Something out of Nothing?

The Influence of 0-0 Studies in Drug Safety Analysis

Zhaohu(Jonathan) Fan Yuejie Chen Nanhua Zhang Dungang Liu

University of CINCINNATI CARL H. LINDNER COLLEGE OF BUSINESS

Cincinnati Children's changing the outcome together

A Motivating Example

How can we perform an analysis on the data set below given we have 50% double-zero studies?

	Off-pump		On-pump	
Study	Strokes	Total Patients	Strokes	Total Patients
Raja 2003	3	150	4	150
PRAGUE-4 2004	0	208	2	192
Legare 2004	2	150	0	150
Lingaas 2004	0	60	2	60
JOCRI 2005	0	81	1	86
Niranjan 2006	1	40	1	40
Motallebzadeh 2006	1	108	5	104
PROMISS 2010	0	73	0	74
Matata 2000	0	10	0	10
Penttila 2001	0	11	0	11
Caputo 2002	0	20	0	20
Zamvar 2002	0	30	0	30
Gasz 2005	0	10	0	20
Ascione 2005	0	10	0	10
Michaux 2006	0	25	0	25
Ascione 2006	0	20	0	20
Tatoulis 2006	0	50	0	50
Ozkara 2007	0	22	0	22
Rasmussen 2007	0	18	0	17
Mandak 2008	0	20	0	20
Vural 1995	0	25	0	25
Gulielmos 1999	0	20	0	20
Czerny 2000	0	15	0	15
Kochamba 2000	0	29	0	29
Wandschneider 2000	0	52	0	67
Czerny 2001	0	40	0	40
Vedin 2003	0	33	0	37
Velissaris 2003	0	27	0	27
Parolari 2003	0	11	0	14
Gasz 2004	0	10	0	10
Synnergren 2004	0	26	0	26
Blacher 2005	0	13	0	15
Rachwalik 2006	0	21	0	21
Malik 2006	0	25	0	25
Cavalca 2006	0	25	0	25
...	...	...	...	...
Gnenc 2006	0	30	0	12
Rainio 2007	0	10	0	10
Kunes 2007	0	17	0	17
Parolari 2007	0	14	0	15
Formica 2009	0	30	0	30
Modine 2010	0	35	0	36

Figure 1: Møller et al. (2012)'s 60 independent studies comparing the off-pump and onpump methods used in coronary artery bypass grafting with regard to the occurrence of postoperative strokes

Background

Given K studies, we assume that in the i-th study, the number of rare events Y_{ic} (control), and Y_{it} (treatment), follow binomial distributions

$$Y_{ic} \sim \text{Binomial}(n_{ic}, p_{ic}),\ Y_{it} \sim \text{Binomial}(n_{it}, p_{it}),\quad i = 1, \cdots, K. \tag{1}$$

The goal is to compare the probability of control group p_{ic} with the probability of treatment group p_{it} to see if there is any difference. To gauge the difference, we consider odds ratios, $\theta_i = \frac{p_{it}}{1-p_{it}} / \frac{p_{ic}}{1-p_{ic}}$. Equivalently, we have a log odds ratio $\delta_i = \log(\theta_i) = \log(\frac{p_{it}}{1-p_{it}}) - \log(\frac{p_{ic}}{1-p_{ic}}) = logit(p_{it}) - logit(p_{ic}) = \mu_{it} - \mu_{ic}$. Thus, we rewrite the binomial model as follows:

$$logit(p_{ic}) = \mu_{ic},\ logit(p_{it}) = \mu_{ic} + \delta_i,\ i = 1, \cdots, K. \tag{2}$$

We assume that the baseline effects μ_{ic} are random-effects. Specifically, the baseline effects vary, and are drawn from a normal distribution $N(a, b^2)$. The treatment effects are assumed to be fixed, namely, the treatment effects are identical across all the studies $\delta_i = \delta$. It is referred to as fixed-effects binomial model. Hence, we reformulate the fixed-effects binomial model as follows:

$$\begin{gathered} logit(p_{ic}) = \mu_{ic},\ logit(p_{it}) = \mu_{ic} + \delta,\ i = 1, \cdots, K, \\ \mu_{ic} \sim N(a, b^2), \delta \sim N[-,-], \\ a \sim N[-,-], b^2 \sim IG[-,-], \end{gathered} \tag{3}$$

where $[-,-]$ denotes a prior distribution to be specified. In the fixed-effects binomial model, the variable of interest is δ in equation (3).

Empirical Study

The value of p_{max} controls the upper bound of baseline probabilities p_{ic} in Table 1. This means that 99.7% of the baseline probability is contained within the p_{max}. Therefore, each such way (p_{max}=1%,0.5% and 0.1%) is in a rare-event (sparse) setting.

Table 1: Summary statistics of baseline probabilities p_{ic}

p_{max}	Mean	Standard Deviation	95% Quantile	99% Quantile
1.00%	0.51%	0.12%	0.73%	0.86%
0.50%	0.26%	0.06%	0.37%	0.43%
0.10%	0.05%	0.01%	0.07%	0.09%

Setting: unbalanced sample size, $n_{ic} > n_{it}$

Table 2: Power for rejecting the null hypothesis H_0: OR=1

OR	Type I error	1.2	1.4	1.6	1.8	2.0
	p_{max}=0.5%					
Partial analysis	10.10%	4.70%	**15.20%**	43.80%	73.90%	90.90%
Full analysis	4.80%	9.80%	**33.50%**	66.20%	88.20%	97.20%
Average number of 0-0 studies	104	100	97	93	90	86
% of 0-0 studies	58%	56%	54%	52%	50%	48%
	p_{max}=1%					
Partial analysis	11.30%	10.30%	**46.90%**	86.40%	97.80%	99.99%
Full analysis	6.00%	21.90%	**67.60%**	95.00%	99.60%	100.00%
Average number of 0-0 studies	65	61	57	54	50	48
% of 0-0 studies	36%	34%	32%	30%	28%	27%
OR	Type I error		2.0	3.0	4.0	
	p_{max}=0.1%					
Partial analysis	9.20%		10.40%	**53.30%**	92.90%	
Full analysis	5.70%		20.90%	**72.60%**	97.00%	
Average number of 0-0 studies	160		152	145	139	
% of 0-0 studies	89%		84%	81%	77%	

Data Example

The goal of this re-analysis is to assess the following assumption: **a double-zero studies in meta-analyses contain negligible or none information for statistical inference.** If the assumption is right, we should not see not any measurable change in the statistical inference. The initial re-analysis of the meta-data shows evidence effects in Kuss (2015). We re-analyze the meta-data by a fixed-effects binomial model. Hence, we consider the following steps in our analysis. More specifically, we first compare inclusion/exclusion of double-zero studies from our approach. Then we scale up the sample size of double-zero studies by 2, 3, 4 and 5.

This reflects that double-zero studies contains information for inference.To this end, the difference between the Kuss (2015) and our analysis arise from two sources: 1) model itself and 2) the sample size of double-zero studies.

Table 3: Re-analysis of Møller et al. (2012)-I

	Estimate		Credible interval	
Scale factor	w/o 0-0s data	Full data	w/o 0-0s data	Full data
1	0.706	0.705	[0.476, 0.947]	[0.469, 0.946]
2	0.707	0.757	[0.467, 0.941]	[0.500, 1.010]
3	0.707	0.776	[0.476, 0.945]	[0.517, 1.044]
4	0.706	0.784	[0.470, 0.939]	[0.524, 1.056]
5	0.706	0.789	[0.473, 0.944]	[0.514, 1.053]

Conclusions

We conclude that double-zero studies contain meaningful information for the statistical inference in meta-analyses. Excluding double-zero studies can mislead inference about odds ratio. Inclusion of double-zero studies can have the following advantages:

1. **type I error rate can be moved toward nominal 5% significance level**.
2. **the testing power can be significantly. increased**.
3. **bias can be decreased**.

Contact Information

- Zhaohu(Jonathan) Fan, PhD candidate
- fanzh@mail.uc.edu