## *An Alternate Method to Find the Confidence Interval for the Difference between Two Proportions of Rare Event*

Ruji Yao, Amarjot Kaur, Qing Li

*Merck & Co., Inc., Rahway, NJ, USA*

In a clinical trial, it is common to have an adverse event summary table to show the difference of event rates and their 95% confidence intervals (CI). The Miettinen, O. and Nurminen, M (M&N) [1] method is often used for this summary. When dealing with a difference of rare events, one possible concern is the normality assumption. In this presentation, we propose a new method for obtaining the CI, which is sometimes more suitable for computing the difference of two proportions of rare event. First, we will use the formula of Exact (Clopper-Pearson) [2] Confidence Limits to define the posterior distributions on observed data. This help us to avoid the need to pick up a Beta (a,b) prior in a Bayesian method. We then calculate the confidence interval directly from these posterior distributions without making a normality assumption for the critical values. Most importantly, our method only makes use of the binomial assumption on the observed data. Compared with the M&N method, we find that our method has very similar results when the differences of rates are in the normal range, but a much narrower widths for the CI when considering differences between rare events.

## Introduction and Issue

First, let us examine an actual AE summary table, which has many CIs of differences of rare events (based on the M&N method).

Analysis of Adverse Event Summary

| | Active | Placebo | Difference in % vs. Placebo† |
|---|---|---|---|
| | n | n | Estimate (95% CI)† |
| Subjects in population | 37 | 34 | |
| with one or more adverse events | 32 | 27 | 7.1 (-11.0, 25.6) |
| with drug-related‡ adverse events | 3 | 2 | 2.2 (-12.3, 16.5) |
| with serious adverse events | 11 | 8 | 6.2 (-14.8, 26.5) |
| with serious drug-related adverse events | 1 | 1 | -0.2 (-12.7, 11.4) |
| who died | 0 | 0 | 0.0 (-10.3, 9.5) |
| discontinued§ due to an adverse event | 2 | 0 | 5.4 (-5.1, 17.8) |
| discontinued due to a serious adverse event | 1 | 0 | 2.7 (-7.7, 13.9) |
| †Based on Miettinen & Nurminen method | | | |

**Idea:** With above examples, we would like to use the observed proportion $\hat{p}$ to get an estimated posterior distribution of $p$. With posterior distributions for both active $p_a$ and placebo $p_p$ group, we can obtain the estimated distribution of differences $(p_a - p_a)$ of two proportions or any function of them and construct a confidence interval without making a normality assumption.

To avoid selecting a pair (α,β) for the Beta prior, used in the Bayesian method to derive a posterior on an observe proportion $\hat{p} = n_1/n$ , we would like to use the formula for the exact Clopper-Pearson confidence limits to define a posterior distribution directly on the observe proportion $\hat{p}$ .

***Exact (Clopper-Pearson) Confidence Limits*** *confidence limits for the binomial proportion are constructed by inverting the equal-tailed test based on the binomial distribution. This method is attributed to Clopper and Pearson (1934). The exact confidence limits pL and pU*

$$pL = (1 + \frac{n - n_1 + 1}{n_1\, F(\alpha/2, 2(n_1 + 1), 2(n - n_1 + 1)})^{-1}$$

$$pU = (1 + \frac{n - n_1 + 1}{(n_1 + 1)F(1 - \alpha/2, 2n_1, 2(n - n_1)})^{-1}$$

*The lower confidence limit equals 0 when event $n_1$ = 0, and the upper confidence limit equals 1 when event $n_1$ = n, where $F(\alpha/2, b, c)$ is the $\alpha/2$ percentile of the F distribution with b and c degrees of freedom.*

*If $\alpha/2$= 0.025 with fixed $n_1$ and n, $pL(0.025)$is a proportion p that has 2.5% probability to observe $n_1$ or above.*

There is a natural way to define a CDF from any formula of confidence limit, based on the observed data. Using $pU$ as an example, with fixed $n_1$ $and$ $n$, it is a function of $(1-\alpha/2)$ and the corresponding CDF can be defined as $\{pU(\mathrm{X}), \mathrm{X}\}$ for X in (0,1) as a posterior distribution.

$$pU = (1 + \frac{n - n_1 + 1}{(n_1 + 1)F(1-\alpha/2, 2n_1, 2(n-n_1)}\,)^{-1}$$

We choose 100 points to illustrate the steps of how to define a discrete posterior distribution, but for any number of points, the procedure is the same.

Steps to define a discrete posterior distribution on the observed $\hat{p} = n_1/n$

1. Define a 100 point sequence of {a}={0.005, 0.015,…,0.995}.

2. Plug each value of {a} into the formula for $pL$ to replace α/2 to get the sequence $pL(\{a\})$.

3. This defines a discrete posterior distribution on $\hat{p} = n_1/n$ . One percent probability is automatically assigned to each value in the sequence of $pL(\{a\})$.

4. Plug each value in {a} into the formula for $pU$ to replace (1- α/2) to get the sequence $pU(\{a\})$.

5. This defines another discrete posterior distribution on $\hat{p} = n_1/n$ . One percent probability is again automatically assigned to each value in the sequence of $pU(\{a\})$.

6. Finally we define the discrete posterior distribution of 200 points on $\hat{p} = n_1/n$ as

$$p = \{pL(\{a\}), pU(\{a\})\}$$

Each point has the same probability of 0.5% for this weighted discrete posterior distribution.

7. if $n_1 = 0$ then $p = pU(\{a\})$

8. if $n_1 = n$ then $p = pL(\{a\})$

**Compare New Method with M&N method**

Analysis of Adverse Event Summary Results

| Width of alternative way over width of M&N (ratio in %) | Active | Placebo | Difference in % vs. Placebo† | Alternative Method |
|---|---|---|---|---|
| | n | n | (95% CI)† | (95% CI) |
| Group total | 37 | 34 | | |
| 97.5% | 32 | 27 | 7.1 (-11.0, 25.6) | (-10.7, 25.0) |
| 92.3% | 3 | 2 | 2.2 (-12.3, 16.5) | (-11.2, 15.4) |
| 99.0% | 11 | 8 | 6.2 (-14.8, 26.5) | (-14.6, 26.3) |
| 85.9% | 1 | 1 | -0.2 (-12.7, 11.4) | (-10.9, 9.8) |
| 80.1% | 0 | 0 | 0.0 (-10.3, 9.5) | (-8.4, 7.6) |
| **89.5%** | **2** | **0** | **5.4 (-5.1, 17.8)** | **(-5.8, 14.7)** |
| 84.3% | 1 | 0 | 2.7 (-7.7, 13.9) | (-7.6, 10.6) |
| 100.2% | 25/50 | 25/50 | (-19.32, 19.32) | (-19.36, 19.36) |
| †Based on Miettinen & Nurminen method | | | | |

Example: How to get the confidence limits of (2/37 – 0/34):

Step 1. Get
$p = \{pL(\{\text{a}\}), pU(\{\text{a}\})\}$ for $n_1 = 2$ and $n = 37$, a K points posterior distribution. Here K=2,000.

Step 2. Get
$p = \{pU(\{\text{a}\})\}$ for $n_1 = 0$ and $n = 34$, a M points posterior distribution. Here M=1,000. And K*M=2,000,000.

Step 3. From K*M pairs data, we get K*M values of differences.

Step 4. $(0.025*K*M)^{th}$ and $(0.975*K*M)^{th}$ values are -5.8 and 14.7.

It takes about 1.8 seconds to get the confidence interval.

**Compare with M&N method (only for upper boundary)**

Width are narrower for rare events.

| Width of alternative way over width of M&N (ratio in %) | A | P | Difference in % vs. Placebo† | Alternative Method |
|---|---|---|---|---|
| | n | n | (95% CI)† | (95% CI) |
| size | 50 | 50 | Upper bound | Upper bound |
| 100.2% | 25 | 25 | 0.19324 | 0.19359 |
| 100.0% | 20 | 20 | 0.18988 | 0.18997 |
| 99.6% | 15 | 15 | 0.17943 | 0.17871 |
| 98.5% | 10 | 10 | 0.16051 | 0.15816 |
| 95.6% | 5 | 5 | 0.12945 | 0.12381 |
| 94.4% | 4 | 4 | 0.12102 | 0.11427 |
| 92.6% | 3 | 3 | 0.11152 | 0.10332 |
| 89.9% | 2 | 2 | 0.10060 | 0.09043 |
| 84.9% | 1 | 1 | 0.08779 | 0.07452 |
| 80.0% | 0 | 0 | 0.07201 | 0.05758 |
| †Based on Miettinen & Nurminen method | | | | |

## Conclusion:

- We proposed a new method to construct a confidence interval for the difference of two proportions of rare events.
- It uses the formula of exact Clopper-Pearson confidence limits to define discrete posterior distributions for both observed proportions and with these two posterior distributions, a confidence interval of their difference can easily be constructed.
- Compared to the conventional M&N method, our approach does not require a normality assumption to get critical values of the confidence limits; compared to the Bayesian method, it does not need to select a prior distribution.
- As an example, we applied our method to the data in a real adverse event summary table and the result showed that it exhibits a much narrower width of confidence intervals.

## References:

[1] Miettinen, O. and Nurminen, M. (1985), “Comparative Analysis of Two Rates,” Statistics in Medicine, 4, 213–226.
[2] Clopper, C. J. and Pearson, E. S. (1934), “The Use of Confidence or Fiducial Limits Illustrated in the Case of the Binomial,” Biometrika, 26, 404–413.