Study of publication bias in meta-analysis using trim and fill method

Maryam Talebi

Department of statistics, shahid chamran university, ahvaz, iran

Corresponding author email: maryam _talebi2090@yahoo.com

ABSTRACT: Meta-analysis includes statistical methods in combining the results of several studies. However, it might be affected by publication bias. Publication bias is the preferred publication of statistically significant results over insignificant ones. Therefore, if the source of meta-analysis studies are published scientific articles, they might face publication bias. Trim and fill is a method for correcting publication bias in meta-analysis. In this article we analyze thirteen studies on cisapride, related to Hartung and Knapp, 2001, and correct their estimated meta-analysis using trim and fill method.

Keywords: meta-analysis, publication bias, trim and fill method

INTRODUCTION

Scientific research and studies are increasing every day, and this increase includes studies with similar topics, so that the same topics are worked on by several centers or different people. The results of these studies might be inconsistent. It may happen that some studies support one assumption, while others support the opposite. Thus, using statistical models and methods is necessary in order to combine such studies. For example, if the results of studies on the effect of a specific diet on weight loss in different societies are collected, by combining the results we will certainly arrive at the real effect of that diet in weight loss. Meta-analysis is the application of statistical methods in combining the reported results of several studies on a common topic. The first methods for combining statistical tests of different studies were introduced by Tippett (1931), and not long after that Fisher (1932) was attracted too. The word meta-analysis was first used by Glass (1976) in psychology. In the second part of this article the parametric meta-analysis methods are examined. In the third part, publication bias and its detection methods are studied. In the fourth part, publication bias correction methods are dealt with. And finally, the application of these methods is shown in an example of a meta-analysis.

Parametric meta-analysis methods

Although different methods have been used for meta-analysis up to now, the one which is very common these days is combining various studies based on published articles in scientific journals which is only by combining presented statistics in articles and does not need primary data. Doing this is made possible by providing estimate of treatment effect and variance estimator and using proper statistical models, then hypotheses are examined and if enough, are combined to make a single hypotheses; this combination of the results has higher statistical power. Meta-analysis studies which aim to estimate population parameters are more noteworthy because of providing more complete information. Therefore, finding methods through which we can combine the results of such studies to achieve more accurate estimations is inevitable. These methods are usually formed based on effect size. Most meta-analysis are about studies in which two groups, an experimental group and a control group, are compared in the form of a set of selected response variables; and the aim of such studies is estimation and inference about the difference between the two groups. In order to do this, it is necessary to select the relevant criterion for the size of treatment effects. There are different types of effect sizes; however, here we limit the discussion to the effect size based on binary data (2×2 tables). Such tables are used when the response variable is binary; win or lose. Outcomes such as death and survival, healing and not healing, infection and no infection, difference between a
new and an old treatment are of this type, and include relative risk, odds ratio, and probability difference (Hartung et al., 2008).

One of the topics of discussion in meta-analysis is the combination of the results for a group of studies based on two models; fixed effects and random effects models. In fixed effects model, the variation between results is only due to random error (within-study variance); fixedness of the treatment effects in all studies is called homogeneity. Suppose we have K independent studies, and \( \tau_i \) is estimation for the \( \theta_i \) effect size, estimation variance of \( \tau_i \) will be \( \hat{\sigma}^2(T_i) \) and \( T_i \) is from the i-th study or population whose sample size is \( n_i \). In large sample sizes, \( T_i \) follows approximate normality in distribution. To test the homogeneity hypothesis \( H_0 : \theta_1 = ... = \theta_k = \theta \), we can use asymptotic chi-square distribution (Borenstein et al., 2009):

\[
Q = \sum_{i=1}^{K} \frac{(T_i - \bar{\theta})^2}{\sigma^2(T_i)} = \sum_{i=1}^{K} \frac{T_i^2}{\sigma^2(T_i)} - \left[ \sum_{i=1}^{K} \frac{T_i}{\sigma^2(T_i)} \right]^2 \sum_{i=1}^{K} \frac{1}{\sigma^2(T_i)}
\]

Homogeneity hypothesis is rejected if \( \chi^2 > \chi^2_{k-1, \alpha} \).

Random effects model also considers presence of between study variation for treatment effect in measurement and estimation of combine effect size. Thus, in the framework of random effects each study has its own population, and random effects of between study parameter variations should be added to the previous source of variation, and the various results between studies should be examined. So it can be said that the results from this model are more generalizable than fixed effects model. This randomness of treatment effect between studies is called heterogeneity. In this model, each study’s effect is shown with \( \theta_i \) and the sufficient statistic in the i-th study is \( T_i \). Effect size in each study is a random variable with normal distribution; in other words, if we have k studies, treatment effects in these k studies, i.e. \( \theta_1, \theta_2, ..., \theta_k \), are independent values from the following distribution: \( \theta_i \sim N(\theta, \tau^2) \) in which \( \tau^2 \) is between-study variance estimated using either method of moments known as DerSimonian and Laird (1986) or method of maximum likelihood (Jackson et al., 2010).

The main purpose is estimating \( \theta \) using \( T_i \). As a result, we have \( T_i \sim N(\theta, \sigma_i^2 + \tau^2) \). We show the reversed sum of variances like this: \( \omega_i = \frac{1}{\sigma_i^2 + \tau^2} \); and the maximum likelihood estimation for \( \theta \) is shown like the following:

\[
\hat{\theta} = \frac{\sum_i w_i T_i}{\sum_i w_i} \quad ; \text{and its variance equals: } Var(\hat{\theta}) = \frac{1}{\sum_i w_i}
\]

**Publication bias**

One of the issues in meta-analysis is that the evidences in studies are those cases which have been published. Published cases are bias samples from all performed studies, because published studies are those which are statistically significant, or in other words have confirmed a new subject. This is what we call publication bias. Thus, failing to consider publication bias causes meta-analyses to be imperfect (Whitehead, 2002).

**Egger’s method in detecting publication bias**

Among important methods of testing for publication bias is correlation tests. A simple formal test can be built by a test of correlation between estimated effects and their variances (usually their standard errors) so as to examine the fact whether publication bias tends to induce a correlation between the two factors or not? Egger et al. (1997) presented a formal test for publication bias based on linear regression analysis. In this regression, \( \frac{T}{\hat{\sigma}(T)} \) is considered as the dependent variable and \( \frac{1}{\hat{\sigma}(T)} \) as the independent variable. Consider the model like the following in which k is the number of studies:

\[
\chi^2 = \left( \frac{T}{\hat{\sigma}(T)} \right) - \left( \frac{1}{\hat{\sigma}(T)} \right) \cdot \left( \frac{1}{\sum \frac{1}{\hat{\sigma}^2(T_i)}} \right)
\]

\[
\chi^2 = \sum \frac{(T_i - \bar{\theta})^2}{\sigma^2(T_i)} - \left[ \sum \frac{T_i}{\sigma^2(T_i)} \right]^2 \sum \frac{1}{\sigma^2(T_i)}
\]

Homogeneity hypothesis is rejected if \( \chi^2 > \chi^2_{k-1, \alpha} \).
Errors are represented as \( e_i \), which are from normal distribution with the average of zero and \( \sigma^2 \) as variance. Consider \( H_0 : \alpha = 0 \) and \( H_1 : \alpha \neq 0 \) as opposite hypotheses. Test statistic is \( \hat{\alpha}/\hat{\sigma}(\hat{\alpha}) \), which follows distribution of \( t \) with \( k-2 \) degree of freedom. Test significance shows the existence of publication bias. Slope of \( \beta \) shows the size and direction of the treatment effect (Egger et al., 1997).

1.1 Begg’s method in detecting publication bias

Another formal test is rank correlation test after standardizing effect sizes, which was presented by Begg-Mazumdar (1994) and have used rank correlation based on Kendall’s tau. They denote an estimate of parameter \( \theta_i \) from study \( i (i = 1, 2, \ldots, k) \) by \( t_i \), and its variances by \( v_i \) and they make the conventional assumption that the variances are known, although in reality these are estimated from the data in each study. The first stage in conducting this test is to stabilize the variances by deriving standardized effect sizes \( t_i^* \), where in these equations \( v_i^* \) is the variance of \( t_i - \bar{t} \).

\[
\begin{align*}
    T_i^* &= (t_i - \bar{t})/(v_i^*)^{1/2} \\
    \bar{t} &= (\sum v_i^{-1} t_i) / \sum v_i^{-1} \\
    v_i^* &= v_i - (\sum v_i)^{-1} 
\end{align*}
\]

This test is based on counting the number of pairs of studies ranked in the same order by two factors (i.e., \( v, t^* \)). If \( x \) presents the number of pairs ranked in the same order and \( y \) the number of pairs in which the order is reversed, the normalized test statistic is as follows (Begg and Mazumdar, 1994):

\[
z = (x - y) / [k(k - 1)(2k + 5)/18]^{1/2}
\]

Funnel plot

A common method to detect publication bias is using a graphic plot of treatment effect against standard error (sometimes based on sample size or precision), for each of the initial studies in a meta-analysis. This plot is known as the funnel plot. The reason for this name is that studies of a smaller volume, due to higher degree of random changes, have a wider distribution of results than studies of a greater volume. This can cause asymmetry in funnel. Therefore, an asymmetrical funnel plot in meta-analysis can suggest publication bias (Whitehead, 2002).

Publication bias correction methods (trim and fill method)

Duval and Tweedie (2000) have introduced a method which corrects meta-analysis parameters according to the issue of publication bias. Firstly, the remote and asymmetric part of funnel is removed after estimate of the number of studies in this part (trimming). Then, from the symmetric remaining part is used for estimating of the true center of the funnel. Finally, the removed studies and their missing counterparts are replaced around the center (filling). In the end, estimated true mean and the variance related to it are achieved, based on the completed funnel plot. The ‘trim and fill’ method is a non-parametric approach that makes strong assumptions about funnel plot asymmetry (Duval and Tweedie 2000a, 2000b).

In the meta-analysis assume that \( k \) single studies which evaluate a similar problem exist. For each \( j = 1, \ldots, k \), study \( j \) produces an effect size of \( t_j \) which estimates \( \theta \); we also have an within-study variance estimate of \( \sigma_j^2 \). Duval and Tweedie assume that in addition to \( k \) observed studies, there are other \( k_o \) studies which have not been observed because of publication bias. The amount of \( k_o \) and the effect sizes taken from these \( k_o \) studies should be reflected in the final analysis of the meta-analysis. In this method uses ranks of absolute values of observed effect sizes and the signs of those effect sizes around \( \theta \). The observed values of \( T_i - \theta \) are shown as
$X_i$, and $X_h$ in the most negative $X_j$. Now if the ranks of observed values $[X_j]$ are shown with $r_j^+$, $r_h^+$ will be the absolute value of the most negative rank of $X_h$. And we will have: $\gamma^+ = k - r_h^+.$

Test statistic of Wilcoxon's rank for $k$ observed values is shown as $T_k = \sum_{X_i > 0} r_j^+$. According to these symbols, Duval and Tweedie have introduced the following three estimators for $k_o$ (Duval and Tweedie, 2000a, 2000b):

$$R_o = \gamma^+ - 1$$

$$L_o = \{4T_k - k(k + 1)\}/(2k - 1)$$

$$Q_o = k - 1/2 - \sqrt{2k^2 - 4T_k + 1}/4$$

Duval and Tweedie have calculated their expected values and variances in the binomial model. They have rounded any one of these estimators to the nearest non-negative integer and have shown the resulting estimates with $L_o^+, R_o^+, Q_o^+$ respectively. Based on simulation and work on original data, it was clear $L_o^+$ work more robust than $R_o^+, Q_o^+$ and for a constant $k$, the area in which $L_o$ is better than $R_o$ can be estimated by $k_o \geq \frac{k - 2}{4}$.

In order to estimate the parameters of the model, the iterative trim and fill algorithm which has been mentioned in Duval and Tweedie (2000a, 2000b)'s article is used.

**Application of trim and fill results to meta-analysis data**

This example is based on data collected by Hartung and Knapp (2001), which are about dyspepsia caused by non-ulcer factors. Among other medicines which are gastroprokinetic, cisapride is as the first experimental treatment. As for the articles on this issue, 13 controlled trials performed throughout the world in connection with the investigation of cisapride were collected by Hartung and Knapp. The collected data is related to effect sizes based on binary data ($2 \times 2$ tables). What was done in the present study about methods of publication bias presence, was applied to this set of data, and the results of the original article Hartung and Knapp (2001), were corrected using trim and fill method.

**Checking existence of publication bias**

By examining publication bias tests on this set of data, according to (2) and (4) relations, effect size of relative risk has the highest significance in both Begg and Egger tests. Using Begg's rank correlation test, statistic value was $Z = 2.5624$ and $P = Value = 0.01040$. As for the results of Egger's linear regression test, statistic value was $t = 2.7913$ and with degree of freedom of 11, probability value was $P = Value = 0.01755$. For the effect size of odds ratio, using Begg's rank correlation test, statistic value was $Z = 1/5862$ and $P = Value = 0/1127$. For Egger's linear regression test, statistic value was $t = 2/0461$, $\hat{\beta} = -0/3588$, and with degree of freedom of 11, probability value was $P = Value = 0/06542$.

**Estimate of missing points by trim and fill method for the log relative risk**

Trim and fill algorithm is based on a qualitative and formulaic method in using funnel plot. In estimating the true effect size for this set of data, both fixed and random effects models and the $L_o^+$ estimator to estimate missing studies according to relation (5) were used. After running the algorithm, this value was $L_o = 4.97$ which was rounded to $L_o^+ = 5$. The estimated fill points are shown as the open circles in figure 1, it is clear that these points cause asymmetry in funnel plot. Therefore, for correction of the parameters of the meta-analysis 18 studies are used.
Estimate of missing points by trim and fill method for the log odds ratio

In the estimate of the true effect size for the log odds ratio, both fixed and random effects models and the $L_0$ estimator to estimate missing studies were used. After running the algorithm this value was $L_0 = 3.48$ which was rounded to $L_0 = 4$. These results are shown as the open circles in figure 2.

Analyzing trim and fill for probability difference, no missing studies were estimated. This is an indicator of a totally symmetric funnel plot. The correction of the meta-analysis model parameters, using trim and fill method is summarized in table 1.

![](image1.png)

Figure 1. Trim and fill plot for the log relative risk

![Image](image2.png)

Figure 2. Trim and fill plot for the log odds ratio

Table 1. The correction of the meta-analysis parameters, using trim and fill method

<table>
<thead>
<tr>
<th></th>
<th>fixed-effect model</th>
<th>LogRiskRatio</th>
<th>LogoddsRatio</th>
<th>risk difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Combined effect size</td>
<td>0.2855</td>
<td>0.946</td>
<td>0.3414</td>
<td></td>
</tr>
<tr>
<td>95% confidence interval</td>
<td>[0.1710, 0.4000]</td>
<td>[0.7088, 1.2804]</td>
<td>[0.2820, 0.4007]</td>
<td></td>
</tr>
<tr>
<td>p-value</td>
<td>&lt;0.0001</td>
<td>&lt;0.0001</td>
<td>&lt;0.0001</td>
<td></td>
</tr>
<tr>
<td>Random-effects model</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Combined effect size</td>
<td>0.3156</td>
<td>1.0585</td>
<td>0.3381</td>
<td></td>
</tr>
<tr>
<td>95% confidence interval</td>
<td>[0.04133, 0.5900]</td>
<td>[0.4449, 1.6720]</td>
<td>[0.2145, 0.4616]</td>
<td></td>
</tr>
<tr>
<td>p-value</td>
<td>0.0241</td>
<td>0.0007</td>
<td>&lt;0.0001</td>
<td></td>
</tr>
</tbody>
</table>
CONCLUSION
Due to being retrospective, meta-analysis is subject to bias; because publication of an article is usually influenced by the significance of the studied problem, and those articles which have significant results get published. It is obvious that if we don’t consider this bias, the results cannot be trusted. Therefore, the consequences of these biases should be specified. In this article, after examining the data set of Hartung and Knapp, it was determined that the effect sizes of relative risk and odds ratio have publication bias, and for improvement of these parameters, trim and fill method was used to correct publication bias.

REFERENCE