# Signal Detection in Small Safety Databases

Katie Spain, Victoria Williams
Rho, Chapel Hill, NC

## Overview

On September 29, 2010, the FDA published a new final rule amending the IND safety reporting requirements under 21 CFR Part 312 to improve safety monitoring in clinical trials. Under this new regulation, effective March 28, 2011, all trials involving an investigational new drug (IND) application are held to more stringent analysis and reporting guidelines for adverse events. One specific requirement of the regulation is that sponsors should have a systematic approach in place for safety surveillance of their entire safety database. This signal detection process extends throughout the investigational lifetime of a drug and is designed to determine if the incidence of any adverse events associated with a study drug is higher than their incidence associated with other drugs or placebos. It is intended to aid in detection of safety signals present at low frequencies that may escape detection by looking solely at individual trials. Several methods for analyzing large safety datasets aggregated across multiple studies have been published, but publications regarding signal detection in smaller safety databases are scarce.

The NIAID, NIDDK, and JDRF -funded Immune Tolerance Network (ITN) includes a portfolio of several smaller clinical trials, many of which are under INDs held by DAIT, NIAID or by the investigators. To comply with the new regulations, attempts were made to extend the methods for analyzing large databases for use on much smaller scale analyses of single studies or groups of studies. This poster will focus on the feasibility of extrapolating a variety of methods used on large databases to smaller studies. We will also display graphical tools developed to enhance the evaluation of possible adverse event signals and

## Signal Detection Assumptions

While there are a myriad of methods used to detect safety signals in large databases, these methods are not easily transferable to smaller databases and there are several factors to consider before determining which methods may work best. Below are some of the factors that were taken into consideration in order to narrow down which methods may work best for our small databases:

**Spontaneous Reporting vs Longitudinal Reporting–** Many of the methods researched involved using a spontaneous reporting system (similar to what is used at the FDA), as signal detection was originally performed on these large nationwide databases. These databases normally included information complied from a large number of studies conducted on various drugs, with a large amount of data about a multitude of adverse events (AEs). There are a more limited number of methods used for longitudinal reporting systems, which are the type of databases used for our individual studies. These are limited to one study and only include information on a single therapeutic regimen.

**Accounting for Subjects with multiple events–** Subjects with multiple events of the same type can confound the signal detection process by amplifying the rate of a specific event. It is important to consider how you will account for this when determining whether a specific AE should be considered a signal.

## Methods Considered

After analyzing the methods available for large databases, we focused our efforts on 3 methods, the Proportional Reporting Ratio, analyzing the Relative Risk at the subject level, and using Poisson Regression at the event level.

**Proportional Reporting Ratio (PRR)- Event-Level Analysis**

| | Treatment of Interest | All Other Drugs |
|---|---|---|
| AE of interest | a | b |
| All other AEs | c | d |

**PRR = [a/(a+c)] / [b/(b+d)]**

- Most useful when the number of adverse events are available, but the overall number of subjects is unavailable.
- Primarily used on databases that have adverse events reported from a variety of drugs and that involve spontaneous reporting; subjects with no AEs not accounted for
- Calculates the proportion of specified reactions for a drug of interest and compares with the all other drugs in the database; Utility of comparing across multiple drugs and multiple studies.
- Interpreted as the risk of a certain AE (among all other AEs) compared between treatment groups, as opposed to looking at the risk of having the AE (versus not having the AE) between treatment groups; the higher the PRR, the greater the signal strength. Most references define a potential signal as PRR ≥ 2.

**Relative Risk (RR)- Subject-level Data Analysis**

| | Treatment A (n= A) | Treatment B (n=B) |
|---|---|---|
| Subjects with AE of interest | a | b |
| Subjects without AE of interest | c | d |

**RR = [a/(A)] / [b/(B)]**

- Straightforward application to the signal detection process
- Takes into account subjects without any AEs and allows for use of all data in the database leading to direct computation of incidences.
- Allows for easy interpretation by reviewers with no statistical background. A relative risk (RR) of 2 implies the subjects in treatment group A are 2 times more likely to have the AE of interest than the subjects in treatment group B.
- With small safety databases we are able to use a Fisher's exact test to compare the RRs between treatment groups.
- After careful consideration, this was determined to be the best method for analyzing AE data at the subject level.

**Poisson Regression (PR)- Event-level Data Analysis**

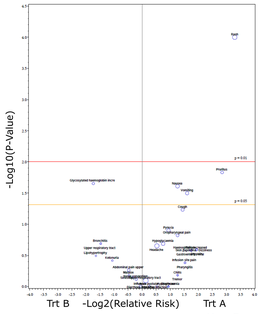
**$Log(\mu) = log(Person\text{-}Years) + \beta_0 + \beta_1 * (Treatment)$**

- Allows for the comparison of event rates between treatment groups while adjusting for the amount of time each subject has been exposed to the treatment of interest.
- Provides a more accurate assessment of the risk to a subject as it will adjust the incidence rate ratio accordingly and differentiate between a subject on treatment 5 years without the AE of interest from a subject with the AE of interest, but only on for 1 day prior to AE onset.

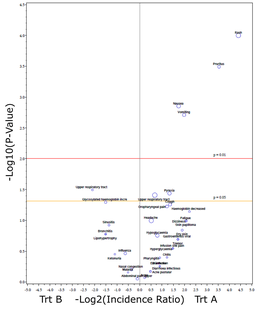
## Signal Detection Analysis

It was determined that the Proportional Reporting Ratio (PRR), would not be an appropriate statistic to analyze because the number of subjects on each treatment regimen is available in the database. Using a combination of the subject-level relative risk estimate, and the incidence rate ratio from the event-level Poisson regression analysis was decided to be the best way to perform signal detection on small databases. While these methods will not replace careful review by the medical monitor, they are meant to aid the reviewer by allowing them to focus in on AEs with the largest difference in risk between the treatment groups. In the volcano plots below, the reviewers can hone in on AEs that appear above the colored lines and are the furthest from the center of the graph.

Subject-Level Comparison (Relative Risks)




Event-Level Comparison (Poisson Regression)




Key Graphical Features

**Y-axis:** Negative $\log_{10}$-transformed p-value.
**Red line–** p= 0.01 level; AE terms above this line represent AEs that are potential safety signals. In this example, reviewers would want to further investigate the preferred terms, Rash, Pruritis, Nausea, and Vomiting.
**Yellow line**– p=0.05 level; AE terms above this line represent AEs that are of concerns, but are not at the level to be considered potential safety signals.
**X-axis -** $Log_2$-transformed relative risk/Incidence Ratio. A Log2(Relative Risk) of 0 represents no difference in risk between the 2 treatment groups, while bubbles to the right indicate a higher risk for subjects in treatment group A (i.e– If Log2(RR)= 1, that is the same as an RR=2), and bubbles to the left indicate a higher risk for subjects in treatment group B. The size of the bubble represents the total number of occurrences of the AE of interest.

**Of Note:** These analyses are performed on the preferred term, which is a coded version of the verbatim term using the Medical Dictionary for Regulatory Activities (MedDRA). Prior to the analysis the adverse event database should be cleaned to ensure that similar verbatim terms are coded to the same preferred term for the most accurate estimation of the relative risk and incidence rate ratio for each adverse event.

Supporting Documentation

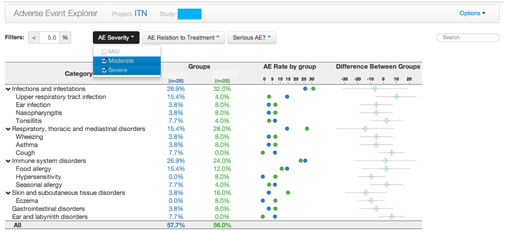
Number of Subjects and Episodes of Adverse Events by System Organ Class and Preferred Term

| System Organ Class<br>Preferred Term | (N=52)<br>No. Pts with >=1 AE | Pct. Pts | No. Events | Event Rate | (N=25)<br>No. Pts with >=1 AE | Pct. Pts | No. Events | Event Rate | Rel. Risk | Subject Level p-value [1] | Inc. Ratio | Event Level p-value [2] |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Gastrointestinal disorders | 43 | | 137 | | 15 | | 23 | | 1.378 | 0.047 | | |
| Nausea | 25 | (48.1) | 40 | 0.404 | 5 | (20.0) | 6 | 0.121 | 2.404 | 0.025 | 3.337 | 0.001 |
| Vomiting | 19 | (36.5) | 34 | 0.319 | 3 | (12.0) | 4 | 0.081 | 3.045 | 0.032 | 3.958 | 0.002 |
| Abdominal pain | 16 | (30.8) | 20 | 0.188 | 0 | | 0 | | | <.001 | | |
| Abdominal pain upper | 9 | (17.3) | 12 | 0.113 | 0 | (24.0) | 6 | 0.121 | 0.721 | 0.545 | 0.931 | 0.887 |
| Constipation | 5 | ( 9.6) | 6 | 0.056 | 0 | | 0 | | | 0.168 | | |
| Diarrhoea | 4 | ( 7.7) | 4 | 0.038 | 1 | ( 4.0) | 1 | 0.020 | 1.923 | 1.000 | 1.863 | 0.556 |
| Skin and subcutaneous tissue disorders | 47 | | 207 | | 10 | | 16 | | 2.260 | <.001 | | |
| Rash | 41 | (78.8) | 91 | 0.854 | 2 | ( 8.0) | 2 | 0.040 | 9.856 | <.001 | 21.168 | <.001 |
| Pruritus | 16 | (28.8) | 28 | 0.286 | 1 | ( 4.0) | 1 | 0.020 | 7.219 | 0.015 | 11.642 | <.001 |
| Skin exfoliation | 10 | (19.2) | 12 | 0.113 | 0 | | 0 | | | 0.020 | | |
| Dry skin | 8 | (15.4) | 8 | 0.075 | 1 | ( 4.0) | 1 | 0.020 | 3.846 | 0.257 | 3.725 | 0.144 |
| Erythema | 8 | (15.4) | 11 | 0.103 | 0 | | 0 | | | 0.048 | | |
| Acne | 5 | ( 9.6) | 5 | 0.047 | 2 | ( 8.0) | 2 | 0.040 | 1.202 | 1.000 | 1.164 | 0.854 |

Along with the volcano plots shown to the left, the reviewers will also be presented the table above. This allows the reviewers to see the underlying data and to analyze the AEs of interest in more detail. AEs highlighted in grey have a p-value less than or equal to 0.01 from either the Fisher's Exact test or the Poisson Regression.

Of Note: Data can also be analyzed by looking at higher level groupings including the High Level Coded Term and the System Organ Class.

## Future Endeavors




In order to further aid in the review of Adverse Event data, Rho has been developing new interactive graphics. These graphics will allow reviewers to filter and sort data in real time while providing both the graphic and the supporting data in the same panel. This effort is currently underway and will be rolled out for ITN studies later this year.

Immune Tolerance Network

Rho® Giving flight to research

NIAID NIDDK National Institute of Diabetes and Digestive and Kidney Diseases JDRF Improving Lives Curing Type 1 Diabetes

This project has been funded with federal funds from the National Institute of Allergy and Infectious Diseases, National Institutes of Health, under contract HHSN272200800029C.
The volcano plots present data from the ITN027AI study.