Published in final edited form as:
Trends Pharmacol Sci. 2013 March ; 34(3): . doi:10.1016/j.tips.2013.01.006.

Informatics confronts drug–drug interactions

Bethany Percha[1] and **Russ B. Altman**[2]

[1]Biomedical Informatics Program, Stanford University, 318 Campus Drive, Clark Center S240, Stanford, CA 94305, USA

[2]Departments of Bioengineering, Genetics, and Medicine, Stanford University, 318 Campus Drive, Clark Center S242, Stanford, CA 94305, USA

Abstract

Drug–drug interactions (DDIs) are an emerging threat to public health. Recent estimates indicate that DDIs cause nearly 74 000 emergency room visits and 195 000 hospitalizations each year in the USA. Current approaches to DDI discovery, which include Phase IV clinical trials and post-marketing surveillance, are insufficient for detecting many DDIs and do not alert the public to potentially dangerous DDIs before a drug enters the market. Recent work has applied state-of-the-art computational and statistical methods to the problem of DDIs. Here we review recent developments that encompass a range of informatics approaches in this domain, from the construction of databases for efficient searching of known DDIs to the prediction of novel DDIs based on data from electronic medical records, adverse event reports, scientific abstracts, and other sources. We also explore why DDIs are so difficult to detect and what the future holds for informatics-based approaches to DDI discovery.

Drug–drug interactions: incidence and impact

In 2007, a meta-analysis of 23 clinical studies from around the world revealed that drug–drug interactions (DDIs) cause approximately 0.054% of emergency room visits, 0.57% of hospital admissions, and 0.12% of rehospitalizations [1]. There are 136.1 million emergency room visits [2] and 34.1 million hospital discharges [3] in the USA alone each year. If these percentages are correct, Americans experience DDI events serious enough to send them to the emergency room almost 74 000 times per year, and hospitals admit nearly 195 000 patients per year because of DDIs. Unsurprisingly, DDIs also contribute to increased cost and duration of hospital stays [4].

We should expect DDI incidence to increase as the simultaneous use of multiple drugs becomes more common. According to the Centers for Disease Control (CDC), the percentage of the US population taking at least one prescription drug within the last 30 days increased from 39.1% in 1988–1994 to 47.5% in 2007–2010. During that same period, the percentage of Americans taking three or more prescription drugs rose from 11.8% to 20.8%, and the percentage taking five or more drugs increased from 4.0% to 10.1% (Figure 1a) [5]. Polypharmacy is particularly common among the elderly, making them especially susceptible to DDIs (Figure 1b). In the 2007 study described above, DDIs caused 4.8% of hospital admissions among the elderly, increasing their risk nearly 8.5-fold relative to the general population.

Recently, the improved availability of large amounts of drug-related information has provided fertile ground for the development of informatics-based approaches for studying

Corresponding author: Altman, R.B. (russ.altman@stanford.edu).

DDIs. Increased access to large-scale databases of electronic medical records (EMRs), scientific articles, population-based reports of adverse events, drug labels, and other sources means that researchers can more easily develop comparative, data-driven techniques to recognize, predict, and explain drug interactions. Here we review the difficulties associated with studying DDIs, the breadth of informatics approaches currently available, and the challenges that lie ahead.

Why DDIs are difficult to study

Many known DDIs involve common medications such as antihypertensives, anti-inflammatories, and anticoagulants (Table 1), so a reasonable question is why so many DDIs go undetected for so long. Drugs have occasionally been pulled from the market because of DDIs, but even in such cases the drugs were usually available to the public for years before withdrawal [6,7]. Although *in vitro* laboratory studies, such as DDI assays, can help to alert drug manufacturers and the scientific community to the presence of new DDIs, the difficulty of recognizing DDIs in the clinic, the dose dependence of many DDIs, the nature of the drug approval process, and natural genetic and demographic variation can all delay DDI recognition.

For example, we cannot realistically expect practicing physicians to notice and document most DDIs on their own. Patients who take multiple drugs are often afflicted with multiple comorbidities, and it is difficult to determine whether adverse events are the result of side effects from a single drug, interactions between two or more drugs, or exacerbations of the patient's underlying disease(s). In addition, the number of patients on a particular drug combination, especially within a single practice or hospital, may be small, preventing physicians from recognizing patterns of interactions within their own patient cohorts. Some DDIs are also dose-dependent, which means that a DDI may be unrecognizable unless a patient is dosed at the high end of the approved range for one or both drugs.

In addition, DDIs are often difficult to observe within the environment of a pre-market clinical trial. Phase III clinical trials, the last stage of investigation before a drug enters the market, usually enroll only 1000–3000 individuals. If the test agent interacts with a drug that is not typically prescribed among members of the study population, or if the interaction is weak or rare, few if any study subjects will experience DDI-related symptoms over the course of the study. It comes as no surprise, therefore, that the DDIs we know about are those that cause the most consistent and serious side effects, or those that occur between very commonly prescribed classes of drugs.

Finally, natural human variability can potentially mask the effects of many DDIs. We already know that individuals with particular genetic complements are more sensitive to the effects of certain drugs [8,9] and have a greater chance of experiencing adverse side effects [10] than others. We might therefore expect DDIs to appear more frequently among certain genetic subpopulations, such as people with transporter gene mutations, for example, or those who are 'fast' or 'slow' metabolizers. We also know that demographic variation – differences in age, gender, race/ethnicity, weight, and height, among other factors – explains much of the variation in dosing requirements for many drugs [8,11,12], so we might expect those factors to confound DDI detection as well. The US Food and Drug Administration (FDA) regularly publishes documents containing advice about the design of DDI studies that addresses many of these complicating factors, summaries of which can be found in [13] and [14].

Why informatics?

As we consider these challenges, it becomes obvious that a new discovery paradigm is needed if we are to predict and understand DDIs on a large scale. Today, most DDIs are still discovered by accident in the clinic or during Phase IV clinical trials that take place once a drug is already on the market. (Several of the DDIs in Table 1 reflect this; for example, the interaction between ketoconazole and terfenidine was first proposed in a 1990 case report of a single patient in a Maryland hospital who experienced tachycardia when taking both drugs simultaneously.) *In vivo* and *in vitro* laboratory studies help to alert the scientific community to likely drug interactions [15], and modeling approaches, such as kinetic models of drug levels within the body [16,17], contribute to our understanding of DDI mechanisms. All of these methods, however, analyze only one or a few drug combinations at a time.

Informatics approaches are different. Their strength lies in their ability to manipulate large amounts of data at once, using computers to sift through and rank thousands of potential DDIs. They depend on the existence of large data repositories containing everything from scientific abstracts to structural information about drug molecules to original clinical documents from EMRs. Informatics methods reflect a changing scientific reality. Huge amounts of clinical and laboratory data relevant to DDI prediction are already being collected and stored daily. Although voluminous, these data may be noisy or incomplete. In addition to collecting more data, therefore, we need new ways to search, organize, and manipulate these data to bring DDIs to light.

Integrating data sources

In recent years, several new sources of data have enabled researchers to better identify, predict, and explain DDIs. The papers we review all use at least one of these sources, which we describe below and in Table 2. Although we cannot provide an exhaustive list of all the data sources that contributed to these papers, we have tried to include those that appear in more than one article and/or reflect the breadth of DDI-related data that are currently available.

Scientific articles and abstracts

Medline, the world's largest repository of scientific articles in the biomedical domain, currently contains over 20 million citations and thousands more are added each day. These records, which include the titles, abstracts, and sometimes the full text of biomedical research articles, allow text mining approaches that seek to understand DDIs based on statements made about drugs in the course of biomedical research. The Entrez Programming Utilities (http://www.ncbi.nlm.nih.gov/books/NBK25500/) provide easy search and retrieval tools for Medline data, and it is also possible to download the entirety of Medline in XML format for access on a local machine.

Adverse event reporting systems

Many governments use spontaneous reporting systems to maintain databases of healthcare-related adverse events as a form of post-marketing surveillance. In the USA, healthcare professionals, drug companies, and consumers report adverse events and medication errors to the government via the FDA Adverse Event Reporting System (FAERS). Other countries, such as Canada and Italy, maintain their own adverse event reporting databases and may have different requirements.

Drug information resources

The full text of all drug labels for both prescription and over-the-counter drugs available in the USA can be obtained from the National Library of Medicine DailyMed website. These labels contain valuable information about formulations, side effects, and known drug interactions. The Japan Pharmaceutical Information Center (JAPIC) provides a similar resource for pharmaceuticals available in Japan (http://www.japic.or.jp).

DrugBank, a database that combines detailed chemical, pharmacological, and pharmaceutical data with sequence, structure, and pathway information about drug targets [18], is often used to build gold-standard sets of known interacting drug pairs for DDI detection methods. It also serves as a central repository of useful information about drug structure and activity, such as Simplified Molecular Input Line Entry System (SMILES) codes [19] and Anatomic, Therapeutic, Chemical Classification (ATC) codes [20]. The KEGG DRUG database is a similar resource for all government-approved drugs in the USA, Europe, and Japan. It includes information on molecular structure, target molecules, metabolizing enzymes and transporters, and interactions with other molecules for drugs [21].

Several databases contain valuable data on drug binding and bioactivity. Directed mainly at the pharmaceutical industry, the WOMBAT (World of Molecular Bioactivity) database contains data on protein–ligand binding from 15 320 papers published in medicinal chemistry journals between 1975 and 2009 [22]. WOMBAT contains 331 872 entries representing 1966 unique targets. ChEMBL [23] contains 5.4 million bioactivity measurements for more than 1 million compounds and 5200 protein targets. BindingDB [24] contains 910 836 binding measurements representing interactions among 6263 protein targets and 378 980 small molecules.

PharmGKB is a comprehensive resource that curates information about the impact of genetic variation on drug response [25]. Directed at clinicians and researchers, PharmGKB publishes lexicons of known drug names and synonyms, as well as gene and disease terms and data on genetic pathways. The OFFSIDES database, a comprehensive database of drug effects mined from adverse event reports, and the TWOSIDES database, a database of DDI side effects, are also available from PharmGKB [26].

Finally, the SIDER [27] database contains drug–side-effect associations extracted from drug package inserts via text mining. At the time of writing this review, it contained information about 4199 side effects, 996 drugs, and 100 049 drug–side-effect pairs.

Electronic medical records

The availability of raw data from EMRs has greatly facilitated informatics approaches to DDI prediction by providing a way to validate predicted DDIs. However, use of EMR data at a particular institution generally requires approval by its institutional review board and a member of the study team who is affiliated with that institution. Some academic medical centers have built interfaces to their EMRs to facilitate translational research (e.g., STRIDE [28]) but the main barrier to this type of data is still access. The Informatics for Integrating Biology and the Bedside (i2b2) Center at Partners HealthCare in Boston has recently developed publicly available software that allows researchers to find patients of interest from EMRs via a query tool interface while maintaining patient privacy [29].

Putting it all together

We found 20 papers published in the past few years that integrated these data sources (and others) to curate known information about DDIs into useful databases, predict novel DDIs, and explain DDI mechanisms. These articles fall into four main groups. The first group

focuses on the assembly and curation of knowledge bases about known DDIs. The second group contains methods for extracting statements about DDIs directly from text. The third and fourth groups concentrate on predicting novel drug interactions based on knowledge mined from other sources of data. The third group uses adverse event reports for prediction, whereas the fourth combines pharmacogenomic or chemoinformatic data from a wide range of other sources.

Group 1: building DDI knowledge bases

Several papers reported use of a combination of automated text-mining algorithms and manual curation to reformat known DDI data into useful tools for clinicians and researchers. These papers represent the most applied edge of the informatics domain and provide a window on how the data from the other three groups might eventually be deployed in practice.

Liver cytochrome P450 enzymes (CYPs) are broadly involved in drug metabolism and mediate many known pharmacokinetic DDIs [30]. (A pharmacokinetic DDI occurs when one drug affects the body's ability to metabolize, absorb, distribute, or excrete another drug.) To build SuperCyp [31], a compendium of information about these important enzymes, Preissner *et al.* screened hundreds of Medline abstracts and other sources for facts about CYP-related genetic mutations and their effects on drug metabolism. Their database contains information on 57 CYPs and their interactions with 1170 drugs, as well as supporting evidence for each interaction that includes the study population, where available. Users can study the effects of drug co-administration on CYP-related drug metabolism and browse the 3D structures of the CYPs themselves.

Takarabe *et al.* took a broader approach and developed a DDI resource that incorporates both pharmacokinetic and pharmacodynamic DDIs [32]. (A pharmacodynamic DDI occurs when one drug interacts with the protein target of another drug or another protein within the same pathway.) They structured their database as a hierarchical network in which DDIs can be searched at any level of granularity from the specific (interactions between individual drugs) to the general (interactions between drug classes). They then integrated this resource with the KEGG DRUG database. Users can search over 200 000 known interactions and see whether others are likely by examining drug pairs that share protein targets or are metabolized by the same enzymes.

For certain classes of drugs, it makes sense to build independent, disease-specific DDI databases. The OncoRX project provides a comprehensive resource of information about anticancer drugs and their interactions with complementary and alternative medicines, especially traditional Chinese medicines, used in the course of oncology [33]. The same authors used their database to investigate the clinical relevance of interactions between anticancer drugs and antidepressants [34].

Finally, some large-scale DDI databases are already in use in the clinic. The real-time clinical use of DDI databases is challenging because such resources must typically first be integrated with existing EMR software. The SFINX project [35] is a database of over 8000 DDIs obtained from seven distinct sources, including Medline. It was recently integrated with the Swedish and Finnish computerized decision support systems to facilitate DDI detection in clinical practice.

Group 2: explicit DDI statement extraction

As the volume of knowledge surrounding DDIs grows, manual curation techniques become increasingly less desirable. A few studies mined drug labels or the scientific literature to obtain statements about known DDIs, automating the database curation process performed

manually in Group 1. All of these papers were generated as part of the DDI Extraction Challenge held in 2011 by the SEPLN 2011 Satellite Workshop (http://labda.inf.uc3m.es/DDIExtraction2011). One describes a rule-based approach to DDI extraction from unstructured text that combines shallow parsing and syntactic simplification with pattern matching to find sentences describing DDIs. Another study performed the same task using a machine-learning based method and the third – the winner of the competition – used an ensemble of different machine learning methods [36–38].

Group 3: mining adverse event reports

All of the approaches in Groups 1 and 2 rely on information about known DDIs, even if it is automatically extracted. However, we have seen that our current paradigm for DDI discovery constitutes a significant bottleneck. The third group of papers contains techniques for automatic mining of adverse event reports, in effect using computers to perform post-marketing surveillance on a large scale. These techniques are exciting because they allow detection of DDIs in situations in which the DDI in question is totally unpredictable, or when a Phase IV clinical trial would be too costly or time-consuming.

Harpaz *et al.* used association rule mining to identify 1167 multi-item adverse drug event associations from a corpus of 162 744 adverse event reports from FAERS; of those, 4% were characterized and validated as DDIs by an expert [39]. Association rule mining is a common data-mining technique most often used in online recommendation systems [40]; this paper represents one of the first attempts to use it in biomedical surveillance. It is based on the concept of a ‘market basket’: if a user buys peanut butter and jelly, for example, an association rule-mining system might recognize that he is also likely to buy bread. Similarly, if an adverse event report contains two drug names, the likelihood that it will also refer to a given adverse event can be calculated.

Leone *et al.* used information from a database containing all reports of suspected adverse drug reactions from five Italian regions to identify adverse events associated with DDIs [41]. One year later, Tatonetti *et al.* mined FAERS reports for side-effect profiles related to glucose homeostasis and uncovered a novel interaction between pravastatin and paroxetine that causes a potentially dangerous increase in blood glucose levels [42]. They validated their results using EMR data from Vanderbilt Hospital, Stanford Hospital, and Partners HealthCare, showing the power of a two-step approach in which researchers use spontaneous reporting databases to predict novel DDIs and EMRs to validate them. These authors also developed new signal detection algorithms for identifying DDIs in adverse event reports [43] and for minimizing the effect of confounding variables, such as a patient’s age and underlying disease state [26].

Group 4: combining pharmacogenomic and chemoinformatic data

The fourth and largest group of papers took the Group 3 approach one step further, attempting to predict novel DDIs based on mechanistic and structural information about the drugs themselves and their interactions with proteins. Although these methods are more difficult to validate than those of Group 3, their advantage lies in the fact that they rely mainly on chemical and bioactivity data from laboratory studies rather than clinical data. As a result, they could potentially be used to predict DDIs before drugs enter the market.

Several of these papers rely on statements from the scientific literature. For example, both Percha *et al.* and Tari *et al.* mined Medline for single sentences containing drug–gene relationships [44,45]. Although their relation extraction approaches differ, in both cases these sentences became raw ‘facts’ that were combined to predict DDIs. In both papers, approximately 80% of predicted DDIs matched those in a gold-standard set from DrugBank.

Duke *et al.* used a hybrid approach in which they first mined the scientific literature for abstracts connecting drug names to CYP enzymes using text analysis and manually curated these abstracts to retrieve relevant gene–drug information [46]. They then used an EMR to restrict their findings to drug combinations that were commonly co-prescribed and found five novel combinations that synergistically increased patient risk of myopathy.

Other papers describe systems that integrate knowledge from a plethora of different sources. The INDI system developed by Gottlieb and colleagues uses a gold-standard set of DDIs from DrugBank and Drugs.com, computes seven distinct drug similarity measures based on chemical, side-effect and target protein similarity data, and then ranks other drug pairs based on how similar they were to known DDIs [47]. Similarly, the Drug Interaction Knowledge Base (DIKB) integrates evidence from retrospective studies, clinical trials, metabolic inhibition identification, metabolic catalysis identification, statements, reviews, and observational reports [48]. Curators manually collect evidence about drugs and drug metabolites and classify this evidence by type and strength. They then apply over 1000 different reasoning strategies to their data and evaluate which strategies are most successful at predicting DDIs. These strategies define the different levels and types of evidence required to infer that a DDI is likely [49].

Finally, Vilar *et al.* used a similarity analysis of drug molecular structures to identify novel DDIs by looking for drugs that were structurally similar to those involved in known DDIs [50]. Their method had sensitivity of 0.68, specificity of 0.96, and precision of 0.26 against a gold standard of known DDIs from DrugBank.

Concluding remarks

Recognizing, explaining, and ultimately predicting DDIs constitute a huge challenge for medicine and public health. Not only are DDIs incredibly diverse, both in terms of their biological mechanisms and their severity, but many also occur rarely enough, or produce effects unpredictable enough, to go unnoticed through the entire drug approval process. Given the rate at which polypharmacy is increasing in the USA and around the world, we can only expect this problem to worsen over time. Fortunately, it has coincided with a tremendous increase in the availability of drug-related data in scientific texts, EMRs, adverse event reports from hospitals and consumers, and dozens of databases covering drug mechanisms, side effects, and chemistry.

In this environment, informatics approaches that seek to organize the data surrounding DDIs, facilitate data extraction and database curation, and use data creatively to predict novel DDIs are already bearing fruit. The future promises progress in two areas: (i) evaluating DDI predictions clinically and deciding if they merit intervention recommendations to prescribers, and (ii) using DDI knowledge as a way to probe the molecular mechanisms of drug response. DDIs are important not only with respect to improving healthcare but also because they offer a window into the molecular mechanisms by which drugs interact with the body and with each other. Understanding the molecular details of a DDI might therefore help to illuminate the ways in which the drugs and their targets participate in the molecular networks leading to drug response. Such knowledge might be useful for understanding the basic pathophysiology of health and disease and may provide insights useful for the development of new drugs, new drug combinations, or new uses for old drugs (drug repurposing).

Early informatics approaches to DDI prediction provide us with an interesting case study of what we should expect as ‘big data’ enters the healthcare realm. Large-scale algorithmic approaches to DDI prediction enable us to organize and process orders-of-magnitude more data than ever before, but no matter how sophisticated an algorithm is, its results will still be

fraught with uncertainty. How do we know that the results we are seeing are not artifacts? How can we ensure that we are focusing our attention on DDIs that are most clinically relevant? How can we confirm or deny our results? We can address some of these questions with increasingly sophisticated statistics, but ultimately even 'big data' predictions will still need to be validated in the clinic and laboratory. We envision a paradigm in which broad, data-driven algorithms alert the healthcare community to new evidence of DDIs, EMRs provide an additional layer of support to boost the best predictions to the top, and laboratory experiments serve as the final arbiter, confirming or denying newly predicted DDIs. It looks as if that day is coming soon.

References

1. Becker ML, et al. Hospitalisations and emergency department visits due to drug–drug interactions: a literature review. Pharmacoepidemiol Drug Saf. 2007; 16:641–651. [PubMed: 17154346]
2. Centers for Disease Control/National Center for Health Statistics. National Hospital Ambulatory Medical Care Survey 2009. Centers for Disease Control/National Center for Health Statistics; 2010.
3. Hall MJ, et al. National hospital discharge survey: 2007 summary. Natl Health Stat Rep. 2010; 29:1–20.
4. Moura CS, et al. Drug–drug interactions associated with length of stay and cost of hospitalization. J Pharm Pharm Sci. 2009; 12:266–272. [PubMed: 20067703]
5. National Center for Health Statistics. Health, United States, 2011: With Special Feature on Socioeconomic Status and Health. National Center for Health Statistics; 2012.
6. Broadhurst P, Nathan AW. Cardiac arrest in a young woman with the long QT syndrome and concomitant astemizole ingestion. Br Heart J. 1993; 70:469–470. [PubMed: 8260280]
7. Rogers R, Prpic R. Profound symptomatic bradycardia associated with combined mibefradil and beta-blocker therapy. Med J Aust. 1998; 169:425–427. [PubMed: 9830391]
8. Klein TE, et al. International Warfarin Pharmacogenetics Consortium. Estimation of the warfarin dose with clinical and pharmacogenomic data. N Engl J Med. 2009; 360:753–764. [PubMed: 19228618]
9. Shuldiner AR, et al. Association of cytochrome P450 2C19 genotype with the antiplatelet effect and clinical efficacy of clopidogrel therapy. J Am Med Assoc. 2009; 302:849–857.
10. Neuvonen PJ, et al. Drug interactions with lipid-lowering drugs: mechanisms and clinical relevance. Clin Pharmacol Ther. 2006; 80:565–581. [PubMed: 17178259]
11. Cheymol G. Effects of obesity on pharmacokinetics: implications for drug therapy. Clin Pharmacokinet. 2000; 39:215–231. [PubMed: 11020136]
12. Xie HG, et al. Molecular basis of ethnic differences in drug disposition and response. Annu Rev Pharmacol Toxicol. 2001; 41:815–850. [PubMed: 11264478]
13. Huang SM, et al. Drug interaction studies: study design, data analysis, and implications for dosing and labeling. Clin Pharmacol Ther. 2007; 81:298–304. [PubMed: 17259955]
14. Zhang L, et al. Predicting drug–drug interactions: an FDA perspective. AAPS J. 2009; 11:300–306. [PubMed: 19418230]
15. Brown HS, et al. Prediction of *in vivo* drug–drug interactions from *in vitro* data: factors affecting prototypic drug–drug interactions involving CYP2C9, CYP2D6 and CYP3A4. Clin Pharmacokinet. 2006; 45:1035–1050. [PubMed: 16984215]
16. Li L, et al. Drug–drug interaction prediction: a Bayesian meta-analysis approach. Stat Med. 2007; 26:3700–3721. [PubMed: 17357990]
17. Zhou J, et al. A new probabilistic rule for drug–dug interaction prediction. J Pharmacokinet Pharmacodyn. 2009; 36:1–18. [PubMed: 19156505]
18. Wishart DS, et al. DrugBank: a knowledgebase for drugs, drug actions and drug targets. Nucleic Acids Res. 2008; 36(Suppl 1):D901–D906. [PubMed: 18048412]
19. Weininger D. SMILES, a chemical language and information system. 1. Introduction to methodology and encoding rules. J Chem Inf Comput Sci. 1988; 28:31–36.

20. Skrbo A, et al. Classification of drugs using the ATC system (Anatomic, Therapeutic, Chemical Classification) and the latest changes. Med Arh. 2004; 58(Suppl 2):138–141. [PubMed: 15137231]
21. Kanehisa M, et al. KEGG for representation and analysis of molecular networks involving diseases and drugs. Nucleic Acids Res. 2010; 38(Suppl 1):D355–D360. [PubMed: 19880382]
22. Olah, M., et al. WOMBAT: world of molecular bioactivity. In: Oprea, TI., editor. Chemoinformatics in Drug Discovery. Wiley-VCH; 2004. p. 223-240.
23. Gaulton A, et al. ChEMBL: a large-scale bioactivity database for drug discovery. Nucleic Acids Res. 2012; 40:D1100–D1107. [PubMed: 21948594]
24. Liu T, et al. BindingDB: a web-accessible database of experimentally determined protein–ligand binding affinities. Nucleic Acids Res. 2007; 35(Suppl 1):D198–D201. [PubMed: 17145705]
25. Klein TE, et al. Integrating genotype andphenotype information: an overview of the PharmGKB project. Pharmacogenomics J. 2001; 1:167–170. [PubMed: 11908751]
26. Tatonetti NP, et al. Data-driven prediction of drug effects and interactions. Sci Transl Med. 2012; 4:125ra31.
27. Kuhn M, et al. A side effect resource to capture phenotypic effects of drugs. Mol Syst Biol. 2010; 6:343. [PubMed: 20087340]
28. Lowe HJ, et al. STRIDE – an integrated standards-based translational research informatics platform. AMIA Annu Symp Proc. 2009; 2009:391. [PubMed: 20351886]
29. Murphy SN, et al. Serving the enterprise and beyond with informatics for integrating biology and the bedside (i2b2). J Am Med Inform Assoc. 2010; 17:124–130. [PubMed: 20190053]
30. Thummel KE, Wilkinson GR. *In vitro* and *in vivo* drug interactions involving human CYP3A. Annu Rev Pharmacol Toxicol. 1998; 38:389–430. [PubMed: 9597161]
31. Preissner S, et al. SuperCYP: a comprehensive database on cytochrome P450 enzymes including a tool for analysis of CYP–drug interactions. Nucleic Acids Res. 2010; 38(Suppl 1):D237–D243. [PubMed: 19934256]
32. Takarabe M, et al. Network-based analysis and characterization of adverse drug–drug interactions. J Chem Inf Model. 2011; 51:2977–2985. [PubMed: 21942936]
33. Yap KYL, et al. An onco-informatics database for anticancer drug interactions with complementary and alternative medicines used in cancer treatment and supportive care: an overview of the OncoRx project. Support Care Cancer. 2010; 18:883–891. [PubMed: 19756775]
34. Chan A, et al. Electronic database to detect drug–drug interactions between antidepressants and oral anticancer drugs from a cancer center in Singapore: implications to clinicians. Pharmacoepidemiol Drug Saf. 2011; 20:939–947. [PubMed: 21732473]
35. Böttiger Y, et al. SFINX – a drug–drug interaction database designed for clinical decision support systems. Eur J Clin Pharmacol. 2009; 65:627–633. [PubMed: 19205683]
36. Segura-Bedmar I, et al. Using a shallow linguistic kernel for drug–drug interaction extraction. J Biomed Inform. 2011; 44:789–804. [PubMed: 21545845]
37. Segura-Bedmar I, et al. A linguistic rule-based approach to extract drug-drug interactions from pharmacological documents. BMC Bioinformatics. 2011; 12 (Suppl 2):S1. [PubMed: 21489220]
38. Thomas, P., et al. Relation extraction for drug–drug interactions using ensemble learning. Proceedings of the First Challenge Task on Drug–Drug Interaction Extraction (DDIExtraction 2011); CEUR Workshop Proceedings; 2011.
39. Harpaz R, et al. Mining multi-item drug adverse effect associations in spontaneous reporting systems. BMC Bioinformatics. 2010; 11 (Suppl 9):S7. [PubMed: 21044365]
40. Agrawal R, et al. Mining association rules between sets of items in large databases. ACM SIGMOD Record. 1993; 22:207–216.
41. Leone R, et al. Identifying adverse drug reactions associated with drug–drug interactions: data mining of a spontaneous reporting database in Italy. Drug Saf. 2010; 33:667–675. [PubMed: 20635825]
42. Tatonetti NP, et al. Detecting drug interactions from adverse-event reports: interaction between paroxetine and pravastatin increases blood glucose levels. Clin Pharmacol Ther. 2011; 90:133–142. [PubMed: 21613990]

43. Tatonetti NP, et al. A novel signal detection algorithm for identifying hidden drug–drug interactions in adverse event reports. J Am Med Inform Assoc. 2012; 19:79–85. [PubMed: 21676938]
44. Percha B, et al. Discovery and explanation of drug–drug interactions via text mining. Pac Symp Biocomput. 2012; 2012:410. [PubMed: 22174296]
45. Tari L, et al. Discovering drug–drug interactions: a text-mining and reasoning approach based on properties of drug metabolism. Bioinformatics. 2010; 26:i547–i553. [PubMed: 20823320]
46. Duke JD, et al. Literature based drug interaction prediction with clinical assessment using electronic medical records: novel myopathy associated drug interactions. PLoS Comput Biol. 2012; 8:e1002614. [PubMed: 22912565]
47. Gottlieb A, et al. INDI: a computational framework for inferring drug interactions and their associated recommendations. Mol Syst Biol. 2012; 8:592. [PubMed: 22806140]
48. Boyce R, et al. Computing with evidence. Part I. A drug-mechanism evidence taxonomy oriented toward confidence assignment. J Biomed Inform. 2009; 42:979–989. [PubMed: 19435613]
49. Boyce R, et al. Computing with evidence. Part II. An evidential approach to predicting metabolic drug–drug interactions. J Biomed Inform. 2009; 42:990–1003. [PubMed: 19539050]
50. Vilar S, et al. Drug–drug interaction through molecular structure similarity analysis. J Am Med Inform Assoc. 2012; 19:1066–1074. [PubMed: 22647690]
51. Monahan BP, et al. Torsades de pointes occurring in association with terfenadine use. J Am Med Assoc. 1990; 264:2788–2790.
52. Okuda H, et al. Lethal drug interactions of sorivudine, a new antiviral drug, with oral 5-fluorouracil prodrugs. Drug Metab Dispos. 1997; 25:270–273.
53. Schmassmann-Suhijar D, et al. Rhabdomyolysis due to interaction of simvastatin with mibefradil. Lancet. 1998; 351:1929–1930. [PubMed: 9654265]
54. Ferslew KE, et al. A fatal drug interaction between clozapine and fluoxetine. J Forensic Sci. 1998; 43:1082–1085. [PubMed: 9729831]
55. Hung IFN, et al. Fatal interaction between clarithromycin and colchicine in patients with renal insufficiency: a retrospective study. Clin Infect Dis. 2005; 41:291–300. [PubMed: 16007523]
56. Schelleman H, et al. Warfarin with fluoroquinolones, sulfonamides, or azole antifungals: interactions and the risk of hospitalization for gastrointestinal bleeding. Clin Pharmacol Ther. 2008; 84:581–588. [PubMed: 18685566]
57. Francis L, et al. Fatal toxic myopathy attributed to propofol, methylprednisolone, and cyclosporine after prior exposure to colchicine and simvastatin. Clin Rheumatol. 2008; 27:129–131. [PubMed: 17628739]

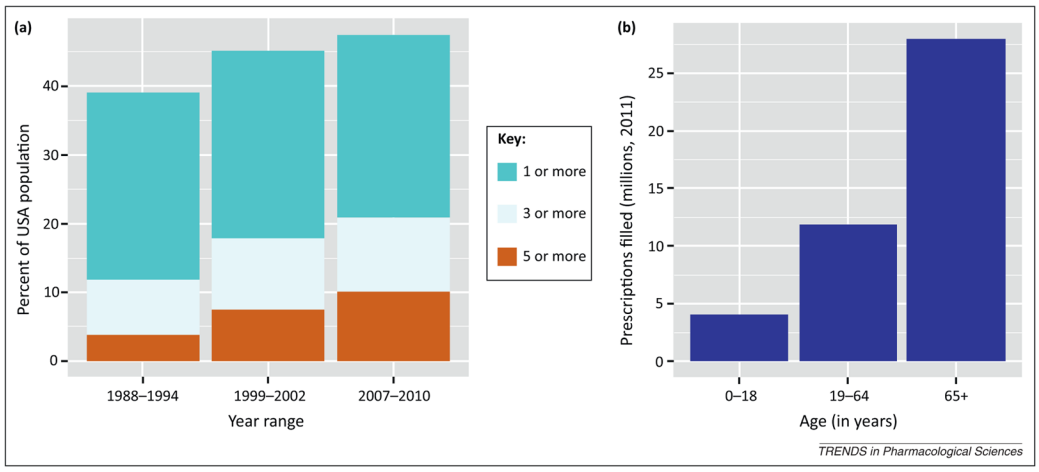

Figure 1.

(a) Number of prescription drugs used in the past 30 days by percentage of the USA population (age-adjusted estimates). Source: National Center for Health Statistics. Health, United States, 2011: With Special Feature on Socioeconomic Status and Health. Hyattsville, MD. 2012. Table 99: Prescription drug use in the past 30 days, by sex, age, race, and Hispanic origin: United States, selected years 1988–1994 through 2007–2010. **(b)** Average number of prescriptions filled in 2011 in the USA by age. The data include both new prescriptions and refills, brand name, and generic drugs. Source: The Kaiser Family Foundation, statehealthfacts.org, accessed September 14, 2012. Data source: SDI Health, LLC: Special Data Request, 2012. Calculations based on 2011 population estimates from the US Census Bureau.

Table 1

Examples of known drug interactions

Year reported	Drug combination	Effect	Refs
1990	• Ketoconazole • Terfenidine	Ventricular arrhythmias	[51]
1993	• Astemizole[a] • CYP3A4 inhibitors (grapefruit juice, erythromycin)	Prolonged QTc interval, arrhythmias	[6]
1997	• Sorivudine • Fluorouracil	Fatal toxicity	[52]
1998	• Mibefradil[a] • Various cardioactive drugs including β blockers and statins	Bradycardia, rhabdomyolysis	[7,53]
1998	• Clozapine • fluoxetine	Death due to depression of nervous, respiratory, and cardiovascular systems	[54]
2005	• Clarithromycin • Colchicine	Death in patients with renal insufficiency	[55]
2008	• Warfarin • Antibiotics	Gastrointestinal bleeding	[56]
2008	• Propofol, methylprednisolone, cyclosporine • Colchicine, simvastatin	Fatal toxic myopathy	[57]
2011	• Pravastatin • Paroxetine	Increased blood glucose	[42]

[a]Mibefradil and astemizole were pulled from the market because of their interactions with other drugs, even though both were considered safe on their own.

Table 2

Data sources useful for the study of drug–drug interactions[a]

Data source	Content	Availability	Original reference	Related papers
DailyMed	Raw drug labels in XML format, for prescription and over-the-counter drugs in the USA (provided by NLM)	http://dailymed.nlm.nih.gov		[36–38]
DrugBank	Chemical, pharmacological, and pharmaceutical drug information; sequence, structure, and pathway information about drug targets	www.drugbank.ca	[18]	[44,45,50]
FAERS	FDA's adverse event reporting system, containing adverse event reports from USA hospitals; full data available in relational database formats	http://www.fda.gov/Drugs/GuidanceComplianceRegulatoryInformation/Surveillance/AdverseDrugEffects/		[39,41,42]
KEGG DRUG	Comprehensive resource for approved drugs in Japan, USA, and Europe containing chemical structures and/or components, drug target, metabolizing enzyme, and other molecular interaction network information	http://www.genome.jp/kegg/drug/	[21]	[32,50]
PharmGKB	Information about the impact of genetic variation on drug response; lexicons of drug, gene, and disease names; pathway information for genes	http://www.pharmgkb.org	[25]	[33,44]
PubMed (Medline)	Scientific papers and abstracts from 1948 to the present day; over 22 million citations and counting; can download all abstracts and metadata in XML format with a license	http://www.ncbi.nlm.nih.gov/pubmed		[44–46]
SIDER	Drug–side-effect frequencies obtained from drug labels and other documents via text mining	http://sideeffects.embl.de	[27]	[26,47]
WOMBAT (World of Molecular Bioactivity)	Small-molecule chemogenomics database containing bioactivity data on protein targets	Proprietary database owned by Sunset Molecular; http://www.sunsetmolecular.com	[22]	[47]

[a]Original reference refers to the original citation for the resource, if one is available. Related papers includes papers we reviewed that used the source directly, as well as those for which the source would be a logical alternative to a source used in the paper.