Chapter 455

Meta-Analysis of Means (Old Version)

Introduction

This module performs a meta-analysis on a set of two-group, continuous-scale, studies. These studies have a treatment group and a control group. Each study's result may be summarized by the sample size, mean, and standard deviation of each of the two groups. The program provides a complete set of numeric reports and plots to allow the investigation and presentation of the studies. The plots include the *forest plot, radial plot,* and *L'Abbe plot.* Both fixed- and random-effects models are available for analysis.

Meta-Analysis refers to methods for the systematic review of a set of individual studies with the aim to combine their results. Meta-analysis has become popular for a number of reasons:

1. The adoption of evidence-based medicine, which requires that all reliable information is considered.

2. The desire to avoid narrative reviews which are often misleading.

3. The desire to interpret the large number of studies that may have been conducted about a specific treatment.

4. The desire to increase the statistical power of the results by combining many small-size studies.

The goals of meta-analysis may be summarized as follows. A meta-analysis seeks to systematically review all pertinent evidence, provide quantitative summaries, integrate results across studies, and provide an overall interpretation of these studies.

We have found many books and articles on meta-analysis. In this chapter, we briefly summarize the information in Sutton et al (2000) and Thompson (1998). Refer to those sources for more details about how to conduct a meta-analysis.

Treatment Effects

Suppose you have obtained the results for *k* studies, labeled *i = 1, ..., k*. Each study consists of a treatment group (T) and a control group (C). The results of each study are summarized by six statistics:

n_{T_i}	the number of subjects in the treatment group.
n_{C_i}	the number of subjects in the control group.
$\bar{x}_{T_i}$	the sample mean of the treatment group which estimates the treatment mean μ_{Ti}.
$\bar{x}_{C_i}$	the sample mean of the control group which estimates the control mean μ_{Ci}.
s_{T_i}	the sample standard deviation of the treatment group.
s_{C_i}	the sample standard deviation of the control group.

Mean Difference

The scales (e.g., blood pressure, pulse rate, volume, etc.) of all studies must be the same. If the logarithm has been taken in one study, it must be taken in all studies. You cannot combine studies with different scales using this program!

The measure of treatment effect for study *i* is

$$\theta_i = \mu_{T_i} - \mu_{C_i}$$

which is estimated by

$$\hat{\theta}_i = \bar{x}_{T_i} - \bar{x}_{C_i}$$

The standard deviation of the difference is

$$V(\hat{\theta}_i) = \sigma_i^2 \left(\frac{1}{n_{T_i}} + \frac{1}{n_{C_i}} \right)$$

The value of σ^2 is estimated by the pooled sample standard deviation given by

$$s_i^2 = \frac{(n_{T_i} - 1)s_{T_i}^2 + (n_{C_i} - 1)s_{C_i}^2}{n_{T_i} + n_{C_i}}$$

so that

$$\hat{V}(\hat{\theta}_i) = s_i^2 \left(\frac{1}{n_{T_i}} + \frac{1}{n_{C_i}} \right)$$

Defining the Study Parameters

Confidence intervals based on the *t* distribution may be defined for θ_i in the usual manner.

$$\hat{\theta}_i \pm t_{n_{T_i}+n_{C_i}-2,1-\alpha/2} \sqrt{\hat{V}(\hat{\theta}_i)}$$

It will be useful in the sequel to make the following definition of the weights.

$$v_i = \hat{V}(\hat{\theta}_i)$$

$$w_i = 1/v_i$$

Hypothesis Tests

Several hypothesis tests have been developed to test the various hypotheses that may be of interest. These will be defined next.

Overall Null Hypothesis

Two statistical tests have been devised to test the overall null hypothesis that all treatment effects are zero. The null hypothesis is written

$$H_0 : \theta_i = 0 \quad i = 1, \cdots, k$$

Nondirectional Test

The nondirectional alternative hypothesis that at least one $\theta_i \neq 0$ may be tested by comparing the quantity

$$X_{ND} = \sum_{i=1}^{k} w_i \hat{\theta}_i^2$$

with a χ_k^2 distribution.

Directional Test

A test of the more interesting directional alternative hypothesis that $\theta_i = \theta \neq 0$ for all *i* may be tested by comparing the quantity

$$X_D = \frac{\left(\sum_{i=1}^{k} w_i \hat{\theta}_i \right)^2}{\sum_{i=1}^{k} w_i}$$

with a χ_1^2 distribution. Note that this tests the hypothesis that all effects are equal to the same nonzero quantity.

Effect-Equality (Heterogeneity) Test

When the overall null hypothesis is rejected, the next step is to test whether all effects are equal, that is, whether the effects are homogeneous. Specifically, the hypothesis is

$$H_0 : \theta_i = \theta \quad i = 1, \cdots, k$$

versus the alternative that at least one effect is different, that is, that the effects are heterogeneous. This may also be interpreted as a test of the study-by-treatment interaction.

This hypothesis is tested using Cochran's Q test which is given by

$$Q = \sum_{i=1}^{k} w_i \left(\hat{\theta}_i - \hat{\theta} \right)^2$$

where

$$\hat{\theta} = \frac{\sum_{i=1}^{k} w_i \hat{\theta}_i}{\sum_{i=1}^{k} w_i}$$

The test is conducted by comparing Q to a χ^2_{k-1} distribution.

Fixed versus Random Effects Combined Confidence Interval

If the effects are be assumed to be equal (homogeneous), either through testing or from other considerations, a *fixed effects model* may be used to construct a combined confidence interval. However, if the effects are heterogeneous, a *random effects model* should be used to construct the combined confidence interval.

Fixed Effects Model

The fixed effects model assumes homogeneity of study results. That is, it assumes that $\theta_i = \theta$ for all *i*. This assumption may not be realistic when combining studies with different patient pools, protocols, follow-up strategies, doses, durations, etc.

If the fixed effects model is adopted, the *inverse variance-weighted* method as described by Sutton (2000) page 58 is used to calculate the confidence interval for θ. The formulas used are

$$\hat{\theta} \pm z_{1-\alpha/2} \sqrt{\hat{V}(\hat{\theta})}$$

where $z_{1-\alpha/2}$ is the appropriate percentage point from the standardized normal distribution and

$$\hat{\theta} = \frac{\sum_{i=1}^{k} w_i \hat{\theta}_i}{\sum_{i=1}^{k} w_i}$$

$$\hat{V}(\hat{\theta}) = \frac{1}{\sum_{i=1}^{k} w_i}$$

Random Effects Model

The random effects model assumes that the individual θ_i come from a random distribution with fixed mean $\bar{\theta}$ and variance σ^2. Sutton (2000) page 74 presents the formulas necessary to conduct a random effects analysis using the *weighted* method. The formulas used are

$$\hat{\bar{\theta}} \pm z_{1-\alpha/2} \sqrt{\hat{V}\left(\hat{\bar{\theta}}\right)}$$

where $z_{1-\alpha/2}$ is the appropriate percentage point from the standardized normal distribution and

$$\hat{\bar{\theta}} = \frac{\sum_{i=1}^{k} \bar{w}_i \hat{\theta}_i}{\sum_{i=1}^{k} \bar{w}_i}$$

$$\hat{V}\left(\hat{\bar{\theta}}\right) = \frac{1}{\sum_{i=1}^{k} \bar{w}_i}$$

$$\bar{w}_i = \frac{1}{\frac{1}{w_i} + \hat{\tau}^2}$$

$$\hat{\tau}^2 = \begin{cases} \dfrac{Q - k + 1}{U} & \text{if } Q > k - 1 \\ \\ 0 & \text{otherwise} \end{cases}$$

$$Q = \sum_{i=1}^{k} w_i \left(\hat{\theta}_i - \hat{\theta}\right)^2$$

$$U = (k - 1)\left(\bar{w} - \frac{s_w^2}{k\bar{w}}\right)$$

$$s_w^2 = \frac{1}{k-1}\left(\sum_{i=1}^{k} w_i^2 - k\bar{w}^2\right)$$

$$\bar{w} = \frac{1}{k}\left(\sum_{i=1}^{k} w_i\right)$$

Graphical Displays

A number of plots have been devised to display the information in a meta-analysis. These include the forest plot, the radial plot, and the L'Abbe plot. More will be said about each of these plots in the Output section.

Data Structure

The data are entered into a dataset using one row per study. Six variables are required to hold the sample size, mean, and standard deviation of each study. In addition to these, an additional variable is usually used to hold a short (3 or 4 character) label. Another variable may be used to hold a grouping variable.

As an example, we will use data referred to in Sutton (2000) page 30 as the dental dataset. This dataset reviews nine randomized clinical trials that were conducted to study the effects of sodium fluoride (NaF) with sodium monofluorophosphate (SMFP). These nine studies were all on the same continuous scale, so their results could be analyzed using the meta-analysis techniques presented in this chapter. These data are contained in the SUTTON30 database. You should load this database to see how the data are arranged.

Example 1 – Meta-Analysis of Means

This section presents an example of how to analyze the data contained in the Sutton30 dataset. This dataset contains data for nine randomized clinical trials that were conducted to study the effect of fluoridation. The NaF data represent the control group and the SMFP represent the treated group.

Setup

To run this example, complete the following steps:

1   Open the Sutton30 example dataset

- From the File menu of the NCSS Data window, select **Open Example Data**.
- Select **Sutton30** and click **OK**.

2   Specify the Meta-Analysis of Means procedure options

- Find and open the **Meta-Analysis of Means** procedure using the menus or the Procedure Navigator.
- The settings for this example are listed below and are stored in the **Example 1** settings file. To load these settings to the procedure window, click **Open Example Settings File** in the Help Center or File menu.

Variables Tab
_____

Treatment - N Variables...................................**SMFPN**
Treatment - Mean Variables ....**SMFPMean**
Treatment - S.D. Variables ....**SMFPSD**
Control - N Variables ....**NAFN**
Control - Mean Variables ....**NAFMean**
Control - S.D. Variables ....**NAFSD**
Label Variable...**Study**

Reports Tab
_____

Summary Report...**Checked**
Heterogeneity Tests..**Checked**
Outcome Detail Reports ....**Checked**

Plots Tab
_____

Forest Plot ....**Checked**
L'Abbe Plot ....**Checked**
Radial Plot ....**Checked**

---

3   Run the procedure

- Click the **Run** button to perform the calculations and generate the output.

Run Summary Section

Run Summary Section

Parameter	Value	Parameter	Value
N Treatment Variable	SMFPN	N Control Variable	NaFN
Mean Treatment Variable	SMFPMean	Mean Control Variable	NaFMean
SD Treatment Variable	SMFPSD	SD Control Variable	NaFSD
Group Variable	None	Number Groups	1
Row Label Variable	Study	Rows Processed	9

Note: Check that the intended variables have been selected and number of rows is correct.

This report records the variables that were used and the number of rows that were processed.

Numeric Summary Section

Numeric Summary Section Using Random Effects Model

Study	NT/NC	Mean T	Mean C	Difference	SD T	SD C
S1	113/134	6.8200	5.9600	0.8600	4.7200	4.2400
S2	151/175	5.0700	4.7400	0.3300	5.3800	4.6400
S3	140/137	2.5100	2.0400	0.4700	3.2200	2.5900
S4	179/184	3.2000	2.7000	0.5000	2.4600	2.3200
S5	169/174	5.8100	6.0900	-0.2800	5.1400	4.8600
S6	736/754	4.7600	4.7200	0.0400	5.2900	5.3300
S7	209/209	10.9000	10.1000	0.8000	7.9000	8.1000
S8	1122/1151	3.0100	2.8200	0.1900	3.3200	3.0500
S9	673/679	4.3700	3.8800	0.4900	5.3700	4.8500
[Combined]						
Average				0.2835		

Note: This report shows the input data for each study in the analysis. The 'Average' values are actually weighted averages with weights based on the effects model that was selected.

This report summarizes the input data. You should scan it for any mistakes. Note that the 'Average' line provides the estimated group average.

NT/NC

These are the count values that were read from the database.

Mean T

These are the input treatment means.

Mean C

These are the input control means.

Difference

These are the computed values of Mean T minus Mean C. These difference values are the effects of interest in the analysis.

SD T

These are the input treatment standard deviations.

SD C

These are the input control standard deviations.

Nondirectional Zero-Effect Test

Nondirectional Zero-Effect Test Using Random Effects Model

Rows	Outcome Measure	Chi-Square	DF	Prob Level
Combined	Mean Difference	14.8417	9	0.0954

Note: This chi-square value tests the null hypothesis that all effects are zero versus the alternative that at least one study had a non-zero effect.

This reports the results of the nondirectional zero-effect chi-square test designed to test the null hypothesis that all treatment effects are zero. The null hypothesis is written

$$H_0 : \theta_i = 0 \quad i = 1, \cdots, k$$

The alternative hypothesis is that at least one $\theta_i \neq 0$, that is, at least one study had a statistically significant result.

Chi-Square

This is the computed chi-square value for this test. The formula was presented earlier.

DF

This is the degrees of freedom. For this test, the degrees of freedom is equal to the number of studies.

Prob Level

This is the significance level of the test. If this value is less than the nominal value of alpha (usually 0.05), the test is statistically significant and the alternative is concluded. If the value is larger than the specified value of alpha, no conclusion can be drawn other than that you do not have enough evidence to reject the null hypothesis.

Directional Zero-Effect Test

Directional Zero-Effect Test Using Random Effects Model

Rows	Outcome Measure	Chi-Square	DF	Prob Level
Combined	Mean Difference	9.4395	1	0.0021

Note: This chi-square value tests the null hypothesis that all effects are zero versus the alternative that all studies had the same, non-zero effect.

This reports the results of the directional zero-effect chi-square test designed to test the overall null hypothesis that all treatment effects are zero. The null hypothesis is written

$$H_0 : \theta_i = 0 \quad i = 1, \cdots, k$$

The alternative hypothesis is that $\theta_i = \theta \neq 0$ for all *i*, that is, that all effects are equal to the same, non-zero value.

Chi-Square

This is the computed chi-square value for this test. The formula was presented earlier.

DF

This is the degrees of freedom. For this test, the degrees of freedom is equal to one.

Prob Level

This is the significance level of the test. If this value is less than the specified value of alpha (usually 0.05), the test is statistically significant and the alternative is concluded. If the value is larger than the specified value of alpha, no conclusion can be drawn other than that you do not have enough evidence to reject the null hypothesis.

Effect-Equality (Heterogeneity) Test

Effect-Equality (Heterogeneity) Test

Rows	Outcome Measure	Cochran's Q	DF	Prob Level
Combined	Mean Difference	5.4022	8	0.7139

Note: This tests the null hypothesis that all effects are equal (homogeneous) versus the alternative that at least one effect had a different effect (heterogeneous). Sometimes this test is used to choose between the use of a Fixed Effect (homogeneous) model and a Random Effects (heterogeneous) model.

This reports the results of the effect-equality (homogeneity) test. This chi-square test was designed to test the null hypothesis that all treatment effects are equal. The null hypothesis is written

$$H_0 : \theta_i = \theta \quad i = 1, \cdots, k$$

The alternative is that at least one effect is different, that is, that the effects are heterogeneous. This may also be interpreted as a test of the study-by-treatment interaction. This test may help you determine whether to use a Fixed Effects model (used for homogeneous effects) or a Random Effects model (heterogeneous effects).

Cochran's Q

This is the computed chi-square value for Cochran's Q statistic. The formula was presented earlier.

DF

This is the degrees of freedom. For this test, the degrees of freedom is equal to the number of studies minus one.

Prob Level

This is the significance level of the test. If this value is less than the specified value of alpha (usually 0.05), the test is statistically significant and the alternative is concluded. If the value is larger than the specified value of alpha, no conclusion can be drawn other than that you do not have enough evidence to reject the null hypothesis.

Mean Difference Detail Section

Mean Difference Detail Section Using Random Effects Model

Study	Mean Difference	Standard Error	95.0% Lower Confidence Limit	95.0% Upper Confidence Limit	Percent Random Effects Weight
S1	0.8600	0.5704	-0.2635	1.9835	2.6172
S2	0.3300	0.5549	-0.7617	1.4217	2.7648
S3	0.4700	0.3516	-0.2221	1.1621	6.8887
S4	0.5000	0.2509	0.0066	0.9934	13.5238
S5	-0.2800	0.5400	-1.3421	0.7821	2.9199
S6	0.0400	0.2752	-0.4993	0.5793	11.2455
S7	0.8000	0.7826	-0.7384	2.3384	1.3900
S8	0.1900	0.1337	-0.0720	0.4520	47.6529
S9	0.4900	0.2782	-0.0554	1.0354	10.9974
[Combined]					
Average	0.2835	0.0923	0.1026	0.4643	

Note: This report presents the difference's value and a confidence interval. The 'Average' line presents the combined estimates for the group. The weights let you determine the influence of each study on the combined results.

This report displays results for the mean difference outcome measure.
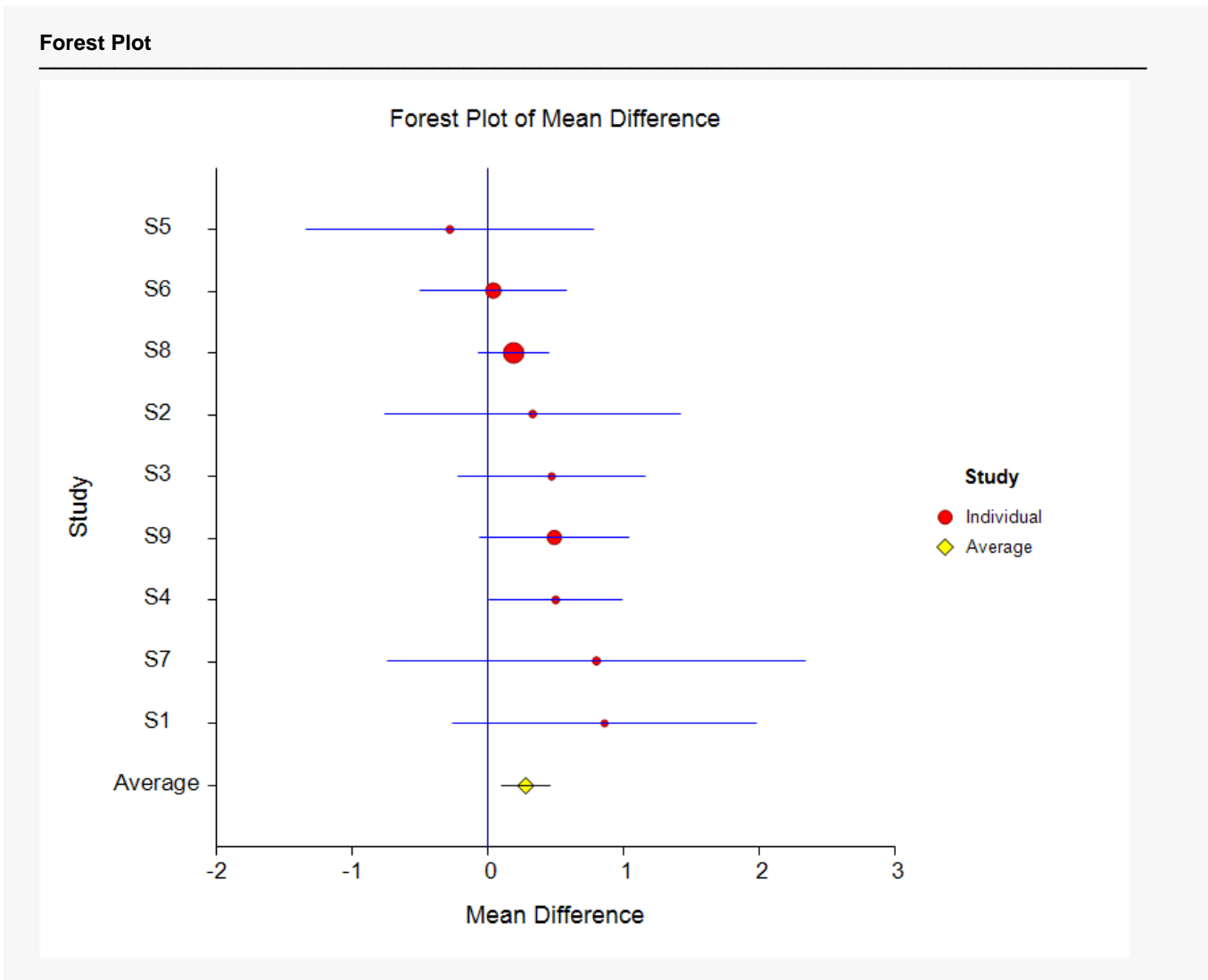
Confidence Limits

These are the lower and upper confidence limits (the formulas were given earlier in this chapter).

Weights

The last column gives the relative (percent) weight used in creating the weighted average. Using these values, you can decide how much influence each study has on the weighted average.

Forest Plot

Forest Plot
_____

This plot presents the results for each study on one plot. The size of the plot symbol is proportional to the sample size of the study. The points on the plot are sorted by the mean difference. The lines represent the confidence intervals about the mean differences. Note that the narrower the confidence limits, the better.

By studying this plot, you can determine the main conclusions that can be drawn from the set of studies. For example, you can determine how many studies were significant (the confidence limits do not intersect the vertical line at 0.0).
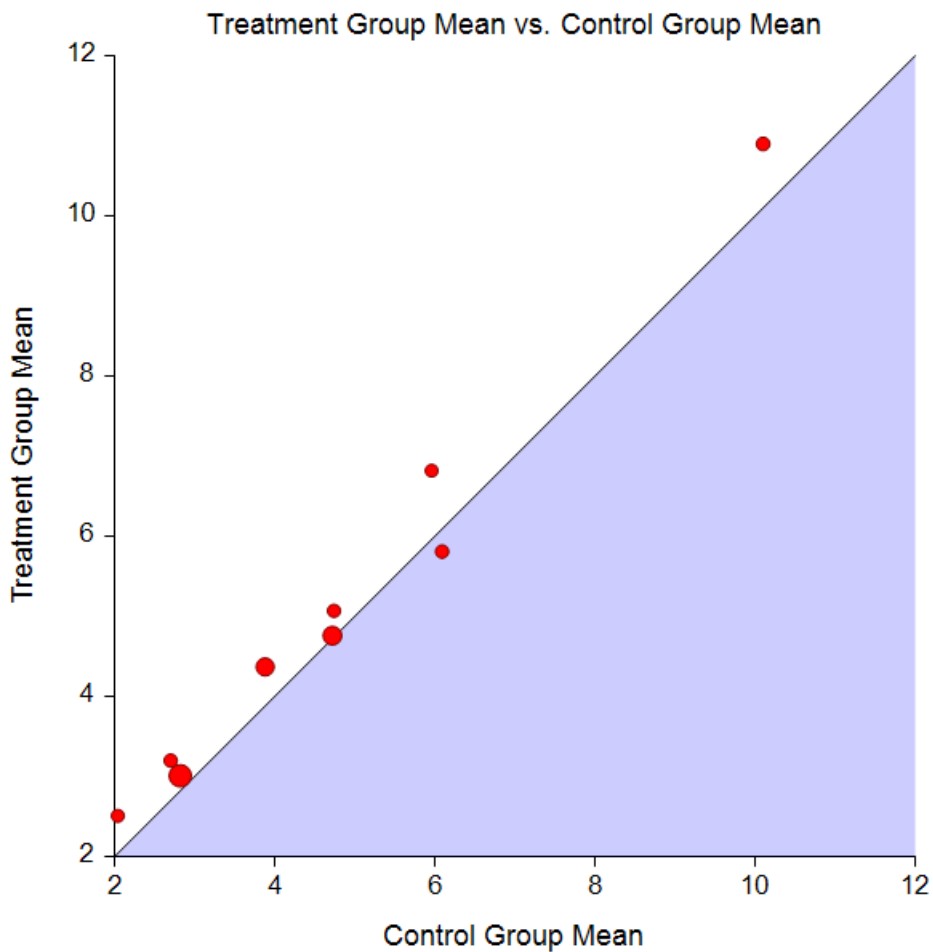
Radial Plot

Radial Plot
_____
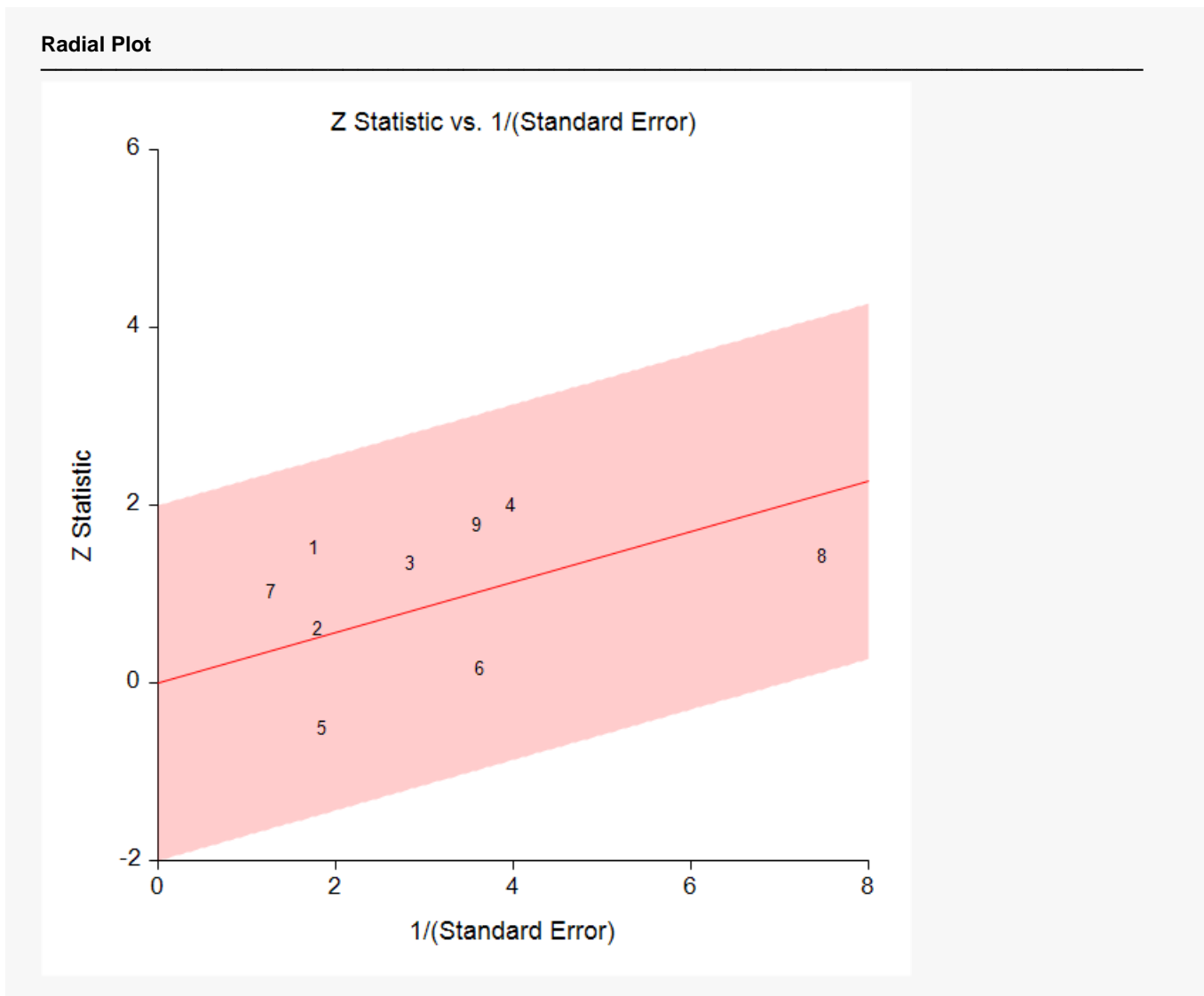
The radial (or Galbraith) plot shows the z-statistic (outcome divided by standard error) on the vertical axis and a measure of weight on the horizontal axis. Studies that have the largest weight are closest to the Y axis. Studies within the limits are interpreted as homogeneous. Studies outside the limits may be outliers.

L'Abbe Plot

L'Abbe Plot
_____

The L'Abbe plot displays the treatment mean on vertical axis versus the control mean on the horizontal axis. Homogenous studies will be arranged along the diagonal line. This plot is especially useful in determining if the relationship between the treatment group and the control group is the same for all values of the control group risk.