A Generalized Likelihood Ratio Test for Estimating Drug Adverse Reactions

Yeqian Liu

Department of Mathematical Sciences, Middle Tennessee State University, Murfreesboro, TN, USA.

Abstract
Likelihood ratio test is widely used for detecting adverse reactions (ARs) of single drug in biomedical studies. However, it is difficult to detect adverse reactions of multiple drugs simultaneously. Corresponding to this, we consider a generalized likelihood ratio test procedure to detect adverse reactions simultaneously for all the drugs of the same class. An extensive simulation study is performed to evaluate the proposed test procedure as well as its power and sensitivity. The simulation study suggest that the proposed test procedure seems to work well for practical situations and an illustrative example is also provided.

Keywords: Likelihood ratio test, Adverse reaction, Type I error, Sensitivity

Introduction
Reporting of drug or medical device related adverse reactions (ARs) is usually voluntary. One of the major postmarket safety surveillance databases is the World Health Organization’s (the “WHO”) global pharmacovigilance database, which contains reports of suspected ADRs, so called Individual Case Safety Reports (ICSRs), collected by national drug authorities in over 110 countries and span over more than 100,000 different medicinal products. Clinical reviewers evaluate adverse reactions reports to look for new safety concerns that might be related to a marketed product, or for a manufacturer’s compliance to reporting regulations.

There are several statistical methods available for adverse reactions detection in pharmacovigilance databases. These methods include the reporting odds ratio (ROR, [7]), proportional reporting ratio (PRR, [3]), multi-gamma Poisson shrinker (MGPS, [2,8]), Bayesian confidence propagation neural network (BCPNN, [1]), Bayesian method based on a new Information Component (IC, [6]), simplified Bayes (sB, [5]), among others. A likelihood ratio test-based method ([4]), that assumes the number of adverse reactions follows a Poisson distribution, was developed to identify signals of adverse events for a particular drug or to identify drugs for a particular adverse event. In post-marketing surveillance, the signals of adverse reactions within a drug class, or drug signals for a group of adverse reactions may be of interest to medical reviewers. In this article, we develop a generalized likelihood ratio test to identify adverse reactions that have high reporting rates compared to other adverse reactions associated with all the drugs of the same class or with similar treatment indications. The drug class refers to a set of drugs which have similar chemical structure such as the antibiotics drug class containing Penicillins, Tetracyclines, Cephalosporins, Quinolones, Lincomycins, Macrolides, Sulfonamides, Glycopeptides, Aminoglycosides and Carbapenems, etc. A group of adverse reactions refers to a set of preferred terms such as hepatic failure, alanine aminotransferase abnormal, ascites, blood bilirubin abnormal, cholestatic liver injury, hepatic atrophy, hepatomegaly, Reye’s syndrome, and so on that are all related to hepatocellular injury.

This article is organized as followed. In section 2, a brief overview of World Health Organization’s global pharmacovigilance database is provided. In section 3, we give a brief review of the likelihood ratio test procedure for adverse reactions detection for a single drug, then we propose a generalized likelihood ratio test procedure, namely GLRT, to detect multiple ARs in a drug class. The performance of GLRT is evaluated using simulated datasets in section 4. In section 5, both the LRT and the GLRT are applied to the 2000-2005 and 2005-2010 data from WHO’s...
global database. Section 6 contains some discussion and concluding remarks.

2 World Health Organization’s Pharmacovigilance Database

The WHO’s global pharmacovigilance database consists of the individual reports with demographical information, route of administration, drug/biological information, medical history, treatment indication, therapy start dates, and end dates. For adverse reaction detection, Medical Dictionary for terminology of preferred terms is often used to identify the adverse events, such as Death, Stroke, Myocardial infarction, and so on. There are also verbatim drug names in the file for drug/biological information. In studying the drug-AE association, the generic name of the drug is used, which refers to the unique chemical makeup of a drug.

The WHO’s global pharmacovigilance database includes reports since 1980, however researchers and reviewers are more interested in data from recent years. In this article we focus on cases reported to WHO between 2000 and 2010 for more than 6500 drugs and 14,000 Adverse reactions. For any particular adverse event, the investigators consider all suspect and concomitant drugs.

3 Inference Procedure

3.1 Test Procedure for Adverse Reactions Detection of A single Drug

After summarizing the data files, the WHO pharmacovigilance data can be presented in a tabular form with, say, adverse reactions (ARs) as the row variable and drugs as the column variable (as in Table 1), with \( n_{ij} \) as the cell count for ith AR and jth drug, \( n_j \) as the sum of counts for ith AR (ith row total) and \( n_i \) as the sum of counts for jth drug (jth column total).

We collapse the \( I \times J \) table into multiple \( 3 \times 3 \) tables. For a fixed jth drug, we have I such tables (Table 2), each associated with an AR (\( i=1,\ldots,I \)). We assume that \( n_j \sim \text{Poisson}(\eta_j) \), where \( \eta_j \) is the reporting rate of jth drug for ith AE; and \( n_{ij} \sim \text{Poisson}(\eta_i \eta_j) \) where \( \eta_i \) is the reporting rate of jth drug for other AEs combined excluding ith AR. We also assume \( n_j \) and \( n_i-n_j \) are independent. Since drug j is fixed, unless stated otherwise, we suppressed the notational dependence of \( \eta_j \) and \( \eta_i \) on jth drug. We define the null hypothesis,

\[
H_0: p_j = p_0 = \text{for all ARs in drug } j, \quad p_0 = p_j \text{ for all ARs in drug } i
\]

and the alternative hypothesis (one sided) as

\[
H_0:p_j > q_j \quad \text{for at least one AR in drug } j, \quad p_j > q_j \quad \text{for all ARs in drug } i
\]

Under the null hypothesis, the maximum likelihood estimates (MLEs) for \( p_j \) and \( q_j \) are \( p_j = \frac{n_{ij}}{n_j} \) and \( q_j = \frac{n_{ij}}{n_j} \). Under the two-sided alternative hypothesis (MLEs) for \( p_j \) and \( q_j \) are \( p_j = \frac{n_{ij}}{n_j} \) and \( q_j = \frac{n_{ij}}{n_j} \) for at least one AR in drug j vs. \( H_0 : p_j < q_j \) for at least one AR.

The maximum likelihood ratio test statistic for testing \( H_0 : p_j = q_j \) for all ARs in drug j vs. \( H_0 : p_j > q_j \) for at least one AR, is the maximum likelihood ratio \( \text{MLR} = \text{max} \left( LR_j \right) \), where the maximum is over ARs \( i=1,\ldots,I \). We calculate

\[
LR_j = \frac{L(p_j=q_j)}{L(p_j>q_j)} = \frac{n_{ij}}{n_j} \text{ and define } \text{MLR} = \text{max} \left( LR_j \right)
\]

where \( \text{log} (LR_j) \) is the log-likelihood ratio for ith AR and jth drug. The likelihood ratio test statistic for testing \( H_0 : p_i = q_j \) for all ARs in drug j vs. \( H_0 : p_j > q_j \) for at least one AR, is the maximum likelihood ratio \( \text{MLR} = \text{max} \left( LR_j \right) \), where the maximum is over ARs \( i=1,\ldots,I \). We calculate

\[
LR_j = \frac{n_{ij}}{n_j} \text{ and define } \text{MLR} = \text{max} \left( LR_j \right)
\]

MLR = max(\( LR_j \)) as the test statistic. For computational convenience, we may sometimes work with the log-likelihood ratio \( \text{log} (LR_j) \) which is the monotone function of \( LR_j \).

The distribution of MLR under \( H_0 \) is not analytically tractable and is obtained using Monte Carlo simulation as defined below. First, the number of cases for each AR, for a given drug j, are simulated under \( H_0 \). Under \( H_0 \), since \( n_j \sim \text{Poisson}(\eta_j) \), given the margin totals \( n_{1j},\ldots,n_{IJ} \) are independent \( \text{Poisson}(n_j,\eta_j) \), the joint distribution of \( n_{ij} \) conditioning on \( n_j \) and \( n_{1j},\ldots,n_{IJ} \) is

\[
(n_{ij},n_{1j},\ldots,n_{IJ}) \sim \text{multinomial}(n_j,\frac{n_{ij}}{n_j},\ldots,\frac{n_{IJ}}{n_j})
\]

Table 1. Data Structure.

<table>
<thead>
<tr>
<th>Drug Type</th>
<th>Drug 1</th>
<th>Drug 2</th>
<th>Drug 3</th>
<th>Drug 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>AR 1</td>
<td>AR 11</td>
<td>AR 12</td>
<td>AR 13</td>
<td>AR 14</td>
</tr>
<tr>
<td>AR 2</td>
<td>AR 21</td>
<td>AR 22</td>
<td>AR 23</td>
<td>AR 24</td>
</tr>
<tr>
<td>AR 3</td>
<td>AR 31</td>
<td>AR 32</td>
<td>AR 33</td>
<td>AR 34</td>
</tr>
<tr>
<td>AR 4</td>
<td>AR 41</td>
<td>AR 42</td>
<td>AR 43</td>
<td>AR 44</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Drug 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>AR 51</td>
</tr>
<tr>
<td>AR 52</td>
</tr>
<tr>
<td>AR 53</td>
</tr>
<tr>
<td>AR 54</td>
</tr>
</tbody>
</table>

Table 2. 3×3 tables for jth drug.

<table>
<thead>
<tr>
<th>ith AR</th>
<th>jth Drug</th>
<th>Other Drugs</th>
<th>Marginal Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>n_{ij}</td>
<td>n_{ij}</td>
<td>n_{ij} - n_{ij}</td>
<td>n_{ij} - n_{ij}</td>
</tr>
<tr>
<td>n_{ij}</td>
<td>n_{ij}</td>
<td>n_{ij} - n_{ij}</td>
<td>n_{ij} - n_{ij}</td>
</tr>
<tr>
<td>n_{ij}</td>
<td>n_{ij}</td>
<td>n_{ij} - n_{ij}</td>
<td>n_{ij} - n_{ij}</td>
</tr>
<tr>
<td>n_{ij}</td>
<td>n_{ij}</td>
<td>n_{ij} - n_{ij}</td>
<td>n_{ij} - n_{ij}</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Marginal Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>n_{ij} - n_{ij}</td>
</tr>
<tr>
<td>n_{ij} - n_{ij}</td>
</tr>
</tbody>
</table>
A total of 4999 datasets under \( H_0 \) are simulated from the multinomial distribution, and 500 MLRs are calculated (using 999 from the simulated datasets and one from the observed dataset). The null hypothesis is rejected at the \( \alpha=0.05 \) level if the value of MLR from the observed dataset is greater than the 95th percentile of the 500 MLR values (threshold, \( T_\alpha \)). The corresponding \( p \)-value is then \( 1-R/500 \), where \( R \) is the rank of the observed MLR among all the 500 MLR values. If the \( p \)-value of the observed MLR is less than \( \alpha \) (say, 0.05), then the AR associated with this MLR is the strongest signal among all ARs for the \( j \)th drug under consideration. Having found the strongest signal, we can then move to the second largest \( LR_{ik} \) and so on, and declare them as signals if their \( LR_{ik} \) are greater than \( T_\alpha \) or the corresponding \( p \)-values are less than \( \alpha \).

The likelihood ratio test is shown, analytically and through extensive simulation study, to control type-I error and false discovery rate (FDR) while retaining good power and sensitivity \([4,5]\). In the next section, we generalize the likelihood ratio test procedure to detect all AR signals in a drug class. The methods to detect drug signals for a set of prespecified ARs can be performed in a similar fashion.

### 3.2 Test Procedure for Adverse Reaction Detection of Multiple Drugs

In order to develop a test statistic that can identify adverse reaction signals in a drug class including multiple drugs, we assume that a drug class has \( k \) different drugs (usually \( k \) is a small number), and we assume that for \( k \) th drug the number of reports for \( i \)th AR and all other ARs (excluding \( i \)th AR) still remains a Poisson distribution:

\[
\begin{align*}
    n_{ik} & \sim \text{Poisson}(n_i, p_{ik}) \\
    n_k - n_{ik} & \sim \text{Poisson}((n - n_i)q_{ik})
\end{align*}
\]

where \( p_{ik} \) is the reporting rate of \( k \)th drug for \( i \)th AR, and \( q_{ik} \) is the reporting rate of \( k \)th drug for the other AEs. The null and alternative hypotheses for detecting AE signal in drug \( k \) are, \( H_{ik}: p_{ik} = q_{ik} = p_{0k} \) for all AEs in drug \( k \) versus \( H_{ik}: p_{ik} > q_{ik} \) for at least one AE.

The null and alternative hypotheses for detecting AE signals among this drug class with drugs are,

\[
H_{ik} \cap \bigcap_{k=1}^{K} p_{ik} = q_{ik} = p_{0k}
\]

where \( i=1, \ldots, I \) and \( k=1, \ldots, K \) versus

\[
H_{ik} \cap \bigcup_{k=1}^{K} p_{ik} > q_{ik}
\]

Another way to write the null and alternative hypotheses are \( H_{ik} \cap H_{ik} \) versus \( H_{ik} \cap H_{ik} \). Under the null hypothesis \( H_{ik} \), the MLE for \( p_{0k} \) is \( p_{0k} = \frac{n_k}{n} \) and the expected number of counts for \( i \)th AE and \( k \)th drug is \( \varepsilon_{ik} = \frac{\varepsilon_i}{n} \). Under the two-sided alternative hypothesis (\( \bigcup_{i} \bigcup_{k} p_{ik} \neq q_{ik} \)), the MLEs for \( p_{ik} \) and \( q_{ik} \) are

\[
\begin{align*}
    p_{ik} &= \frac{n_{ik}}{n_k} \text{ and } q_{ik} = \frac{n_k - n_{ik}}{n - n_k}.
\end{align*}
\]

The maximum likelihood under both the null and the two-sided alternative hypotheses are obtained by replacing the parameters with their MLEs in the likelihood functions, leading to the likelihood ratio for \( i \)th AR in \( k \)th drug as:

\[
LR_{ik} = \frac{L_i(p_{ik}, q_{ik})}{L_i(p_{0k})} = \frac{\left(\frac{n_{ik}}{n^n}\right)^{n_{ik}} \left(\frac{n_k - n_{ik}}{n^n}ight)^{n_k - n_{ik}}}{\left(\frac{n_k}{n^n}\right)^{n_k}} \sim \frac{\left(\frac{n_{ik}}{n_k}ight)^{n_{ik}} \left(\frac{n_k - n_{ik}}{n_k}ight)^{n_k - n_{ik}}}{\left(\frac{n_k}{n_k}ight)^{n_k}} = \left(\frac{n_{ik}}{n_k}ight)^{n_{ik}} \left(\frac{n_k - n_{ik}}{n_k}ight)^{n_k - n_{ik}}.
\]

The likelihood ratio test statistic for testing \( H_{0k}: p_{ik} = q_{ik} = p_{0k} \) versus, \( H_{ik}: p_{ik} > q_{ik} \) for at least one AR, is the maximum likelihood ratio MLR = \( \max \{ LR_{ik} \} \), \( i=1, \ldots, I \). The test statistic for testing \( H_{0k}: \bigcup_{i=1}^{I} p_{ik} = q_{ik} = p_{0k} \) versus \( H_{ik}: \bigcup_{i=1}^{I} p_{ik} > q_{ik} \) is

\[
\text{MLR} = \max \{ (LR_{ik} : (p_{ik} > q_{ik})) \}
\]

where \( i=1, \ldots, I \) and \( k=1, \ldots, K \).

Because the distribution of MLR under \( H_0 \) is not analytically tractable, we still use a Monte Carlo simulation to obtain its distribution. For each drug \( k \) in the drug class under \( H_0 \) we generate 500 datasets using

\[
(n_{1k}, n_{2k}, \ldots, n_{Ik}) | n_k \sim \text{multinomial}(n_k, (\frac{n_k}{n}, \ldots, \frac{n_k}{n}))
\]

and compute 500 values of MLR including the one from the real data, for \( k=1, \ldots, K \). This results into \( 500 \times K \) MLR values. The null hypothesis is rejected at \( \alpha=0.05 \) level if the value of MLR from the observed dataset is greater than the \( (1-\alpha) \) th percentile of the \( 500 \times K \) MLR values \( T_\alpha \). After AR associated with the largest \( LR_{ik} \) is identified as signal \( LR_{ik} > T_\alpha \), we move to the AR with the second largest value of \( LR_{ik} \), determine if it is a signal and so on. This way, the generalized likelihood ratio test procedure controls Type-I error. It also controls the false discovery rate (FDR) with \( FDR \leq \alpha \).

### 4 Applications

In the following, we present the results from applying the likelihood ratio test procedure discussed in Section 3 to the \"Monoamine oxidase inhibitors\" (MAOIs). The MAOIs are used to treat several conditions. They include, but are not limited to: depression, generalized anxiety disorder, agitation, obsessive-compulsive disorders (OCD), manic-depressive disorders, childhood enuresis (bedwetting), major depressive disorder, diabetic peripheral neuropathic pain, neuropathic pain, social anxiety disorder, posttraumatic stress disorder (PTSD) etc.

The drug class includes Nardil (phenelzine), Parnate (tranylcypromine), Marplan (isocarboxazid), Emsam (selegiline), etc. We select four MAOIs labeled as MAOI1, MAOI2, MAOI3, MAOI4 and MAOI5 (not in any specific order to mask their names) using the WHO 2000–2005 and 2005–2010 data set.
The purpose of this analysis is to identify the ARs signals (with high disproportionality rates) associated with MAOIs drug class. We apply the likelihood ratio test (LRT) and generalized likelihood ratio test (GLRT) for detecting Adverse Reactions.

The results of MAOIs drug class using both the LRT and GLRT are listed in Table 3. By using the likelihood ratio test procedure to each of the four drugs in the drug class, there are 66, 37, 74, 45 ARs detected for the four MAOIs drugs; while using the generalized likelihood ratio test, there are 61, 32, 68, 39 ARs detected, respectively. Across the four drugs, the GLRT detects less ARs than the LRT. By cross-checking the ARs in the four MAOIs drugs, there are 23 common ARs detected within this drug class. The top ARs are listed in Tables 4 and 5, and postural hypotension, high blood pressure, fainting, abnormal heart rhythm, dizziness, headache, drowsiness are the most strong ARs for this MAOIs class.

### Table 3. Number of signals detected by the LRT and GLRT in Monoamine oxidase inhibitors (MAOIs) drug class.

<table>
<thead>
<tr>
<th>Drug</th>
<th>2000-2005</th>
<th>2005-2010</th>
</tr>
</thead>
<tbody>
<tr>
<td>LRT</td>
<td>GLRT</td>
<td>LRT</td>
</tr>
<tr>
<td>MAO1</td>
<td>66</td>
<td>61</td>
</tr>
<tr>
<td>MAO2</td>
<td>37</td>
<td>32</td>
</tr>
<tr>
<td>MAO3</td>
<td>74</td>
<td>68</td>
</tr>
<tr>
<td>MAO4</td>
<td>45</td>
<td>39</td>
</tr>
</tbody>
</table>

### 5 A Simulation Study
#### 5.1 Data Simulation
We then study the performance of the generalized likelihood ratio test (GLRT) using simulated datasets. We simulate datasets based on the four drugs in the monoamine oxidase inhibitors drug class in WHO’s global pharmacovigilance database.

Under the null hypothesis, the data are simulated from multinomial distribution (3). Under the alternative hypothesis, data are generated as follow,

\[
\begin{align*}
n_{1}, \ldots, n_k | n_x: n_{x1}, \ldots, n_{xk} & \sim \text{multinomial}(n_x, (r_{x1} \times n_x / n), \ldots, (r_{xk} \times n_x / n)) \\
\end{align*}
\]

where \(k=1, \ldots, 5\), and \(r_{x1}, \ldots, r_{xk}\) are the relative reporting rates for \(AE_1, \ldots, AE_i\) in \(k\) drugs with constraints \(0 \leq r_{x1}^{ik} \times r_{xk}^{ik} \times n_{x} / n\) and \(\sum_{i=1}^{l} r_{xk}^{ik} \times r_{xk}^{ik} \times n_{x} / n = 1\).

Relative reporting rates \(r_{xk}^{ik}\) are specified as follows: \(r_{xk}^{ik}\) are assigned a value; higher than 1 for ARs selected as signals and 1 for all other ARs not selected as signals. \(r_{xk}\) can be regarded as baseline risks for drug \(k\), and \(r_{xk}\) can be different from one drug to another.

We evaluate how the relative reporting rate \(r\), the sample size \(n_x\) and the number of signals affect the performance of the GLRT through the following four scenarios:

- **Scenario 1**: one signal is randomly assigned to one drug, and the remaining of other four drugs are free of signals. Without loss of generality, we assign one signal to the drug with the column total as 12000.
- **Scenario 2**: we randomly assign 30 common signals in each drug over the drug class with homogeneous relative reporting rate.
- **Scenario 3**: we randomly assign 30 signals in each drug using homogeneous relative reporting rates \(r\) across the drug class, but signals are not necessarily common between drugs.
- Scenario 4: we take a similar process as Scenario 3, randomly select 30 signals for each drug independently, but we use inhomogeneous $rr$. A rate of $2 \times rr$ is assigned to those AR signals for which $n_i$ (the total number of reports for the ARs) fall between 35,000 and 40,000, a rate of $3 \times rr$ to those AR signals for which $n_i$ fall between 20,000 and 25,000, a rate of $4 \times rr$ to those AR signals for which fall between 15,000 and 20,000, and a rate of $5 \times rr$ to those AR signals for which $n_i$ fall between 6,000 and 12,000. $rr$ is assigned to 1 for those ARs that are not selected as signals.

In each simulation, we generate 1,000 datasets.

5.2 Performance Characteristics Evaluation

The performance of the proposed methods is evaluated by using Power, sensitivity (ST) and false discovery rate (FDR). First, power is defined as:

$$\text{Power} = \frac{\text{Number of times reject } H_0}{L},$$

where $L=1,000$ is the total number of simulations. $H_0$ will be rejected when at least one AR in any one drug (in the drug class) is signal.

The sensitivity of a test is the proportion of positive results that are correctly identified. In our case, sensitivity is defined as:

$$\text{Sensitivity} = \frac{V}{V + S},$$

where $V$ is the number of falsely detected signals, $S$ is the number of correctly detected signals. FDR is defined as:

$$\text{FDR} = \frac{V}{V + S}.$$
If $rr$ is fixed, as the number of selected signals increases, the power increases but the FDR decreases.

### 6 Discussion and Concluding Remarks

In this paper we generalized the likelihood ratio test procedure to detect adverse event for a class of drugs and applied it to the World Health Organization's global pharmacovigilance database. The proposed methods can also be used to detect drug adverse reactions in a group of pre-specified adverse reactions by renaming the row and column variables. One of the advantages of the generalized likelihood ratio test presented here is that the methods can be used to find multiple adverse reactions with both the Type-I error and false discovery rates controlled while retaining good power and sensitivity. We note that the GLRT tends to detect less adverse reactions than the LRT method. This is to be expected, since the threshold in the GLRT of the drug class is greater than or equal to those from each individual drug using the LRT, thus it is more conservative.

The generalized likelihood ratio test procedure provides a useful tool to identify potential adverse reactions in pharmacovigilance database. However, the final discovery of the true adverse reactions should also be based on a thorough review of all available medical records.

### Competing interests

The author declares that he has no competing interests.

### Publication history

Editor: Qiang Shawn Cheng, Southern Illinois University, USA.
Received: 30-Aug-2019 Final Revised: 18-Oct-2019
Accepted: 22-Oct-2019 Published: 24-Oct-2019

### Table 8. The effect of the number of true signals on the performance of the GLRT.

<table>
<thead>
<tr>
<th>Scenario 2.1: Randomly select 10 common signals with homogeneous $rr$</th>
<th>marginal count</th>
<th>$n_1$</th>
<th>$rr$</th>
<th>Power</th>
<th>Sensitivity</th>
<th>FDR</th>
</tr>
</thead>
<tbody>
<tr>
<td>NA</td>
<td>AS 1</td>
<td>0.071</td>
<td>NA</td>
<td>0.068</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NA</td>
<td>AS 2</td>
<td>0.236</td>
<td>0.076</td>
<td>0.059</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NA</td>
<td>AS 3</td>
<td>0.596</td>
<td>0.14081</td>
<td>0.021</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NA</td>
<td>AS 5</td>
<td>1</td>
<td>0.45608</td>
<td>0.006</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NA</td>
<td>AS 7</td>
<td>1</td>
<td>0.09961</td>
<td>0.0030</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NA</td>
<td>AS 9</td>
<td>1</td>
<td>0.88292</td>
<td>0.0025</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| Scenario 2.2: Randomly select 20 common signals with homogeneous $rr$ |
|---|---|---|---|---|---|---|
| NA | AS 1 | 0.059 | NA | 0.058 |
| NA | AS 2 | 0.535 | 0.01639 | 0.005 |
| NA | AS 3 | 1 | 0.1601 | 0.0048 |
| NA | AS 5 | 1 | 0.5631 | 0.0029 |
| NA | AS 7 | 1 | 0.72354 | 0.00145 |
| NA | AS 9 | 1 | 0.85018 | 0.00137 |

| Scenario 2.3: Randomly select 30 common signals with homogeneous $rr$ |
|---|---|---|---|---|---|---|
| NA | AS 1 | 0.063 | NA | 0.063 |
| NA | AS 2 | 0.837 | 0.01416 | 0.0149 |
| NA | AS 3 | 1 | 0.1629 | 0.0052 |
| NA | AS 5 | 1 | 0.5048 | 0.0012 |
| NA | AS 7 | 1 | 0.7333 | 0.0009 |
| NA | AS 9 | 1 | 0.9461 | 0.0008 |

AS: actual sample size; NA: not applicable.

### References


### Citation: