META-ANALYSIS FOR $2 \times 2$ TABLES WITH
MULTIPLE TREATMENT GROUPS

by

LEON J. GLESER
University of Pittsburgh

and

INGRAM OLKIN
Departments of Statistics
Stanford University

April 1999

Department of Statistics
Sequoia Hall
STANFORD UNIVERSITY
Stanford, California 94305-4065

http://www-stat.stanford.edu
META-ANALYSIS FOR $2 \times 2$ TABLES WITH MULTIPLE TREATMENT GROUPS

by
Leon J. Gleser and Ingram Olkin
University of Pittsburgh and Stanford University

Abstract

Meta-analytic methods have been developed for determining the effectiveness of a single treatment versus a control. However, in large studies more than one competing treatment may be tested for its effectiveness. Further, different sites may use different subsets of treatments. This leads to a model with missing data and with correlated effect sizes. A regression procedure is developed that yields an estimate of the overall effect.

Key words: risk differences, odds ratios, arcsin transformation, multiple comparisons

1. PRELIMINARIES

Meta-analytic methods have been developed for determining the effectiveness of a single treatment versus a control or standard by combining such comparisons over a number of studies. When the endpoint is dichotomous (e.g., effective or not effective, side effects present or absent, life or death), effectiveness is typically measured in terms of differences in risks (proportions) or odds ratios. In large studies, more than one treatment may be involved, with each treatment being compared to the common control. This is particularly true of pharmaceutical studies, in which the effects of several drugs or drug doses are compared in order to identify the most promising choices. Because investigators may have different goals, or are prevented by financial or other constraints from testing all possible treatments, different studies may involve different treatments. When later a meta-analytic review is attempted of all studies that involve the treatments of interest to the researcher, the facts that some studies may be missing one or more of the treatments and that (because of the common control) the effect sizes within studies are correlated need to be accounted for in the statistical analysis.

The present Chapter illustrates, in the context of an example, how information from studies can be combined to estimate increments in proportions (or in log odds ratios) due to various treatments. Also shown is how to construct appropriate simultaneous confidence intervals for such increments (and for contrasts of increments). The approach illustrated is approximate, with the approximation being best when studies have large sample sizes. It can be regarded as the specialization to sample proportions of the general approach in Gleser and Olkin (1994).
2. EXAMPLE

Suppose that one is interested in combining the results of several studies of the effectiveness of one or more of three anti-hypertension therapies in preventing heart disease in males or females considered “at risk” because of excessive high blood pressure readings. For specificity, suppose that the therapies are $T_1 =$ use of a new beta-blocker, $T_2 =$ use of a drug that reduces low-density cholesterol and $T_3 =$ use of a calcium channel blocker, with the endpoint being the occurrence of coronary heart disease. In every study, the effects of these therapies are compared to that of using only a diuretic (the control). Although the data are hypothetical, they represent an actual therapy.

A statistical problem arises from the fact that a comparison of therapy $T_i$ with the control and a comparison of therapy $T_j$ with the control, in the same study, have the control in common. If the control rate of occurrence of heart disease is high, then it is likely that both therapies will show an effect: if the control rate is low, it will be difficult for the therapies to do better (be effective). Consequently, the comparisons of the therapies with the control are positively correlated. and information about the effect size for therapy $T_i$ can be used to predict the effect size for any therapy $T_j$ in the same study, and vice versa.

To illustrate the model, suppose a thorough search of published and unpublished studies (using various medical registers) yields only 5 studies. Although all of these studies use a similar control, each study need not include all of the treatments, as exemplified by the hypothetical data given in Table 1.

<table>
<thead>
<tr>
<th>Study</th>
<th>Control</th>
<th>$T_1$</th>
<th>$T_2$</th>
<th>$T_3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>20/1000</td>
<td>100/4000</td>
<td>150/4000</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>(0.0200)</td>
<td>(0.0250)</td>
<td>(0.0375)</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>10/200</td>
<td>15/400</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>(0.0500)</td>
<td>(0.0375)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>40/450</td>
<td>—</td>
<td>30/400</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>(0.0889)</td>
<td></td>
<td>(0.0750)</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>150/2000</td>
<td>—</td>
<td>80/1000</td>
<td>50/1000</td>
</tr>
<tr>
<td></td>
<td>(0.0750)</td>
<td></td>
<td>(0.0800)</td>
<td>(0.0500)</td>
</tr>
<tr>
<td>5</td>
<td>60/400</td>
<td>—</td>
<td>—</td>
<td>50/400</td>
</tr>
<tr>
<td></td>
<td>(0.1500)</td>
<td></td>
<td></td>
<td>(0.1250)</td>
</tr>
</tbody>
</table>

TABLE 1

Number and proportions of subjects exhibiting heart disease for a control and three therapies
For our illustration, effect sizes will be measured by differences in proportions:

\[
effect \text{ size} = \text{proportion control} - \text{proportion therapy} \equiv p_c - p_t.
\]

Modification of the methods for other measures of effect sizes are discussed in Sections 4 and 5. The effect size estimates for risk differences, arranged by study, are:

\[
\begin{align*}
    d_1 &= (-0.0050, -0.0175, \ldots) & d_2 &= (0.0125, \ldots, \ldots) \\
    d_3 &= (\ldots, 0.0139, \ldots) & d_4 &= (\ldots, -0.0050, 0.0250) \\
    d_5 &= (\ldots, \ldots, 0.0250),
\end{align*}
\]

where the dashes indicate effect sizes that are unavailable from the particular study.

The (estimated) variance of a difference of proportions \( p_0 - p_i \) is

\[
\begin{align*}
\hat{\text{Var}}(p_0 - p_i) &= \frac{p_0(1-p_0)}{n_0} + \frac{p_i(1-p_i)}{n_i},
\end{align*}
\]

where \( n_0 \) is the number of subjects given the control treatment and \( n_i \) is the number of subjects given therapy \( T_i \). The covariance between two such differences \( p_0 - p_i \) and \( p_0 - p_j \) is \( p_0(1-p_0)/n_0 \). The following are estimated covariance matrices for each study, where only entries corresponding to effect sizes that can be estimated in the study are given:

\[
\begin{align*}
\hat{\Psi}_1 &= 10^{-9} \begin{bmatrix} 25694 & 19600 \\ 19600 & 28623 \end{bmatrix}, & \hat{\Psi}_4 &= 10^{-9} \begin{bmatrix} 108288 & 34688 \\ 34688 & 82188 \end{bmatrix}, \\
\hat{\Psi}_2 &= 10^{-9}(327734), & \hat{\Psi}_3 &= 10^{-9}(353410), & \hat{\Psi}_5 &= 10^{-9}(592188).
\end{align*}
\]

3. A REGRESSION PROCEDURE FOR RISK DIFFERENCES

Following the approach in Gleser and Olkin (1994), we combine the vectors \( d_i \) of effect sizes into a single vector, omitting all missing comparisons. We thus obtain from (1) the vector

\[
\begin{align*}
    d &= (-0.0050, -0.0175, 0.0125, 0.0139, -0.0050, 0.0250, 0.0250).
\end{align*}
\]

The estimated covariance matrix of \( d \) is obtained from (3), where the blocks represent the studies.

3
(5) \[
\begin{bmatrix}
25694 & 19600 & 00000 & 00000 & 00000 & 00000 \\
19600 & 28623 & 00000 & 00000 & 00000 & 00000 \\
00000 & 00000 & 327734 & 00000 & 00000 & 00000 \\
00000 & 00000 & 00000 & 353410 & 00000 & 00000 \\
00000 & 00000 & 00000 & 00000 & 108288 & 34688 \\
00000 & 00000 & 00000 & 00000 & 34688 & 82188 \\
00000 & 00000 & 00000 & 00000 & 00000 & 592188 \\
\end{bmatrix}
\]

\[\hat{\Psi} = 10^{-9}\]

Let \(\beta_i\) be the effect size (assumed common to all studies) for therapy \(T_i\), \(i = 1, 2, 3\), and let \(\beta = (\beta_1, \beta_2, \beta_3)\). We can now write a regression model for \(d\), namely

\[d = X\beta + \text{error},\]

with design matrix

\[
X = \begin{bmatrix}
1 & 0 & 0 \\
0 & 1 & 0 \\
1 & 0 & 0 \\
0 & 1 & 0 \\
0 & 1 & 0 \\
0 & 0 & 1 \\
0 & 0 & 1
\end{bmatrix},
\]

in which the columns represent the three therapies and the rows correspond to groups of individuals who receive one of the therapies, two groups in studies 1 and 4 and one group in studies 2, 3, and 5.

The estimates of effect sizes (here, rate differences) are obtained from a weighted least squares fit of this regression model:

(7) \[
\hat{\beta} \equiv (\hat{\beta}_1, \hat{\beta}_2, \hat{\beta}_3) = (X'\hat{\Psi}^{-1}X)^{-1}X'\hat{\Psi}^{-1}d,
\]
where \( \hat{\Psi} \) is the sample covariance matrix given in (5). The needed vector-matrix operations can easily be carried out using most standard statistical software packages. For our example,

\[
\hat{\beta}_1 = -0.0010808, \quad \hat{\beta}_2 = -0.0125192, \quad \hat{\beta}_3 = 0.0228495.
\]

The estimated variance-covariance matrix of the effect size estimates is

\[
C = \text{Cov}(\hat{\beta}) = (X'\hat{\Psi}^{-1}X)^{-1} = 10^{-9} \begin{pmatrix}
20806 & 13599 & 3889 \\
13599 & 20604 & 5893 \\
3889 & 5893 & 65145
\end{pmatrix}.
\]

Let \( c_{ij} \) be the \((i, j)\)-th element of \( C = \text{Cov}(\hat{\beta}) \). We will need \( C \) to construct approximate simultaneous confidence intervals for the effect sizes or linear combinations of the effect sizes.

Simultaneous 95\% confidence intervals for the individual therapy effects are of the form:

\[
\hat{\beta}_i \pm [\chi^2_q(0.95)c_{ii}]^{1/2}, \quad i = 1, 2, 3,
\]

where \( \chi^2_q(a) \) is the 100\(a\)-th percentile of the chi-squared distribution with \( q \) degrees of freedom and \( q \) is the dimension of the vector \( \beta \); in our example, \( q = 3 \). Table 2 gives these intervals for our example. (Alternatively, the Bonferroni or maximum modulus methods could be used; see Seber (1977). For multiple comparison methods see Hochberg and Tamhane (1987).)

Simultaneous 95\% confidence intervals for the three possible comparisons \( \beta_1 - \beta_2 \), \( \beta_1 - \beta_3 \) and \( \beta_2 - \beta_3 \) of therapy effect sizes are given by

\[
\hat{\beta}_i - \hat{\beta}_j \pm [\chi^2_3(0.95)(c_{ii} + c_{jj} - 2c_{ij})]^{1/2}, \quad 1 \leq i < j \leq 3.
\]

Table 3 gives these intervals for our example.

These confidence intervals can also be used for simultaneous two-sided tests of hypotheses about the values of the effect sizes \( \beta_i \), or of comparisons between the effect sizes. Thus, for example, on the basis of Table 2 we reject the null hypothesis that the effect size for therapy 3 is 0, but do not reject null hypotheses that the effect sizes for therapies 1 and 2 are 0, at the simultaneous 5\% level of significance. Similarly, from Table 3 we conclude that only the pair of therapies 2 and 3 have significantly different effect sizes at the simultaneous 5\% level of significance.

**TABLE 2**

<table>
<thead>
<tr>
<th>Effect Size</th>
<th>Confidence Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \beta_1 )</td>
<td>(-0.00108 \pm 0.01275 ) or ([-0.0138, 0.0117])</td>
</tr>
<tr>
<td>( \beta_2 )</td>
<td>(-0.01252 \pm 0.01269 ) or ([-0.0252, 0.0002])</td>
</tr>
<tr>
<td>( \beta_3 )</td>
<td>(0.02285 \pm 0.02256 ) or ([0.0003, 0.0454])</td>
</tr>
</tbody>
</table>
TABLE 3
Simultaneous 95% confidence intervals for $\beta_i - \beta_j$, $1 \leq i < j \leq 3$

<table>
<thead>
<tr>
<th>Comparison</th>
<th>Confidence Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta_1 - \beta_2$</td>
<td>$0.01144 \pm 0.01054$ or $[-0.0096, 0.0220]$</td>
</tr>
<tr>
<td>$\beta_1 - \beta_3$</td>
<td>$-0.02393 \pm 0.02472$ or $[-0.0486, 0.0008]$</td>
</tr>
<tr>
<td>$\beta_2 - \beta_3$</td>
<td>$-0.03537 \pm 0.02404$ or $[-0.0594, -0.0113]$</td>
</tr>
</tbody>
</table>

4. **REGRESSION FOR LOG ODDS RATIOS**

If the effect size used for the therapies is the log odds ratio

$$d_i^* = \log(p_0/(1 - p_0)) - \log(p_i/(1 - p_i))$$

instead of the risk differences $p_i - p_0$, then the large-sample variance of $d_i^*$ is

$$\psi_{ii} = \frac{1}{n_0p_0(1 - p_0)} + \frac{1}{n_ip_i(1 - p_i)},$$

and the large-sample covariance between $d_i$ and $d_j$ is

$$\psi_{ij} = \frac{1}{n_0p_0(1 - p_0)}.$$

A regression model of the form given in Section 3 for the vector $d$ of estimated effect sizes can now be used to estimate the common effect sizes for the therapies. The design matrix $X$ given by (6) remains the same, and the computation of $\hat{\beta}$ in (7) uses the new values of $\psi_{ij}$.

Here

$$d = (-0.2283, -0.6466, 0.3008, 0.1850, -0.0700, 0.4321, 0.2113),$$

$$\hat{\Psi}_1 = \begin{pmatrix} 0.06128 & 0.05102 \\ 0.05102 & 0.05795 \end{pmatrix}, \quad \hat{\Psi}_2 = 0.17403, \quad \hat{\Psi}_3 = 0.06348,$$

$$\hat{\Psi}_4 = \begin{pmatrix} 0.02051 & 0.00693 \\ 0.00693 & 0.02798 \end{pmatrix}, \quad \hat{\Psi}_5 = 0.04246,$$

which from (7) yields

$$\hat{\beta}_1, \hat{\beta}_2, \hat{\beta}_3 = (0.2169, -0.1500, 0.3315)$$
with approximate covariance matrix

\[ C = (X'\hat{\Psi}X)^{-1} = 10^{-7}\begin{pmatrix} 2240 & 9258 & 1997 \\ 9258 & 11503 & 2482 \\ 1997 & 2482 & 1657 \end{pmatrix}. \]

Consequently, simultaneous 95% confidence intervals for the \(\beta\)'s are:

\[
\beta_1: [-0.2015, 0.6353], \\
\beta_2: [-0.4498, 0.1498], \\
\beta_3: [-0.0284, 0.6913],
\]

and simultaneous 95% confidence intervals for paired differences are:

\[
\beta_1 - \beta_2: [0.020, 0.714], \\
\beta_1 - \beta_3: [-0.637, 0.408], \\
\beta_2 - \beta_3: [-0.906, -0.057].
\]

Confidence intervals on the \(\beta\)'s (log odds ratios) can be converted to confidence intervals for the odd ratios, \(\omega = e^\beta\):

\[
\omega_1: [0.818, 1.888], \\
\omega_2: [0.638, 1.162], \\
\omega_3: [0.972, 1.996].
\]

5. Regression Using a Variance Stabilizing Transformation

It is well known that the transformation \(2\arcsin\sqrt{p}\) stabilizes the variance. For large samples the variance of \(2\arcsin\sqrt{p}\) (measured in radians) is \(1/n\). With this transformation the effect sizes as

\[ d_i = 2\arcsin\sqrt{p_0} - 2\arcsin\sqrt{p_i}, \]

in which case the large-sample variance of \(d_i\) is \(\psi_{ii} = (1/n_0) + (1/n_i)\) and the large-sample covariance between \(d_i\) and \(d_j\) is \(\psi_{ij} = 1/n_0\). The regression analysis now proceeds as in Section 3 with these values of \(\psi_{ij}\).

For this model

\[ d = (-0.0338, -0.1060, 0.0613, 0.0507, -0.0187, 0.1038, 0.0727), \]

\[
\hat{\Psi}_1 = \begin{pmatrix} 0.00125 & 0.00100 \\ 0.00100 & 0.00125 \end{pmatrix}, \quad \hat{\Psi}_2 = 0.00750, \quad \hat{\Psi}_3 = 0.04722,
\]

\[
\hat{\Psi}_4 = \begin{pmatrix} 0.00150 & 0.00050 \\ 0.00050 & 0.00150 \end{pmatrix}, \quad \hat{\Psi}_5 = 0.00500,\]
which from (7) yields

$$(\hat{\beta}_1, \hat{\beta}_2, \hat{\beta}_3) = (0.0058, -0.0612, 0.0860)$$

with approximate covariance matrix

$$C = (X'\hat{\Psi}X)^{-1} = 10^{-8} \begin{pmatrix} 78369 & 47589 & 12524 \\ 47589 & 63056 & 16594 \\ 12524 & 16594 & 109630 \end{pmatrix}.$$ 

Consequently, simultaneous 95% confidence intervals for the $\beta$'s are:

$$\beta_1: [-0.0725, 0.0840],$$

$$\beta_2: [-0.1314, 0.0090],$$

$$\beta_3: [-0.0065, 0.1786],$$

and simultaneous 95% confidence intervals for paired differences are:

$$\beta_1 - \beta_2: [0.0065, 0.1267],$$

$$\beta_1 - \beta_3: [-0.1935, 0.0322],$$

$$\beta_2 - \beta_3: [-0.2517, -0.0429].$$

6. DISCUSSION

There were five studies, each involving a subset of the three therapy treatments. We now summarize the confidence intervals for the various statistical methods: difference of proportions (Section 3), arc sin transformation of proportions (Section 5), and the log odds ratio (Section 4).

<table>
<thead>
<tr>
<th>Therapy</th>
<th>Differences</th>
<th>Arcsin</th>
<th>Log Odds Ratio</th>
<th>Odds Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>(-0.0138, 0.0117)</td>
<td>(-0.0725, 0.0840)</td>
<td>(-0.2015, 0.6353)</td>
<td>(0.818, 1.888)</td>
</tr>
<tr>
<td>2</td>
<td>(-0.0252, 0.0002)</td>
<td>(-0.1314, 0.0090)</td>
<td>(-0.4498, 0.1498)</td>
<td>(0.638, 1.162)</td>
</tr>
<tr>
<td>3</td>
<td>(0.0003, 0.0454)</td>
<td>(-0.0065, 0.1786)</td>
<td>(-0.0284, 0.6913)</td>
<td>(0.972, 1.996)</td>
</tr>
</tbody>
</table>

The three methods agree that therapies 1 and 2 are not efficacious with respect to the control. The arc sin and log odds ratio (or odds ratio) methods agree that therapy 3 is not efficacious, whereas the difference method shows a slight positive effect.

With respect to a comparison of the effectiveness of the treatments, the majority of the analyses show that therapies 2 and 3 are not significantly different from therapy 1, but that these therapies are different from each other.

The three methods are based on large sample theory. The difference method and the log odds ratio also require using sample estimates in the variance, whereas the arc sin transformation does not. Thus, in terms of a testing procedure, in contrast to an estimation procedure, the arc sin method is perhaps the most reliable.
ACKNOWLEDGEMENTS

The research of Leon Gleser was partially supported under Grant DMS-9504924 from the National Science Foundation. The research of Ingram Olkin was supported in part by the Centers for Disease Control and Prevention and the National Science Foundation, Grant DMS-9301366.

REFERENCES

