NSIDES: DRUG EFFECT DISCOVERY USING THE FDA ADVERSE REPORTING SYSTEM

RAMI S VANGURI
Department of Biomedical Informatics, Columbia University,
New York, NY 10032 USA
E-mail: r.vanguri@columbia.edu

JOSEPH D ROMANO
Department of Biomedical Informatics, Columbia University,
New York, NY 10032 USA
E-mail: r.vanguri@columbia.edu

TAL LORBERBAUM
Department of Biomedical Informatics, Columbia University,
New York, NY 10032 USA
E-mail: r.vanguri@columbia.edu

VICTOR NWANKWO
Department of Biomedical Informatics, Columbia University,
New York, NY 10032 USA
E-mail: r.vanguri@columbia.edu

CHOONHAN YOUN
San Diego Supercomputer Center, University of California, San Diego,
La Jolla, CA 92093 USA
E-mail: cyoun@sdsc.edu

NICHOLAS P TATONETTI
Department of Biomedical Informatics, Columbia University,
New York, NY 10032 USA
E-mail: nick.tatonetti@columbia.edu

Adverse drug events are a leading cause of morbidity and mortality around the world. Regulatory agencies, such as the Food and Drug Administration (FDA), maintain large collections of adverse event reports. These spontaneous reporting systems provide an opportunity to retrospectively study drug effects from patient population data. The FDA Adverse Event Reporting System contains millions of reports, thousands of drugs and thousands of effects. As a result, calculating the statistical significance of every possible side effect is computational intensive. The problem grows exponentially when considering drug interactions. By using a combination of distributed high computing computing made available by the Open Science Grid and designing an on-demand system to study higher order drug effects, we create an extensive database of drug effects.

1. Introduction

Spontaneous reporting systems such as the FDA Adverse Event Reporting System (FAERS) are important resources for detecting drug adverse events after a drug is approved (pharmacovigilance). However, pharmacovigilance algorithms often lead to many false positive and false negative findings due to issues of confounding, and detection of drug-drug interactions is an even greater challenge. We previously developed databases for off-label drug effects (Offsides) and drug interactions (Twosides) that account for these limitations using a novel Statistical Correction for Uncharacterized Bias (SCRUB).[1] These databases are available for download to the public. In addition to updates to the algorithm, we have developed a gateway to access the databases, known as nSides. nSides aims to make these databases accessible to researchers, clinicians, and patients alike and contains additional features related to drug safety. Since it is not feasible to generate models for every possible drug combination, we develop a novel model request system which submits jobs to the Open Science Grid and appends the results to the databases for future access.

2. Data Sources

There are several data sources which are involved in nSides. We use a curated version of the FDA Adverse Event Reporting System (FAERS) known as Adverse Event Open Learning through Universal Standardization (AEOLUS).[2] AEOLUS aims to clean and normalize the data by removing duplicate cases. This is done by applying standardized vocabularies in the form of RxNorm to map drug names and SNOMED-CT to map outcomes. The AEOLUS dataset is publicly available.

The algorithm used to develop the databases used for nSides are similar to the previously developed Offsides and Twosides databases using raw FAERS data.[1] A standard signal detection algorithm involves conducting a disproportionality analysis by comparing the observed reporting frequency of a drug and outcome to the expected reporting frequency of all other drugs and the outcome. The metric is known as a Proportional Reporting Ratio (PRR). If the outcome occurred by chance, the frequencies will be equal and the PRR will be one. If the PRR is much larger than one, the null hypothesis is rejected. To reduce sampling variance and selection bias, propensity score matching is implemented on the FAERS data to form the groups used in the disproportionality analysis. This procedure, known as SCRUB, matches cases and controls between patients exposed and not exposed to a particular drug (Offsides) or two drugs (Twosides) to mitigate confounding biases.

3. Methods

3.1. *Computational Challenge*

References

1. N. P. Tatonetti, P. P. Ye, R. Daneshjou and R. B. Altman, *Science Translational Medicine* **4**, 125ra31 (2012).
2. J. M. Banda, L. Evans, R. S. Vanguri, N. P. Tatonetti, P. B. Ryan and N. H. Shah, *Scientific data* **3** (2016).