

ADVANCES IN DRUG DEVELOPMENT

Current Developments in Oncology Drug Research

Section Editor: Mark J. Ratain, MD

Mining the Internet for Drug Information



Nigam Shah, MBBS, PhD
Assistant Professor of Medicine
Stanford University School of Medicine
Stanford, California

H&O How did you come to research the use of the Internet to study drug safety?

NS Dr Eric Horvitz, a distinguished scientist at Microsoft and an alum of our graduate program, gave the keynote speech at one of our graduate program retreats. Over a campfire, Eric, Dr Russ Altman (a professor of bioengineering, genetics, and medicine at Stanford University) and I were speaking about drug safety and how common it is for people to search the Internet for information about drug safety and side effects. Google Flu Trends, a tool for tracking flu cases, had just been released, and we began wondering if it might be possible to detect adverse events in the same way that others were detecting flu. That speculation is how the project started.

H&O Could you describe the study you then conducted?

NS Russ Altman had a student, Dr Nicholas Tatonetti (now an assistant professor of biomedical informatics at Columbia University) who had identified a drug-drug interaction by analyzing adverse event reports submitted to the US Food and Drug Administration (FDA) to uncover new drug-drug interactions. We used this work as our way in. With a predicted interaction in hand—an association between paroxetine and pravastatin that leads to high blood glucose—we tried to investigate whether we could have found the same interaction on the Internet via analyzing search logs.

Because finding just 1 instance of this interaction online would have been anecdotal, I proposed that we look at a set of approximately 30 known drug-drug inter-

actions and 30 drug pairs that did not have any known interaction. Among these samples, how often would we correctly identify pairs with interactions? The answer to that question was: approximately 80% of the time. The results of this work were published in the *Journal of the American Medical Informatics Association* in 2013.

H&O You followed up this work with another, larger study. Could you discuss that research?

NS After these interesting results, we conducted a more exhaustive evaluation, looking at more than 300 drug adverse event associations. We expanded our search beyond drug-drug interactions because we wanted to investigate whether this approach might be a viable pharmacovigilance tool for phase 4 surveillance after a drug is approved based on a phase 3 clinical trial.

H&O What were your study methods?

NS We were trying to uncover toxicities or other issues related to FDA-approved medications using data collected from searches people did on the Internet. If an individual searched for a particular drug and a particular condition, then that became information we could collect.

We created a model that distinguished between patients and healthcare professionals based on the language used in the query. For example, patients are more likely to search for “heart attack,” whereas healthcare professionals are more likely to search for “myocardial infarction.” With that broad distinction, we categorized searchers as patient or professional.

We also catalogued the variety of search terms that would lead different users to the same Wikipedia page. What other terms would bring a person to the page for myocardial infarction? If we track a search and find that a person entered one term but then landed on the Wikipedia page for myocardial infarction, then we assumed that this information was what the searcher was looking for. Through this approach, we generated sets of words that consistently pointed to a certain Wikipedia topic across large numbers of users. These groups of words included technical and colloquial terms for symptoms and adverse events.

These 2 approaches—creating a model that distinguished patients from healthcare professionals, and building a synonym set of colloquially used words to describe medical symptoms—were applied to a very large data set. The data set, which comprised the search activity of 80 million users, and the methods were compared using a known set of true positives and true negatives. As a result, we could quantify how many drug associations we would have identified and also how many toxic effects would have been picked up before the FDA recognized them.

H&O With so many drugs and potential interactions and side effects, how do you decide where to focus?

NS When we consider any tool to detect drug adverse events, 2 crucial questions are: (1) How soon can we detect it? and (2) What is the impact of detecting it? Uncovering a side effect that is serious and affects many people is more important than, say, finding nausea in 3 people. Both might be occurring, but our efforts are best directed toward finding severe adverse events in large numbers of people. Thus a tool that can uncover a serious toxicity in large numbers of people 6 months before it might otherwise be found would be more valuable than a tool to find a mild effect in a few people 1 year early. To put it another way, if you can only prevent one crime, would you prevent a homicide or the theft of a garden hose?

Historically, a scientist would look at 10 true positives and 10 true negatives and ask: How many of the true positives did we get correct? But instead, we are weighting the true positives based on the notion of importance, which is defined in terms of severity and population size.

The context is also important. Nausea, vomiting, and fainting spells from medication for late-stage colon cancer raises different considerations than heart attacks associated with an over-the-counter drug.

The question of context brings us to the other major consideration about this research: What is the impact of detecting the adverse event? If a researcher uncovers leg cramps as a side effect of antineoplastic drugs, are people likely to stop taking those medications? The answer is no.

Alerting the authorities and changing the product label would serve no real purpose. It might be necessary and helpful, but it is not urgent. By contrast, if an over-the-counter drug is causing even a modest increase in heart attacks, we need to sound the alarm.

H&O Your research on Internet pharmacovigilance is continuing. Could you describe the study you are now conducting?

NS We are pushing the notion of “holistic pharmacovigilance,” which is to say, gathering data from all possible sources. The FDA has its adverse event reporting system, which healthcare professionals report to. A patient’s electronic health record contains an individual’s medical history, so would include even those events that are not reported to the FDA. We also have access to data from UpToDate.com, a resource used by health professionals that summarizes the latest state of medicine and contains a search engine. And there are general searches by individuals through search engines like Google, Yahoo!, and Bing, which we access via our collaboration with Microsoft Research.

We have a grant from the National Institutes of Health to develop methods to combine data from these multiple sources. With this approach, we can look at what is being reported to the FDA, what is being observed in patients, what patients are searching for, and what doctors are searching for. If we cross-reference all 4, then we can probably zero in on the highest priority drug safety issues.

H&O How can healthcare professionals and patients make practical use of such findings?

NS The FDA has initiated a data mining group, which my colleague Rave Harpaz (of the department of biomedical informatics at Columbia University) and I interact with closely. Our methods need to be validated, though, and that is what this current 5-year research program is focusing on. But the FDA is following our work as it proceeds.

H&O Could you describe some of the mathematics behind this approach to data mining?

NS As we discussed, there are at least 4 sources of information. The simplest way to combine that information is to take the average—add the number of incidents reported via each source and divide it by 4.

But when taking the average, the sources need to be weighted by reliability. To ascertain reliability, we need to know how often an individual source accurately identifies a true positive or true negative effect. The importance of a source can be weighted on a case-by-case basis. For a given source, we can assess whether the confidence in a particular

drug-toxicity association is very tight or very wide. We want our calculations to reflect the relatively greater importance of a source that delivers very tight confidence intervals.

We published a method for combining signals from multiple sources using this type of meta-analysis approach where credit was assigned to each individual source based on how often the information generated from it is accurate. We presented that work at the Knowledge Discovery and Data Mining (KDD) conference in 2013.

H&O Do you ever try to identify characteristics of the people behind the searches? Would it be important to know whether a search was done by a female older than 50 years vs younger than 25 years, for example?

NS We are not pursuing this approach right now, given the rules around privacy. Because issues of surveillance are still fresh in people's minds, a proposal to work with identified user data, even though useful, would likely be viewed as an invasion of privacy.

H&O If your research confirms the efficacy of mining the Internet for drug safety information, then should patients be encouraged to report side effects online to help uncover adverse events?

NS Yes. We have drug safety reporting mechanisms on the Internet, but the average person is not aware of them. Also, if a several-page form needs to be completed and faxed, then many people will be reluctant to follow through.

However, it is unlikely that our one study will make any dent because the issue is so contentious. It is difficult to understand why people will readily enter credit card information in a website, but will not provide prescription information. We protect health information more than financial information. Such paranoia might be the result of horror stories around denial of insurance, which until recently was a real danger. I am hopeful that the public perceptions around the value of sharing health information will change soon.

H&O What is needed to speed up the research so that patients learn of adverse events more quickly?

NS My group at Stanford is advocating for "real-life" drug safety surveillance. Most efforts monitor single drugs at a time. In phase 4 surveillance, a new drug comes on the market and the pharmaceutical company that sells it must monitor its safety.

But real patients have multiple comorbidities and take more than 1 drug. Few companies monitor drug-

drug interactions. If there are 40 million people with an average of 3 comorbid conditions and taking 3 or more drugs, then the number of potential interactions is very large. The potential number of interactions in someone taking 3 drugs and having 3 comorbid conditions is greater than the estimated number of grains of sand on the planet. However, combing through patient data to count what interactions actually occur yields a relatively small number. We need to search for interactions within the context of a disease and within the context of combinations of drugs. Searching for the next rofecoxib—a single drug with a single toxicity that was recognized late—will not be very productive.

We need to focus on the riskiest prescriptions. For example, if an individual has diabetes, hypertension, and high cholesterol, and we know that this patient is prescribed lisinopril for hypertension, metformin for diabetes, and a statin for high cholesterol, we can search for interactions among these different prescriptions. The results can also inform doctors' prescribing habits; some combinations of drugs may be safer than others.

H&O Could data mining be done on a global basis?

NS In theory, yes. I am a member of a consortium called Observational Health Data Sciences and Informatics, which includes researchers at Stanford and Columbia University, individuals from a few pharmaceutical companies, and then a broader circle of interested scientists. There is a consensus among the consortium members that one-drug-at-a-time safety monitoring is not productive, and we need to search in a more comprehensive manner. We are working on creating an infrastructure so that if an association between 2 drugs is detected at Stanford, then colleagues at Columbia or a pharmaceutical company can confirm or refute that finding. If an association is detected at 3 sites, then we know that we need to investigate further.

Suggested Reading

Iyer SV, Harpaz R, LePendu P, Bauer-Mehren A, Shah NH. Mining clinical text for signals of adverse drug-drug interactions. *J Am Med Inform Assoc*. 2014;21(2):353-362.

Jung K, LePendu P, Chen WS, et al. Automated detection of off-label drug use. *PLoS ONE*. 2014;9(2):e89324.

Leeper NJ, Bauer-Mehren A, Iyer SV, Lependu P, Olson C, Shah NH. Practice-based evidence: profiling the safety of cilostazol by text-mining of clinical notes. *PLoS ONE*. 2013;8(5):e63499.

LePendu P, Iyer SV, Bauer-Mehren A, et al. Pharmacovigilance using clinical notes. *Clin Pharmacol Ther*. 2013;93(6):547-555.

White RW, Harpaz R, Shah NH, Dumouchel W, Horvitz E. Toward enhanced pharmacovigilance using patient-generated data on the Internet [published online April 8, 2014]. *Clin Pharmacol Ther*. doi:10.1038/clpt.2014.77.

White RW, Tatonetti NP, Shah NH, Altman RB, Horvitz E. Web-scale pharmacovigilance: listening to signals from the crowd. *J Am Med Inform Assoc*. 2013;20(3):404-408.