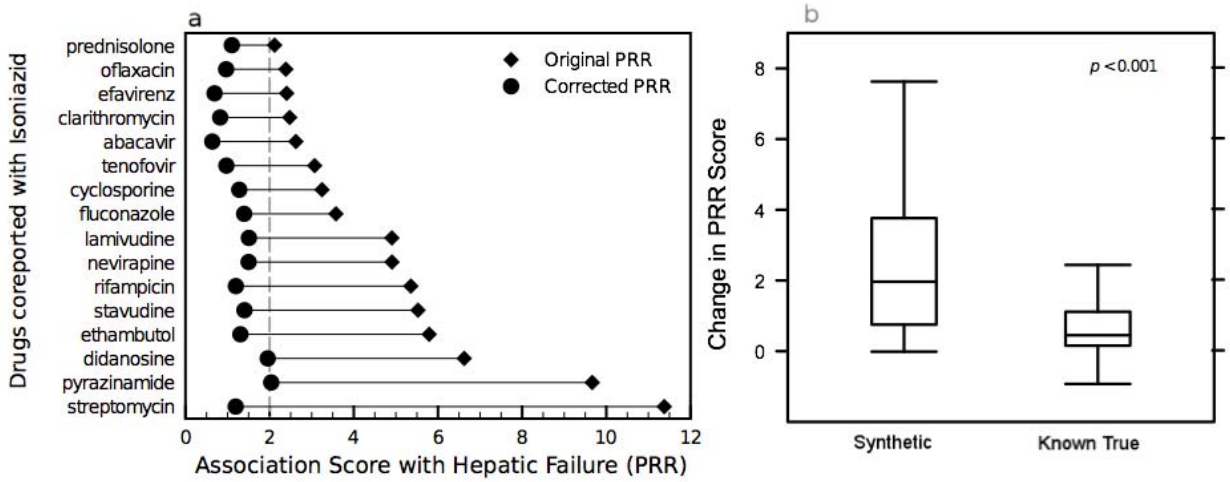

Fig. S2. Case study of the correction of bias. Drugs co-reported with isoniazid are more likely than other drugs to be falsely associated with hepatic failure (one of isoniazid's known effects). SCRUB preferentially dampens the signals of these likely false positive associations (a) while maintaining the signal from known true associations (those for drugs known to also cause hepatic failure). The dashed line is the traditional significance cutoff.

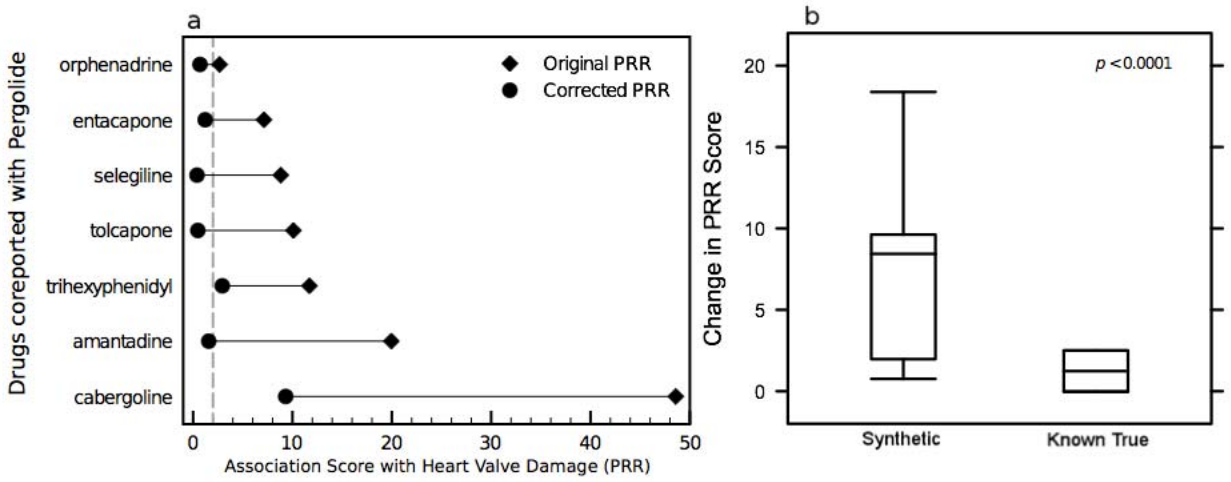

Fig. S3. Case study of the correction of bias. Drugs co-reported with pergolide are more likely than other drugs to be falsely associated with heart valve damage (an effect of pergolide). SCRUB preferentially dampens the signals of these likely false positive associations (a) while maintaining the signal of known associations. The dashed line is the traditional significance cutoff.

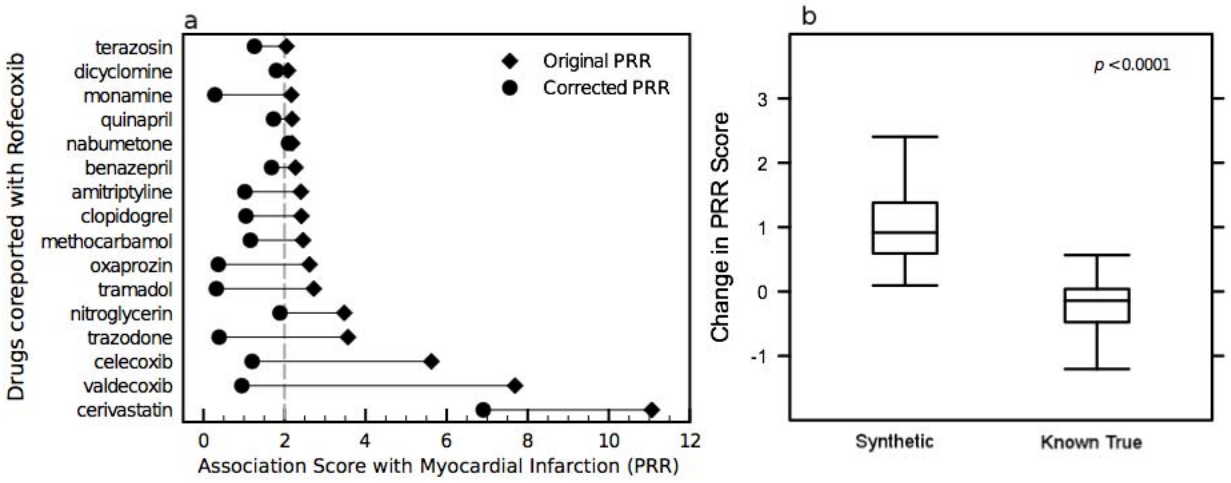

Fig. S4. Case study of the correction of bias. Drugs co-reported with rofecoxib are more likely than other drugs to be falsely associated with myocardial infarction (an effect of rofecoxib). SCRUB preferentially dampens the signals of these likely false positive associations (a) while maintaining the signal of known associations. The dashed line is the traditional significance cutoff.

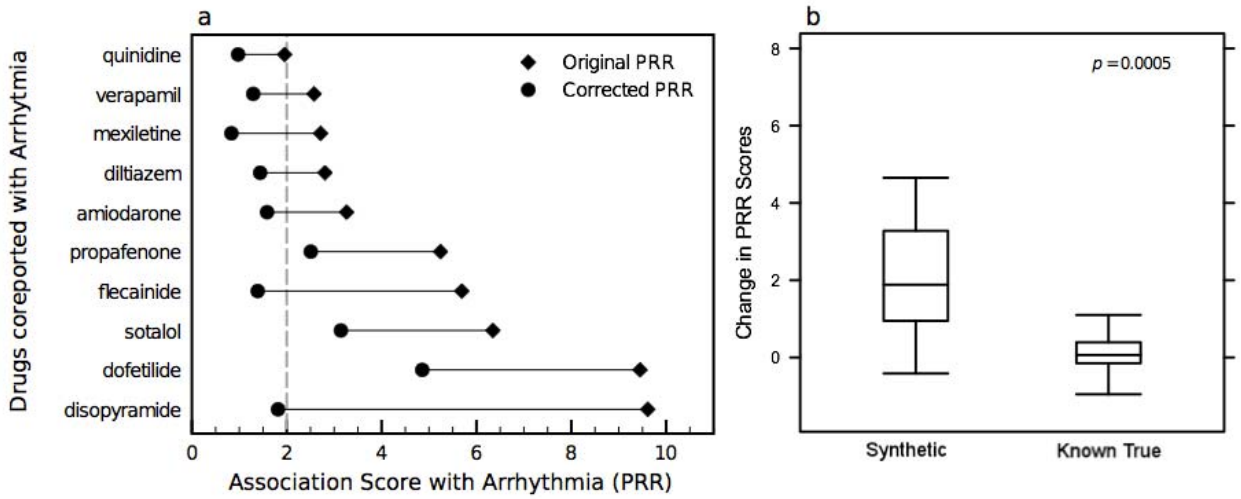

Fig. S5. Case study of the correction of indication bias. Drugs that are commonly given to patients with arrhythmias (i.e., anti-arrhythmics) are more likely than other drugs to be associated with proarrhythmic effects falsely. SCRUB corrects for this noise by dampening the signal of those likely false positive associations. The method does so while preserving the signal of known true positives (b). In fact, the three drugs with scores that remain above the significance threshold are all known to have proarrhythmic effects.

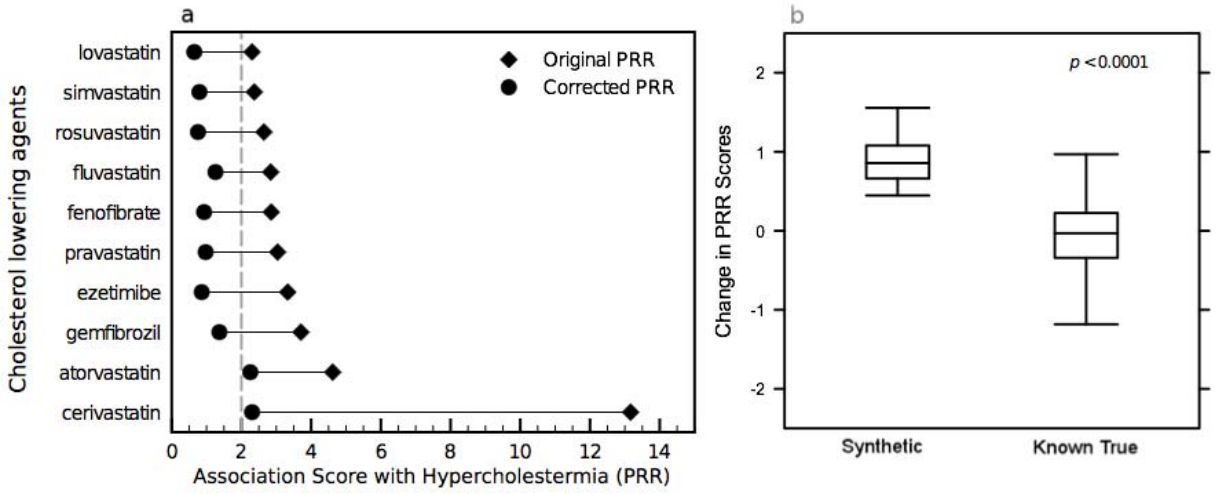

Fig. S6. Cholesterol lowering agents are commonly (and falsely in most cases) associated with hypercholesterolemia. SCRUB corrects for this indication-based bias by dampening the signal of these drugs. The method is specific so that it dampens the false positive signals while maintaining known true positive associations (b).

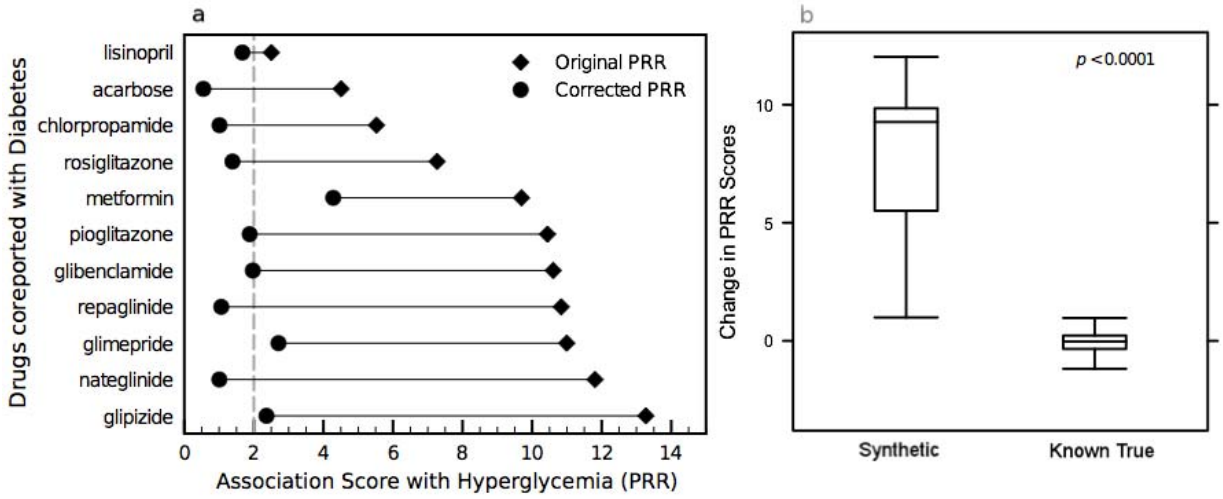

Fig. S7. Drugs often given to diabetics are commonly associated with hyperglycemia. Many of these drugs actually lower blood glucose levels and so this association is false and nonsensical. SCRUB corrects for this bias (a) and does so selectively so that known true associations are not lost (b).

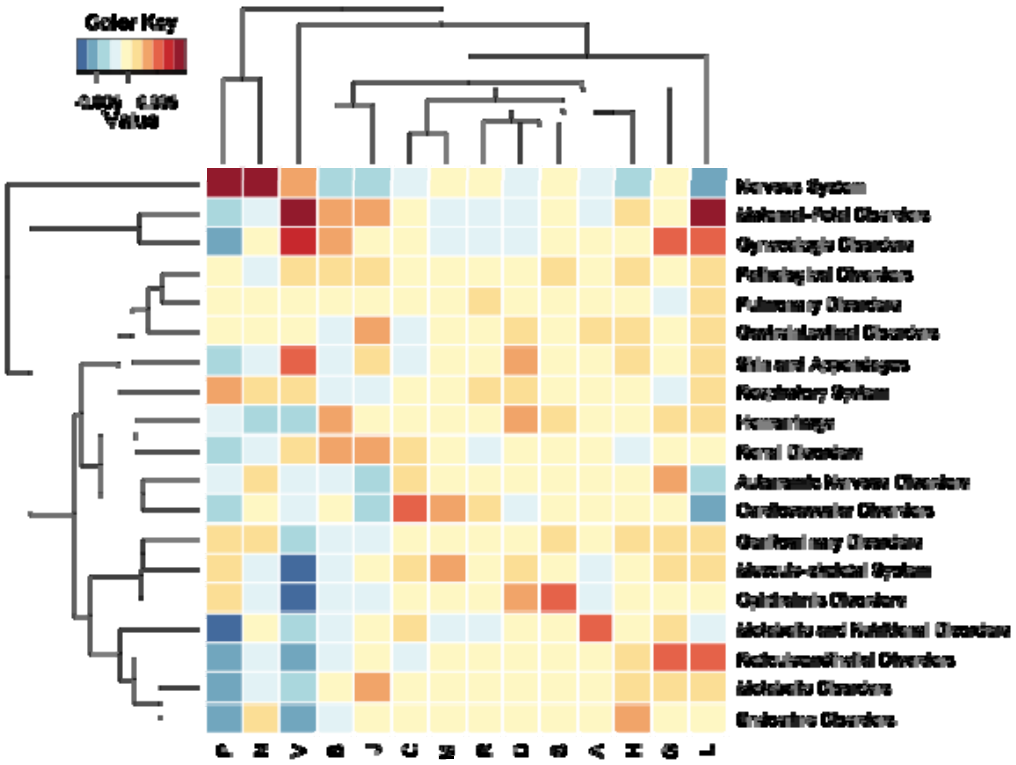

Fig. S8. Heat map of the interaction coefficients between adverse event categories and drug classes. Positive associations are shown as red hues and negative associations are shown as blue hues. (A) Alimentary tract and metabolism, (B) Blood and blood forming organs, (C) Cardiovascular system, (D) Dermatologicals, (G) Genito-urinary system and sex hormones, (J) Antiinfectives for systemic use, (L) Antineoplastic and immunomodulating agents, (M) Musculo-skeletal system, (N) Nervous system, (P) Antiparasitic products, (S) Sensory organs, (V) Various.

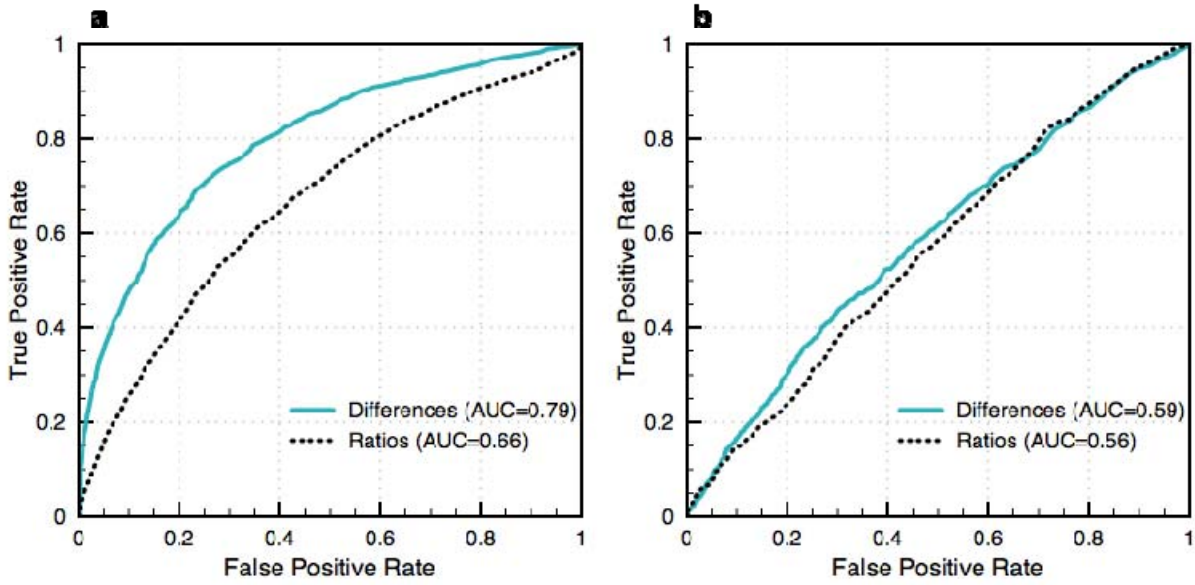

Fig. S9. Receiver operating characteristic curves comparing the performance of the difference based signal detection statistic versus the ratio based. (a) Evaluation against a set of adverse effects reported with pairs of drugs to the Canadian MedEffect database. (b) Evaluation against a list of critical drug-drug interactions maintained by the Veterans Affairs Medical System. In both cases the difference method significantly outperforms the ratio based method.

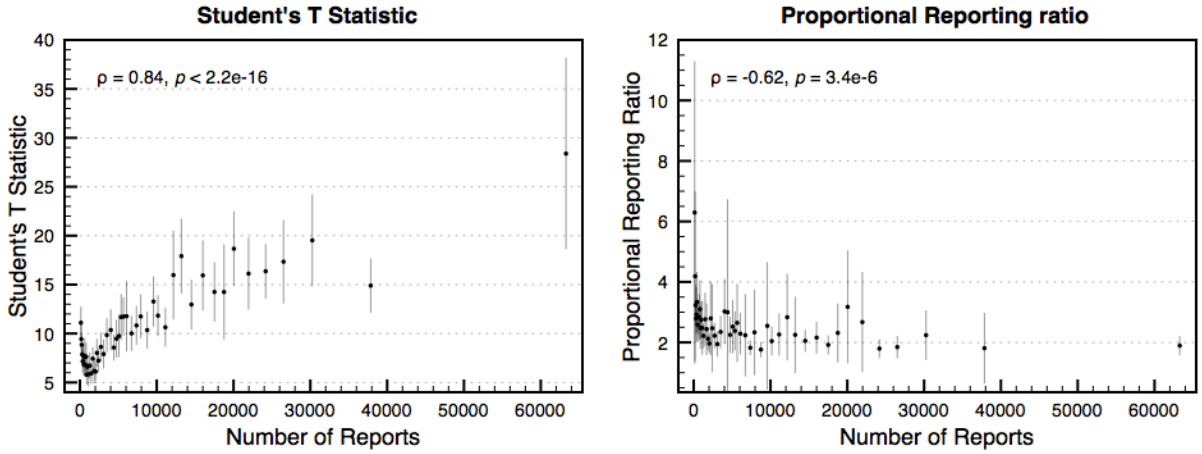

Fig. S10. The difference-based association statistic increases in magnitude as the number of reports supporting the association increases (binned mean ± SEM) (left). The same is not true for the ratio-based statistic, which is biased towards drugs (and events) that have relatively few reports (binned mean ± SEM) (right). This limits the ratio-based method's ability to identify rare events for common drugs or common events for rare drugs (or drug combinations). The difference-based method does not suffer from that bias, and therefore, is more appropriate for use in these scenarios.

C. Supplementary Tables

Table S1. Adverse events primarily reported with males.

UMLS ID	Adverse Event	Proportion Male	Number of Males Reported	Number of Females Reported
C0403673	retrograde ejaculation	1.0000	104	0
C0194790	prostatic operation	1.0000	144	0
C1282799	Penile swelling	1.0000	107	0
C0004509	azoospermia	0.9936	157	1
C0010417	cryptorchidism	0.9909	220	2
C0029191	orchitis	0.9845	194	3
C0855411	coagulation factor viii level decreased	0.9521	146	7
C1688637	floppy iris syndrome	0.9431	123	7
C0853117	blood testosterone decreased	0.9011	354	35
C0001175	acquired immune deficiency syndrome	0.8000	130	26
C1096726	hepatitis b virus	0.7967	123	25
C0274869	metal poisoning	0.7944	287	59
C0276535	kaposi's sarcoma	0.7940	398	82
C0948060	iridocele	0.7767	103	23
C0312420	hypersexuality	0.7014	144	43
C0154446	autism	0.6749	203	66
C0748052	erythrodermic psoriasis	0.6535	101	35
C1142166	brugada syndrome	0.6476	105	37
C0085315	cerebral toxoplasmosis	0.6299	127	47
C1096691	bk virus infection	0.5686	102	44
C0035012	reiter's syndrome	0.5660	106	46
C0877192	lipodystrophy acquired	0.5365	274	127
C0236663	alcohol withdrawal syndrome	0.5350	157	73
C0341702	fanconi syndrome	0.5079	254	125
C1096260	enterococcal sepsis	0.5079	126	62
C1096272	vanishing bile duct syndrome	0.4950	101	51
C0376620	pouchitis	0.4895	190	97
C0877072	cytomegalovirus chorioretinitis	0.4650	157	84
C0877020	anti-erythropoietin antibody positive	0.4425	226	126
C0919710	cytomegalovirus antigen	0.4363	204	115
C1112375	cataract operation complication	0.4352	108	61
C0856437	adenovirus infection	0.4050	121	72
C0043541	mucormycosis	0.3846	182	112
C0036508	dermatitis seborrheic	0.3775	204	127
C1739372	acute graft versus host disease	0.3537	311	201
C0745528	klebsiella sepsis	0.3303	109	73
C0035232	diaphragmatic paralysis	0.3302	106	71
C1868882	human herpesvirus 6 infection	0.3235	102	69
C1167782	lung hyperinflation	0.3217	115	78

Table S2. Age-related adverse events.

UMLS ID	Adverse Event	Average Age of Patient Reported with Event	Absolute Difference From Database Average
C0000832	premature separation of placenta	30.840491	22.689509
C0007860	cervicitis	36.171429	17.358571
C0008313	Cholangitis sclerosing	40.186033	13.343967
C0011551	depersonalisation	36.941368	16.588632
C0014179	Endometritis	29.565574	23.964426
C0014583	episcleritis	43.257426	10.272574
C0016436	folliculitis	41.267051	12.262949
C0017638	glioma	24.916667	28.613333
C0021099	impetigo	40.393939	13.136061
C0042900	skin depigmentation	43.088435	10.441565
C0042998	vulval oedema	38.041420	15.488580
C0085166	bacterial vaginitis	32.560510	20.969490
C0085315	cerebral toxoplasmosis	42.462963	11.067037
C0152459	Skin Striae	30.385827	23.144173
C0155107	corneal deposits	65.025316	11.495316
C0156533	abortion incomplete	27.331019	26.198981
C0156892	twin pregnancy	32.258333	21.271667
C0233681	delusion of grandeur	38.888889	14.641111
C0236650	vascular dementia	71.160714	17.630714
C0238425	sickle cell anaemia with crisis	30.957983	22.572017
C0259817	xerosis	33.140351	20.389649
C0267465	intestinal stenosis	42.174312	11.355688
C0332687	Burns Second Degree	42.643312	10.886688
C0349231	phobia	38.514286	15.015714
C0376620	pouchitis	23.408284	30.121716
C0520561	small intestinal haemorrhage	65.140693	11.610693
C0542518	kidney enlargement	42.233333	11.296667
C0547058	pseudopolyp	28.703488	24.826512
C0566948	vaginal swelling	32.708333	20.821667
C0678201	ileitis	32.415274	21.114726
C0852792	blood cortisol decreased	39.654545	13.875455
C0854062	blood fibrinogen	43.216049	10.313951
C0854406	glomerulonephritis proliferative	40.325581	13.204419
C0856437	adenovirus infection	29.271127	24.258873
C0867389	chronic graft versus host disease	31.720126	21.809874
C0877192	lipodystrophy acquired	42.350722	11.179278
C0919718	mitral valve calcification	64.094697	10.564697
C0919746	engraftment syndrome	38.415842	15.114158
C0948441	venoocclusive disease	29.994487	23.535513

Table S2. Age-related adverse events.

UMLS ID	Adverse Event	Average Age of Patient Reported with Event	Absolute Difference From Database Average
C1096272	vanishing bile duct syndrome	30.462857	23.067143
C1096691	bk virus infection	37.158333	16.371667
C1112375	cataract operation complication	74.588235	21.058235
C1141896	congenital mitochondrial cytopathy	38.940678	14.589322
C1167782	lung hyperinflation	40.572519	12.957481
C1328409	propofol infusion syndrome	28.883495	24.646505
C1739372	acute graft versus host disease	35.457837	18.072163
C1868882	human herpesvirus 6 infection	43.120805	10.409195
C2363744	epstein-barr virus associated lymphoproliferative disorder	41.669421	11.860579

Table S3. Effect of commonly co-prescribed drugs on association with thiazides and SSRIs.

Co-Rx Drug	Combination vs Drug P Value (ANOVA)	Drug Class 1 vs Combo P Value	Drug Class 2 vs Combo P Value
acetaminophen	1.66E-06	9.82E-08	5.47E-05
ondansetron	8.93E-07	3.84E-08	2.00E-05
aspirin	2.16E-07	8.44E-09	1.76E-05
midazolam	7.51E-08	3.40E-09	6.18E-09
chlorothiazide	2.68E-06	1.05E-07	5.64E-06
thyroxine	1.63E-07	6.09E-09	2.62E-08
metformin	9.82E-08	3.63E-09	1.24E-08
omeprazole	1.15E-07	4.58E-09	1.38E-08
trazodone	4.28E-06	2.61E-07	9.44E-07
metoprolol	4.12E-07	1.09E-08	5.06E-07
morphine	3.08E-07	9.69E-09	1.11E-07
lisinopril	6.68E-08	3.01E-09	6.85E-09
lidocaine	1.52E-06	5.30E-08	1.75E-06
KCl	2.76E-07	1.01E-08	5.49E-08
amlodipine	1.43E-07	3.93E-09	4.54E-08
galactose	3.81E-06	1.87E-07	2.39E-05
salbutamol	1.16E-06	4.94E-08	3.42E-07
lorazepam	1.78E-06	7.18E-08	8.79E-07
fentanyl	4.24E-07	1.76E-08	7.73E-08
warfarin	1.24E-06	5.19E-08	1.29E-05
naloxone	1.48E-08	1.72E-09	3.57E-10
zolpidem	3.16E-07	1.00E-08	1.11E-07
hydralazine	1.86E-06	6.63E-08	2.79E-06
famotidine	1.32E-07	5.23E-09	1.61E-08
vancomycin	6.12E-07	2.36E-08	1.57E-07
pantoprazole	2.76E-06	1.53E-07	4.16E-05
ciprofloxacin	2.53E-07	9.39E-09	4.52E-08
simvastatin	1.16E-07	4.05E-09	1.86E-08
calcium	1.92E-07	5.74E-09	5.26E-08
metoclopramide	3.90E-08	2.48E-09	2.02E-09
hydromorphone	3.59E-08	2.22E-09	1.74E-09
diphenhydramine	1.91E-07	5.40E-09	6.27E-08
heparin	1.80E-06	7.36E-08	1.20E-05
promethazine	3.06E-08	1.95E-09	1.35E-09
bisacodyl	1.14E-07	4.62E-09	1.37E-08
nitroglycerin	9.94E-07	3.15E-08	2.38E-06
furosemide	2.23E-05	1.18E-05	0.0063
gabapentin	5.61E-07	2.35E-08	1.13E-07

Table S4. Covariates in cox-regression model of thiazides and SSRIs.

Covariate	Coefficient	z	Pr(>|z|)
age	-0.03398	-19.196	< 2e-16
sex	0.01448	0.258	0.796
race	-0.68845	-13.957	< 2e-16
co-rx	1.06606	14.809	< 2e-16

Table S5. Drug class-adverse event category associations in OFFSIDES				
ATC Level 1	**Drug Class Name**	**Adverse Event Category**	**P Value**	**Interaction Coefficient**
A	Alimentary tract and metabolism	Metabolic and Nutritional Disorders	3.193E-07	0.004508759438308
B	Blood and blood forming organs	Nervous System	8.302E-05	-0.004192781375536 88
C	Cardiovascular system	Cardiovascular Disorders	3.868E-25	0.005242108 59892963
D	Dermatologicals	Nervous System	5.776E-06	-0.002957192 91060774
D	Dermatologicals	Skin and Appendages	1.719E-05	0.003250374463832 63
G	Genito-urinary system and sex hormones	Gynecologic Disorders	2.220E-05	0.005435049 34699865
G	Genito-urinary system and sex hormones	Reticuloendthelial Disorders	1.484E-04	0.005366706 27470556
J	Antiinfectives for systemic use	Cardiovascular Disorders	6.834E-09	-0.003312559315378 26
J	Antiinfectives for systemic use	Gastrointestinal Disorders	3.168E-06	0.002508055 16728877
J	Antiinfectives for systemic use	Nervous System	4.768E-09	-0.003638849 28372045
L	Antineoplastic and immunomodulating agents	Cardiovascular Disorders	6.649E-12	-0.004824671 41299813
L	Antineoplastic and immunomodulating agents	Gynecologic Disorders	4.333E-05	0.004772704267665 32
L	Antineoplastic and immunomodulating agents	Maternal-Fetal Disorders	3.960E-06	0.008250389 72731614
L	Antineoplastic and immunomodulating agents	Nervous System	2.094E-12	-0.005340686 59437577
L	Antineoplastic and immunomodulating agents	Reticuloendthelial Disorders	1.767E-06	0.004727600 9840146
M	Musculo-skeletal system	Cardiovascular Disorders	2.445E-08	0.004031660 89647497
N	Nervous system	Hemorrhage	3.374E-06	-0.003498070 21243736
N	Nervous system	Nervous System	1.403E-81	0.009172877952327
N	Nervous system	Pathological Disorders	5.105E-10	-0.002683197 15171321

Table S5. Drug class-adverse event category associations in OFFSIDES				
ATC Level 1	**Drug Class Name**	**Adverse Event Category**	**P Value**	**Interaction Coefficient**
N	Nervous system	Skin and Appendages	1.446E-04	-0.002790563 66160221
P	Antiparasitic products	Nervous System	1.322E-08	0.009538116 41730558
S	Sensory organs	Ophthalmic Disorders	6.396E-13	0.004855924 11287241

Table S6. Labs used as markers for adverse event categories.

Laboratory Code	Adverse Event Category
ALB	Metabolic and Nutritional Disorders
ALB	Gastrointestinal Disorders
ALKP	Gastrointestinal Disorders
ALKP	Musculo-skeletal System
ALT	Gastrointestinal Disorders
AST	Gastrointestinal Disorders
BASO	Reticuloendthelial Disorders
BASOAB	Reticuloendthelial Disorders
BD	Pulmonary Disorders
BE	Pulmonary Disorders
BUN	Metabolic Disorders
BUN	Musculo-skeletal System
BUN	Renal Disorders
CA	Endocrine Disorders
CA	Cardiovascular Disorders
CA	Metabolic Disorders
CAION	Pulmonary Disorders
CHOL	Metabolic Disorders
CHOL	Metabolic and Nutritional Disorders
CHOL	Endocrine Disorders
CHOLHDL	Metabolic Disorders
CHOLHDL	Metabolic and Nutritional Disorders
CL	Metabolic Disorders
CL	Renal Disorders
CO_2	Metabolic Disorders
CO_2	Pulmonary Disorders
CR	Metabolic Disorders
CR	Metabolic and Nutritional Disorders
CR	Genitourinary Disorders
CR	Renal Disorders
CULT	Reticuloendthelial Disorders
DI	Hemorrhage
ECGIMPRESS	Cardiovascular Disorders
ECGIMPRESS	Metabolic Disorders
EOS	Reticuloendthelial Disorders
EOSAB	Reticuloendthelial Disorders
FIO_2	Pulmonary Disorders
GLOB	Metabolic and Nutritional Disorders
GLU	Metabolic Disorders
GLU	Endocrine Disorders

Table S6. Labs used as markers for adverse event categories.

Laboratory Code	Adverse Event Category
HCO3A	Pulmonary Disorders
HCT	Hemorrhage
HDL	Metabolic and Nutritional Disorders
HDL	Metabolic Disorders
HEARTRATE	Metabolic Disorders
HEARTRATE	Cardiovascular Disorders
HGB	Hemorrhage
INR	Hemorrhage
K	Metabolic Disorders
K	Endocrine Disorders
K	Metabolic and Nutritional Disorders
K	Cardiovascular Disorders
K	Renal Disorders
LDL	Metabolic Disorders
LDL	Metabolic and Nutritional Disorders
LEUKEST	Genitourinary Disorders
LYM	Reticuloendthelial Disorders
LYMAB	Reticuloendthelial Disorders
MCH	Metabolic and Nutritional Disorders
MCH	Hemorrhage
MCHC	Hemorrhage
MCHC	Metabolic and Nutritional Disorders
MCV	Hemorrhage
MCV	Metabolic and Nutritional Disorders
MG	Gynecologic Disorders
MG	Cardiovascular Disorders
MG	Endocrine Disorders
MONO	Reticuloendthelial Disorders
MONOAB	Reticuloendthelial Disorders
MORPH	Reticuloendthelial Disorders
NA	Renal Disorders
NA	Endocrine Disorders
NA	Metabolic and Nutritional Disorders
NA	Metabolic Disorders
NEUT	Reticuloendthelial Disorders
NEUTAB	Reticuloendthelial Disorders
NHDL	Metabolic and Nutritional Disorders
NHDL	Metabolic Disorders
NITRITE	Reticuloendthelial Disorders
NITRITE	Genitourinary Disorders
O2SATA	Pulmonary Disorders

Table S6. Labs used as markers for adverse event categories.

Laboratory Code	Adverse Event Category
PAXIS	Cardiovascular Disorders
PAXIS	Metabolic Disorders
PCCOM	Metabolic Disorders
PCO2A	Pulmonary Disorders
PHA	Pulmonary Disorders
PHOS	Metabolic Disorders
PHOS	Endocrine Disorders
PLT	Hemorrhage
PLT	Renal Disorders
PO2A	Pulmonary Disorders
PRINTERVAL	Cardiovascular Disorders
PRINTERVAL	Metabolic Disorders
PRO	Hemorrhage
PT	Hemorrhage
PTT	Hemorrhage
PTT	Gastrointestinal Disorders
QRSAXIS	Cardiovascular Disorders
QRSAXIS	Metabolic Disorders
QRSDINTERVAL	Metabolic Disorders
QRSDINTERVAL	Cardiovascular Disorders
QTCINTERVAL	Cardiovascular Disorders
QTCINTERVAL	Metabolic Disorders
QTINTERVAL	Metabolic Disorders
QTINTERVAL	Cardiovascular Disorders
RBC	Hemorrhage
RBC	Metabolic and Nutritional Disorders
RDW	Metabolic and Nutritional Disorders
RDW	Hemorrhage
RPT	Reticuloendthelial Disorders
RR	Metabolic Disorders
RR	Cardiovascular Disorders
SDES	Reticuloendthelial Disorders
SI	Hemorrhage
SLREV	Reticuloendthelial Disorders
SOU	Metabolic Disorders
SPECMN	Genitourinary Disorders
SPG	Genitourinary Disorders
SREQ	Reticuloendthelial Disorders
TBIL	Gastrointestinal Disorders
TCO2A	Pulmonary Disorders

Table S6. Labs used as markers for adverse event categories.	
Laboratory Code	**Adverse Event Category**
TGL	Metabolic and Nutritional Disorders
TGL	Metabolic Disorders
TGL	Endocrine Disorders
TP	Metabolic and Nutritional Disorders
TSH	Metabolic Disorders
TWAVEAXIS	Metabolic Disorders
TWAVEAXIS	Cardiovascular Disorders
UBACT	Genitourinary Disorders
UBLOOD	Genitourinary Disorders
UBT	Hemorrhage
UCLAR	Genitourinary Disorders
UCMT	Genitourinary Disorders
UCOL	Genitourinary Disorders
UGLU	Genitourinary Disorders
UKET	Genitourinary Disorders
UNT	Hemorrhage
UPH	Genitourinary Disorders
UPROT	Genitourinary Disorders
URBC	Genitourinary Disorders
UWBC	Genitourinary Disorders
WBC	Reticuloendthelial Disorders
XMR	Hemorrhage

Table S7. Normal lab value ranges.			
Laboratory Code	**Minimum Normal Value**	**Maximum Normal Value**	**Units**
ALB	3.9	4.8	g/dL
CA	8.9	10.4	mg/dL
EOSAB	0.1	0.79	K/uL
GLOB	2.3	3.5	g/dL
HCT	37	52	%
HGB	12	16.2	g/dl
K	3.6	5	mEq/L
MCV	78	102	fL
MG	1.6	2.7	mg/dL
NA	137	145	mEq/L
PLT	140	450	K/ul
PO2A	75	100	mmHg
RBC	3.8	6.2	x10^6/ul
RDW	11.5	14.5	%
TP	6.3	7.9	g/dL
UPH	5	7	same
WBC	0	3	K/ul
EOS	0	7	%
PCO2A	32	48	mmHg
PHOS	3	4.5	mg/dL
ALT	8	37	U/L
ALKP	44	147	U/L
AST	10	34	U/L
BASO	0	2	%
BUN	7	12	mg/dL
CL	99	108	mmol/L
CO2	20	29	mmol/L
CR	0.8	1.4	mg/dL
GLU	65	110	mg/dl
MCH	31	35	g/dl
MCHC	31	35	g/dl
MONO	4	10	%
MONOAB	0	0.08	K/ul
NEUT	45	74	%
PCCOM	65	110	mg/dl
PTT	18	28	sec
TBIL	0.2	1.3	mg/dL
QTCINTERVAL	380	440	ms
QRSAXIS	-30	90	degrees
INR	0.8	1.2	INR
PT	11	16	s

Table S7. Normal lab value ranges.			
Laboratory Code	**Minimum Normal Value**	**Maximum Normal Value**	**Units**
CHOL	0	200	mg/dL
CHOLHDL	3.5	5	Units
HDL	40	60	mg/dL
LYM	18	52	%
TGL	0	150	mg/dL
TSH	0.5	5	uIU/mL

References

S1. F. Montastruc *et al.*, The importance of drug-drug interactions as a cause of adverse drug reactions: A pharmacovigilance study of serotoninergic reuptake inhibitors in France. *European journal of clinical pharmacology* (2011), doi:10.1007/s00228-011-1156-7.

S2. A. Bate, S. J. W. Evans, Quantitative signal detection using spontaneous ADR reporting. *Pharmacoepidemiol Drug Saf* **18**, 427–436 (2009).

S3. G. Norén, R. Sundberg, A. Bate, A statistical methodology for drug-drug interaction surveillance. *Stat Med.* **27**, 3057–3070 (2008).

S4. W. DuMouchel, D. Pregibon, Empirical Bayes screening for multi-item associations. *Proceedings of the seventh ACM international conference on Knowledge discovery and data mining KDD01*. **10**, 67–76 (2001).

S5. E. L. Olvey, S. Clauschee, D. C. Malone, Comparison of critical drug-drug interaction listings: The Department of Veterans Affairs medical system and standard reference compendia. *Clin Pharmacol Ther* **87**, 48–51 (2010).

S6. H. Bagheri, E. Simiand, J.-L. Montastruc, J.-F. Magnaval, Adverse drug reactions to anthelmintics. *The Annals of pharmacotherapy* **38**, 383–388 (2004).

S7. K. M. Hayden *et al.*, Occupational exposure to pesticides increases the risk of incident AD: The Cache County study. *Neurology* **74**, 1524–1530 (2010).

S8. G. S. Ang, S. A. Simpson, A. R. Reddy, Mycophenolate mofetil embryopathy may be dose and timing dependent. *Am. J. Med. Genet. A* **146A**, 1963–1966 (2008).

S9. M. T. Anderka, A. E. Lin, D. N. Abuelo, A. A. Mitchell, S. A. Rasmussen, Reviewing the evidence for mycophenolate mofetil as a new teratogen: Case report and review of the literature. *Am. J. Med. Genet. A* **149A**, 1241–1248 (2009).

S10. A. E. Lin *et al.*, An additional patient with mycophenolate mofetil embryopathy: Cardiac and facial analyses. *Am. J. Med. Genet. A* **155A**, 748–756 (2011).

S11. N. T. Yedlinsky, F. C. Morgan, P. W. Whitecar, Anomalies associated with failed methotrexate and misoprostol termination. *Obstet Gynecol* **105**, 1203–1205 (2005).

S12. C. R. Wira *et al.*, Sex hormone regulation of innate immunity in the female reproductive tract: The role of epithelial cells in balancing reproductive potential with protection against sexually transmitted pathogens. *Am. J. Reprod. Immunol.* **63**, 544–565 (2010).

S13. C. Kaushic, K. L. Roth, V. Anipindi, F. Xiu, Increased prevalence of sexually transmitted viral infections in women: The role of female sex hormones in regulating susceptibility and immune responses. *J. Reprod. Immunol.* **88**, 204–209 (2011).

S14. 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).

S15. B. L. Nørgaard *et al.*, Efficacy and safety of intravenously administered dofetilide in acute termination of atrial fibrillation and flutter: A multicenter, randomized, double-blind, placebo-controlled trial. Danish Dofetilide in Atrial Fibrillation and Flutter Study Group. *Am. Heart J.* **137**, 1062–1069 (1999).

S16. N. Freemantle, C. Lafuente-Lafuente, S. Mitchell, L. Eckert, M. Reynolds, Mixed treatment comparison of dronedarone, amiodarone, sotalol, flecainide, and propafenone, for the management of atrial fibrillation. *Europace* **13**, 329–345 (2011).

S17. M. Jastrzebski, Proarrhythmic effect of propafenone in patients with atrial fibrillation and atrial flutter. *Kardiol Pol* **66**, 1221–1224; discussion 1225 (2008).

S18. F. Femenia, J. Palazzolo, M. Arce, M. Arrieta, Proarrhythmia induced by propafenone: What is the mechanism? *Indian Pacing Electrophysiol J* **10**, 278–280 (2010).

S19. N. P. Tatonetti *et al.*, Detecting drug interactions from adverse-event reports: Interaction between paroxetine and pravastatin increases blood glucose levels. *Clin Pharmacol Ther*. **90**, 133–142 (2011).