Bin Lu
Chris Mynatt
Srija Ummaneni
Victoria Golden

A Review of FAERS Data for Estimation of Risk Prone Populations

Introduction:

With the rapid development of medical science, increasing quantities of available drugs have reduced the suffering of more and more people. However, over time there has been an increase in reports of drug adverse reactions [1]. In recent years, countries have been committed to continuously strengthening the management of the safe use of drugs; but, the incidence of phytotoxicity has increased. Drug safety issues are a global concern, as an increasing number of people become disabled or die each year due to adverse drug reactions. Presently, many countries have established laws, regulations and quality standards for drug supervision; however, the situation facing drug safety is still serious. According to the World Health Organization (WHO), nearly six to twelve percent of the annual inpatients are hospitalized because of adverse drug reactions [2]. Among patients who have been hospitalized, approximately 10% to 20% have drug adverse reactions [3]. In general, most adverse drug reactions are mild and tolerable, but adverse reactions to certain drugs can cause disability or death. Within the United States, adverse drug reactions are among the top 10 leading causes of death [4]. The variable nature of drug reactions is caused by individual differentiation in drug tolerance. When drugs enter the market, clinical trials are needed to reduce the occurrence of adverse reactions. However, most clinical trials exclude children, the elderly, pregnant women and other infected groups from the test population [5], so the results of clinical trials do not always accurately represent all patient cohorts.

Within the history of drug development, people's understanding of pharmacology was not profound because of the limited levels of knowledge at that time. Even currently, imperfections in the regulatory system comprise one of the most important reasons patients may encounter adverse drug reactions. Arguably, the most famous example is the Thalidomide event that occurred in the 1960s [6]. The occurrence of this incident prompted multiple nations to begin to recognize the importance of regulating drug safety. The United States government has amended its federal food and drug regulations to require drug manufacturers to submit safety information to the FDA for approval.

Based on this situation, analyzing the metadata of drug adverse reactions to estimate risk prone populations has great significance. Firstly, it can promote clinical rational use of drugs. Analyzation of metadata from adverse drug reaction reports to identify the applicable population and high-risk groups of different drugs carries multiple benefits, as it also helps to promote good communication with drug manufacturers, jointly publish drug safety information, improve drug specifications, identify potential safety hazards, and provide risk warnings. This helps medical practitioners to rationally choose drugs and minimize the occurrence of adverse reactions. In addition, the analysis and evaluation of drug reactions provides a basis for the FDA to rectify or phase out drugs from the market [7]. Government agencies can report on adverse drug reactions based on analysis reports and give safety recommendations. For drugs with serious adverse reactions, policymakers must determine whether it is necessary to suspend the medicine’s distribution. In scenarios where no alternative medicine is available, the presence of serious adverse reactions could promote the development of new alternatives.

Therefore, the purpose of our research is to identify characteristics of high-risk populations that are prone to adverse drug reactions by using the data collected from the FDA Adverse Event Reporting System

(FAERS), with the goal of providing suggestions to medical institutions to reduce the incidence of adverse drug reactions. The proposal and final report can be found at the website link: http://mason.gmu.edu/~blu4/.

Problem Statement:

The importance of determination of high-risk demographics of patients is well understood from an insurance perspective but is often nebulous in definition when being assigned by a practitioner to a patient. This is partially due to the qualitative nature of a typical high-risk patient diagnosis [8][9]. Typically, identification of a high-risk patient is determined by a combination of multiple variables, including age, weight, number and type of medications, and reported symptoms [10]. However, while efforts exist to expand current knowledge regarding identification of high-risk patients, there is currently no deterministic methodology formally approved by the FDA for medical practitioners in which to formally characterize a patient as being generally high risk especially when considered within the context of patients that may experience adverse drug reactions [11][12].

As such, the authors suggest the following research question to be addressed:

1. Within the United States, what characterizes the typical patient for which an adverse drug reaction report has been submitted to the United States Food & Drug Administration?

Objectives:

The current objective of this research is to determine a general profile of a U.S. based patient for which an adverse drug reaction report has been generated and submitted to the FAERS. The project has the following sub-objectives:

1. Analyze relationships between individual physiological metadata and drug usage counts.
2. Identify differentiation in trends between male or female populations.
3. Evaluate FAERS reports with geospatial information to review trends in patient demographics between varying locations.

Results from this study may prove valuable to the healthcare industry when determining high risk patients, as utilization for this data could affect the prescribing of higher risk approved drugs to higher risk demographics of patients. We hope that this effort proves significant, as adverse drug reactions comprise the fourth leading cause of death in the United States, and costs individuals and the industry billions of dollars annually [13][14][15].

Literature Review:

The first reporting systems for adverse drug effects were introduced in the 1960's by the European Molecular Biology Organization (EMBO) [16]. At the time, reports were usually reserved for healthcare professionals and provided no venue for patients to report safety issues or analyze datasets. In Europe, ADR reporting system has only changed within the past decade, as the Netherlands, Denmark, the United Kingdom, and Sweden opened their ADR reporting systems to the public [16]. Within Europe, studies have been conducted which dealt to distinguish between adverse drug effect and adverse drug event along with their classifications [17].

After a long period of development and progress, the drug adverse reaction reporting system has become more sophisticated and mature, allowing much research to be done in this area. Peng [18] and her associates used data mining methods to study adverse reactions described in the FDA Adverse Event

Reporting System (FAERS) and analyzed the distribution of gender and age in the report. They found that among patients with adverse reactions in the digestive system, 40 to 75 years old patients accounted for the majority, and most of the adverse reactions occurred within half a year of medication. Motiur and Rahman [19] also used data from FAERS but focused on analyzing the difference between brand name and generic antiepileptic medications and they found that generic lamotrigine has higher risk to cause adverse reactions. Additionally, Yip et al. proposed that association of HLA and carbamazepine hypersensitivity as an adverse drug reaction demonstrates sensitivity regarding ethnicity and phenotype specificity [20]. But, it was found that most research activities carried out on adverse drug effects focused on drug classification, and effects of drug interactions related to specific adverse reactions, but limited progress has been made regarding generalized demographics who are at increased risk of susceptibility to having an adverse drug reaction, to include determination of age groups, sex, and weight. As such, the authors suggest that categorization and identification of demographics and patient cohorts who report adverse drug effects to FAERS may be of value.

Outside the scope of the FDA's AER system, significant research has been done on global populations with respect to characterization of adverse drug reaction prone population traits. General population characterizations assumed to be at higher risk are pregnant women, women, ethnicity, disease states, polypharmacy, and age-related individuals to include elderly and children [5][21]. Additionally, two of the characterized groups, pregnant women and children, tend to be underrepresented with clinical trial results due to ethical standards. However, while adverse drug reactions are reported for all aforementioned characterizations within the FAERS dataset, this type of analysis has not yet been performed on the FAERS dataset.

Although most researchers believe that FAERS data is very important for studying disease characteristics [22], some scholars still suggest data collected from FAERS has some limitations: there is data to suggest that adverse drug reaction reports are underreported when compared to spontaneous reporting via alternative mediums. The Drug Safety Research Unit in Southampton, UK, determined that the median under-reporting rate of ADRs across 12 nations was 94% (with an interquartile range of 82-98%) [23]. Their findings suggest that there may be significant under reporting in many nations, which can lead to a lack of insight regarding affected demographics and the types and volume of adverse drug reactions for which they may encounter. The reason why this happened is that FAERS uses a voluntary reporting system, and adverse events and medication error reports are usually submitted voluntarily by health care professionals and consumers, so there are a large number of unsubmitted adverse drug events. Unfortunately, this means the FDA cannot collect all reports of all adverse events or medication errors for a drug [24]. In addition, at the time of submission, the FDA does not require proof of causality between adverse events and drugs, and the report does not usually include details of adverse events. Since we do not intend to use this data to calculate the incidence of adverse events or medication errors in the US population, these issues will have no impact on our findings.

Proposed methodology:

Data to be utilized for this project will be sourced from the FDA Adverse Event Reporting System (FAERS), a computerized information database containing event reports, medication error reports, and product quality complaints resulting in adverse events to be submitted to the FDA. This data is utilized by the FDA currently within their post marketing safety surveillance program, which aims to monitor and identify correlations between all approved drugs and reported adverse drug events or medication errors [25].

FAERS data is made available to the general public through multiple ingest mechanisms. For the purpose of conducting this project, FAERS data will be sourced through the use of openFDA which is a

government sponsored project which provides easy access to a wide variety of resources including source code, reports, and relevant government sourced datasets. OpenFDA provides FAERS data, stored in order temporally in JSON format starting in 2004. The FAERS data sourced via openFDA is updated quarterly and is minimally altered from source format in order to fit a standardized JSON structure.

There are limitations to what may be accomplished by reviewing the data contained within the FAERS data. The FDA outlines four major limitations to be taken into consideration when utilizing the dataset [25]. These include:

1. Duplicate or incomplete reports may be present.
2. Drugs suggested as causal factors for adverse reactions are not definitive and cannot be proven on an individual report basis.
3. Report content has not been verified by the FDA.
4. Incidence rates may not be established by the reports due to incomplete data.

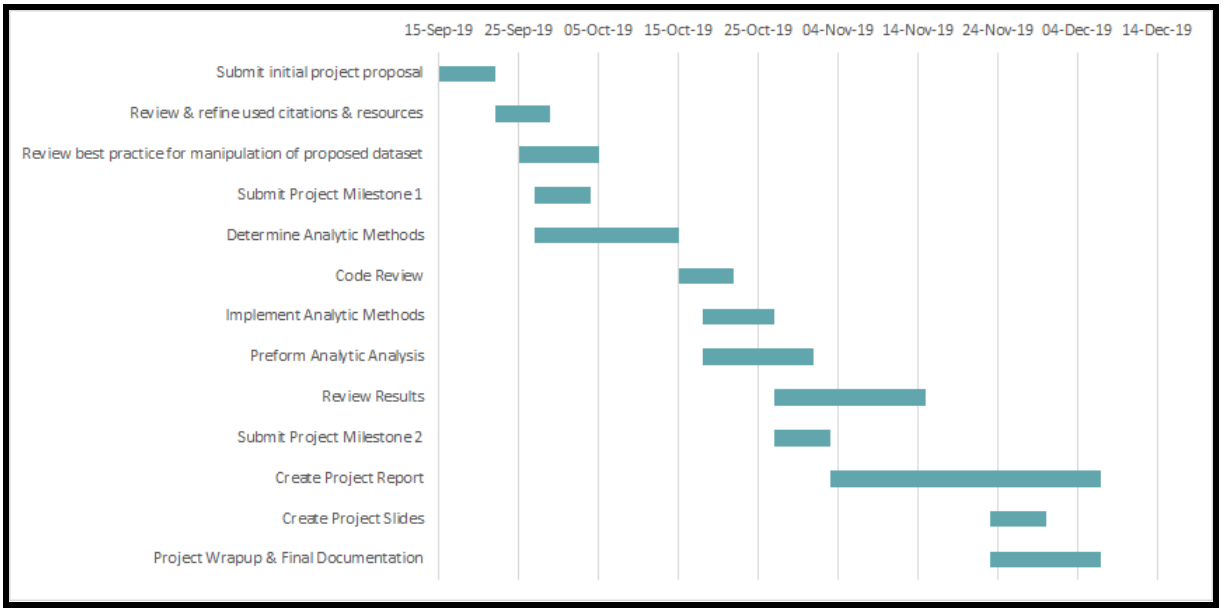
However, while the provided limitations do prevent the analysis of many drug-related venues of interest, this data is still relevant for the determination of characteristics relating to patient demographics and health.

FAERS data can be queried via openFDA using their published API; and, adverse event reports are stored in JSON structure. For the purposes of this research, the FAERS data will be consumed via the openFDA API and stored in a relational database to allow for a definitively static dataset as well as easing data classification and organization. The data utilized for analysis will include all FAERS data reports from inception of the program until the second quarter of 2019. Each FAERS record is comprised of three major sections, consisting of patient information, drugs taken during the event experienced, and adverse reaction recorded by the practitioner; and, each major section is comprised of a variable number of fields relating to data available to the practitioner on duty during the time of adverse reaction event.

To address the objectives outlined within this proposal, the authors will begin by performing standard data wrangling techniques such as data discovery, cleaning, enrichment, and validation. This will be necessary as the format for the adverse reaction reports has changed since their inception in the early 2000's, and not all fields are standardized between years. Once a consistent dataset is formulated, initial characterization of the data will be assessed. Safety signal identification through the use of Signal Detection Algorithms (SDAs) will be evaluated for efficacy and relevance to our primary object, as they have been shown to be useful when determining surrogate measures of statistical association between objects contained in FAERS reports [26]. To characterize cohorts present within the FAERS data, basic descriptive analysis will be conducted to include characterization of available individual metadata which contains values like height, weight, age, and sex. Where possible, investigation into geo-semantic parsing will be conducted to provide an approximate geo-location for patients affected by adverse drug reactions. If successful, this will aid in drawing additional conclusions that may be affected by geographic location.

Project Development Timeline

The timeline below will provide an approximate schedule for execution of project goals, starting with the initial project proposal creation and submission. Each section of the projected timeline may be extended or contracted as necessary to maintain an accurate project roadmap.



References:

[1] Food and Drug Administration, *openFDA Drug Drug Adverse Events Dataset*: U.S. Department of Health and Human Services, 2019. Accessed on: October 5, 2019. Available: https://open.fda.gov/apis/drug/event/download/

[2] Parameswaran Nair N, et al. "Hospitalization in Older Patients Due to Adverse Drug Reactions - the Need for a Prediction Tool." Clinical Interventions in Aging, vol. 11, Dove Medical Press, May 2016, pp. 497–505.

[3] Kongkaew, C., Noyce, P. R., & Ashcroft, D. M. "Hospital Admissions Associated with Adverse Drug Reactions: A Systematic Review of Prospective Observational Studies". Annals of Pharmacotherapy, 42(7–8), (2008): 1017–1025. https://doi.org/10.1345/aph.1L037

[4] Johnson, J., Bootman, J. Drug-related morbidity and mortality: a cost-of-illness model. Arch Intern Med. 1995:155:1949-56

[5] Ahmad, Syed Rizwanuddin,M.D., M.P.H., "Adverse drug event monitoring at the food and drug administration: Your report can make a difference," Journal of General Internal Medicine, vol. 18, (1), pp. 57-60, 2003. Available: https://search-proquest-com.mutex.gmu.edu/docview/923381714?accountid=14541. DOI: http://dx.doi.org.mutex.gmu.edu/10.1046/j.1525-1497.2003.20130.x.

[6] James H. Kim, Anthony R. Scialli. "Thalidomide: The Tragedy of Birth Defects and the Effective Treatment of Disease", Toxicological Sciences, Volume 122, Issue 1, July 2011, 1–6, https://doi.org/10.1093/toxsci/kfr088

[7] "Preventable Adverse Drug Reactions: A Focus on Drug Interactions", U.S. Food and Drug Administration, FDA, Available: https://www.fda.gov/drugs/drug-interactions-labeling/preventable-adverse-drug-reactions-focus-drug-interactions

[8] Braun, Lt et al. "Diagnostic Errors by Medical Students: Results of a Prospective Qualitative Study." *Bmc Medical Education* 17 (2017): n. pag. Web.

[9] Graber ML. The incidence of diagnostic error in medicine. BMJ quality & safety. 2013:bmjqs-2012-001615.

[10] Sankar A, Beattie WS, Wijeysundera DN. How can we identify the high-risk patient?. *Curr Opin Crit Care*. 2015;21(4):328–335. doi:10.1097/MCC.0000000000000216

[11] Bates DW, Miller EB, Cullen DJ, et al. Patient Risk Factors for Adverse Drug Events in Hospitalized Patients. *Arch Intern Med*. 1999;159(21):2553–2560. doi:10.1001/archinte.159.21.2553

[12] van den Bemt, P.M.L.A. et al. "Risk Factors for the Development of Adverse Drug Events in Hospitalized Patients." *Pharmacy World and Science* 22.2 (2000): 62–66. Web.

[13] Lazarou, J., Pomeranz, B.H. & Corey, P.N. "Incidence of adverse drug reactions in hospitalized patients: a meta-analysis of prospective studies". JAMA 279 (1998): 1200–1205

[14] Classen, D.C., Pestotnik, S.L., Evans, R.S., Lloyd, J.F. & Burke, J.P. "Adverse drug events in hospitalized patients. Excess length of stay, extra costs, and attributable mortality". JAMA 277 (1997): 301–306.

[15] Harpaz, R., DuMouchel, W., Shah, N.H., Madigan, D., Ryan, P. & Friedman, C. "Novel data-mining methodologies for adverse drug event discovery and analysis. Clin. Pharmacol". Ther. 91 (2012): 1010–1021.

[16] Weigmann, Katrin. "Consumer reporting of adverse drug reactions: Systems that allow patients to report side effects of the drugs they are taking have yielded valuable information for improving drugs safety and health care." EMBO reports vol. 17,7 (2016): 949-52. doi:10.15252/embr.201642616

[17] Bouvy, J.C., De Bruin, M.L. & Koopmanschap, M.A. "Epidemiology of Adverse Drug Reactions in Europe: A Review of Recent Observational Studies". Drug Safety 38 (2015): 437. https://doi.org/10.1007/s40264-015-0281-0

[18] Yuan Peng, Yan Zhong, Fu Liu, "Detection and Analysis of Engel's Net Safety Signal Based on US Adverse Event Reporting System" Journal of drug epidemiology, vol. 27 (8), pp. 509-517, 2008.

[19] Rahman, Md. Motiur et al. "Comparison of Brand Versus Generic Antiepileptic Drug Adverse Event Reporting Rates in the U.S. Food and Drug Administration Adverse Event Reporting System (FAERS)." Epilepsy Research 135 (2017): 71–78. Web.

[20] V. L. Yip et al, "Genetics of Immune-Mediated Adverse Drug Reactions: a Comprehensive and Clinical Review," Clin. Rev. Allergy Immunol., vol. 48, (2-3), pp. 165-175, 2015. Available: https://search-proquest-com.mutex.gmu.edu/docview/1672261012?accountid=14541. DOI: http://dx.doi.org.mutex.gmu.edu/10.1007/s12016-014-8418-y.

[21] G. Kaufman, "Adverse drug reactions: classification, susceptibility and reporting," Nursing Standard (2014+), vol. 30, (50), pp. 53, 2016. Available: https://search-proquest-com.mutex.gmu.edu/docview/1814729924?accountid=14541. DOI: http://dx.doi.org.mutex.gmu.edu/10.7748/ns.2016.e10214.

[22] H Fang, et al. "Exploring the FDA Adverse Event Reporting System to Generate Hypotheses for Monitoring of Disease Characteristics." Clinical Pharmacology & Therapeutics, vol. 95, no. 5, Nature Publishing Group, Jan. 2014, pp. 496–98, doi:10.1038/clpt.2014.17.

[23] Hazell, L. & Shakir, S.A. "Under-Reporting of Adverse Drug Reactions". Drug-Safety (2006) 29: 385.

[24] Zhang Dandan, "FDA's Adverse Event Reporting System - How much do you know about FAERS?" Center for International Food, Drug, Policy and Law, Shenyang Pharmaceutical University, https://www.sohu.com/a/148778442_803087

[25] "Questions and Answers on FDA's Adverse Event Reporting System (FAERS)." U.S. Food and Drug Administration, FDA, www.fda.gov/drugs/surveillance/questions-and-answers-fdas-adverse-event-reporting-system-faers.

[26] Harpaz, R., DuMouchel, W., LePendu, P., Bauer-Mehren, A., Ryan, P. & Shah, N.H. "Performance of pharmacovigilance signal-detection algorithms for the FDA adverse event reporting system". Clin. Pharmacol. Ther. 93 (2013): 539–546