Chapter 458

Meta-Analysis of Hazard Ratios

Introduction

This module performs a meta-analysis on a set of two-group, time to event (survival), studies in which some data may be censored. These studies have a treatment group and a control group. Each study's result may be summarized by the log hazard ratio and its standard error. 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* and *radial 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 be 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.

As for the particular topic of combining hazard ratio studies in a meta-analysis, the book by Parmar and Machin (1995) and the paper by Parmar *et al.*(1998) are essential reading. The paper provides instructions on how to obtain estimates of the hazard ratio and its standard error from trials that do not report these items explicitly (a situation that is common).

Treatment Effect – Hazard Ratio

The most recommended single summary statistic for quantifying the treatment effect in studies using survival data is the (log) hazard rate. This statistic is chosen because it can be calculated from time-to-event data with censoring and because it measures the size of the difference between two Kaplan-Meier curves.

The Cox-Mantel estimate of the *hazard ratio* is formed by dividing the hazard rate under treatment by the hazard rate under control. Thus, it measures the change in risk of treatment versus control over the follow-up period. Since the distribution of the log hazard ratio is nearly normal, the log transformation is applied. The formula for the hazard rate is

$$HR_{CM} = \frac{H_T}{H_C}$$
$$= \frac{O_T / E_T}{O_C / E_C}$$

where O_i is the observed number of events (deaths) in group i, E_i is the expected number of events (deaths) in group i, and H_i is the overall hazard rate for the ith group. The calculation of the E_i is explained in Parmar and Machin (1995).

A confidence interval for *HR* is found by first transforming to the log scale which is better approximated by the normal distribution, calculating the limits, and then transforming back to the original scale. The calculation is made using

$$\ln(HR_{CM}) \pm z_{1-\alpha/2}\left(SE_{\ln HR_{CM}}\right)$$

where

$$SE_{\ln HR_{CM}} = \sqrt{\frac{1}{E_T} + \frac{1}{E_C}}$$

An alternative estimate of *HR* that is sometimes used is the Mantel-Haenszel estimator which is calculated using

$$HR_{MH} = \exp\left(\frac{O_T - E_T}{V}\right)$$

where *V* is the hypergeometric variance. For further details, see Parmar and Machin (1995). A confidence interval for *HR* is found by first transforming to the log scale which is better approximated by the normal distribution, calculating the limits, and then transforming back to the original scale. The calculation is made using

$$\ln(HR_{MH}) \pm z_{1-\alpha/2}\left(SE_{\ln HR_{MH}}\right)$$

where

$$SE_{\ln HR_{MH}} = \sqrt{\frac{1}{V}}$$

If the log hazard ratio and its standard error are not reported in a particular study it will have to be estimated from the logrank test statistic, p-value, or from the Kaplan-Meier curves. Details of how to do this are presented in Parmar *et al.* (1998).

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 two statistics:

$\ln(HR_i)$ the log hazard ratio.

$SE_{\ln(HR_i)}$ the standard error of the log hazard ratio.

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

$$v_i = \left(SE_{\ln HR}\right)^2$$

$$w_i = 1/v_i$$

Hypothesis Tests

In the discussion below, we let θ_i represent $\ln HR_i$. Several hypothesis tests have be 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

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

Nondirectional Test

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

$$X_{ND} = \sum_{i=1}^{k} w_i \theta_i^2$$

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 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

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

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

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

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

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

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

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

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

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

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

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

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

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

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. Two variables are required to hold the log hazard ratio and its standard error. 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 a dataset giving the results for survival studies. The results of these studies are recorded in the MetaHR dataset. You should load this database to see how the data are arranged.

Procedure Options

This section describes the options available in this procedure.

Variables Tab

The options on this screen control the variables that are used in the analysis.

Variables

Log(Hazard Ratio) Variable

Specify the variable containing the log hazard ratio of each study. Each row of data represents a separate study. Note that the base of the logarithm (e or 10) is arbitrary. However, it must be consistent throughout the dataset.

S.E. Log(Hazard Ratio) Variable

Specify the variable containing the standard error of the log hazard ratio of each study. Each row of data represents a separate study. Note that the base of the logarithm (e or 10) is arbitrary. However, it must be consistent throughout the dataset.

Variables – Optional Variables

Label Variable

Specify an optional variable containing a label for each study (row) in the database. This label should be short (< 8 letters) so that it can fit on the plots.

Group Variable

Specify an optional variable containing a group identification value. Each unique value of this variable will receive its own plotting symbol on the forest plots. Some reports are sorted by these group values.

Combine Studies Method

Combine Studies Using

Specify the method used to combine treatment effects.

Use the **Fixed Effects** method when you do not want to account for the variation between studies.

Use the **Random Effects** method when you want to account for the variation between studies as well as the variation within the studies.

Reports Tab

The options on this screen control the appearance of the reports.

Select Reports

Summary Report - Outcome Detail Reports

Indicate whether to display the corresponding report.

Alpha Level

This setting controls the confidence coefficient used in the confidence limits. Note that 100 x (1 - alpha)% confidence limits will be calculated. This must be a value between 0 and 0.5. The most common choice is 0.05, which results in 95% confidence intervals.

Report Options

Show Notes

Indicate whether to show the notes at the end of reports. Although these notes are helpful at first, they may tend to clutter the output. This option lets you omit them.

Precision

Specify the precision of numbers in the report. A single-precision number will show seven-place accuracy, while a double-precision number will show thirteen-place accuracy. Note that the reports are formatted for single precision. If you select double precision, some numbers may run into others. Also note that all calculations are performed in double precision regardless of which option you select here. Single precision is for reporting purposes only.

Variable Names

This option lets you select whether to display only variable names, variable labels, or both.

Report Options – Decimal Places

Probability Values – Z Values

This setting controls the number of digits to the right of the decimal place that are displayed when showing this item.

Plots Tab

The options on this panel control the inclusion and the appearance of the forest plot and the radial plot.

Select Plots

Forest Plot – Radial Plot

Indicate whether to display the corresponding plot. Click the plot format button to change the plot settings.

Storage Tab

These options let you specify if, and where on the dataset, various statistics are stored.

Warning: Any data already in these columns are replaced by the new data. Be careful not to specify columns that contain important data.

Data Storage Options

Storage Option

This option controls whether the values indicated below are stored on the dataset when the procedure is run.

- **Do not store data**

  No data are stored even if they are checked.

- **Store in empty columns only**

  The values are stored in empty columns only. Columns containing data are not used for data storage, so no data can be lost.

- **Store in designated columns**

  Beginning at the *First Storage Variable*, the values are stored in this column and those to the right. If a column contains data, the data are replaced by the storage values. Care must be used with this option because it cannot be undone.

Store First Item In

The first item is stored in this column. Each additional item that is checked is stored in the columns immediately to the right of this variable.

Leave this value blank if you want the data storage to begin in the first blank column on the right-hand side of the data.

Warning: any existing data in these columns is automatically replaced, so be careful.

Data Storage Options – Select Items to Store with the Dataset

Log(HR) - Weights

Indicate whether to store these row-by-row values, beginning at the column indicated by the *Store First Item In* option.

Example 1 – Meta-Analysis of Hazard Ratios

This section presents an example of how to analyze the data contained in the MetaHR dataset. This dataset contains data for sixteen randomized clinical trials with survival endpoints.

You may follow along here by making the appropriate entries or load the completed template **Example 1** by clicking on Open Example Template from the File menu of the Meta-Analysis of Hazard Ratios window.

1   Open the MetaHR dataset.
- From the File menu of the NCSS Data window, select **Open Example Data**.
- Click on the file **MetaHR.NCSS**.
- Click **Open**.

2   Open the Meta-Analysis of Hazard Ratios window.
- Using the Analysis menu or the Procedure Navigator, find and select the **Meta-Analysis of Hazard Ratios** procedure.
- On the menus, select **File**, then **New Template**. This will fill the procedure with the default template.

3   Select the variables.
- Select the **Variables tab**.
- Set the **Log(Hazard Ratio) Variable** to **LogHR**.
- Set the **S.E. Log(Hazard Ratio) Variable** to **SELogHR**.
- Set the **Label Variable** to **Study**.

4   Specify the reports.
- Select the **Reports tab**.
- Check the **Summary Report** option box.
- Check the **Heterogeneity Tests** option box.
- Check the **Outcome Detail Reports** option box.
- On the **Plots** tab, check the **Forest Plot** option box.
- Check **Radial Plot** option box.

5   Run the procedure.
- From the Run menu, select **Run Procedure**. Alternatively, just click the green Run button.

Run Summary Section

Parameter	Value	Parameter	Value
Log HR Variable	LogHR	SE(Log HR) Variable	SELogHR
Group Variable	None	Number Groups	1
Row Label Variable	Study	Rows Processed	16

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

Numeric Summary Section

Study	Log HR	SE(Log HR)
S1	-0.1350	0.0799
S2	-0.2570	0.0734
S3	-0.4610	0.0492
S4	0.2030	0.0401
S5	-0.7980	0.1203
S6	-0.3240	0.0933
S7	-0.5510	0.0577
S8	-0.6820	0.1084
S9	-0.3340	0.1385
S10	-0.3840	0.0472
S11	0.0564	0.0671
S12	-0.9910	0.0528
S13	-0.7230	0.0319
S14	-0.4240	0.0289
S15	0.0178	0.0817
S16	-0.1870	0.0203

[Combined]
Average   -0.3712

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

Nondirectional Zero-Effect Test

Rows	Outcome Measure	Chi-Square	DF	Prob Level
Combined	Log(Hazard Ratio)	1554.1876	16	0.0000

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

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

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

Rows	Outcome Measure	Chi-Square	DF	Prob Level
Combined	Log(Hazard Ratio)	902.6977	1	0.0000

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

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

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 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

Treatment	Outcome Measure	Cochran's Q	DF	Prob Level
Combined	Log(Hazard Ratio)	651.4927	15	0.0000

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.

Log(Hazard Ratio) Detail Section

Study	Log Hazard Ratio	Standard Error	95.0% Lower Confidence Limit	95.0% Upper Confidence Limit	Percent Random Effects Weight
S1	-0.1350	0.0799	-0.2916	0.0216	6.1953
S2	-0.2570	0.0734	-0.4009	-0.1131	6.2557
S3	-0.4610	0.0492	-0.5574	-0.3646	6.4428
S4	0.2030	0.0401	0.1244	0.2816	6.4960
S5	-0.7980	0.1203	-1.0338	-0.5622	5.7451
S6	-0.3240	0.0933	-0.5069	-0.1411	6.0591
S7	-0.5510	0.0577	-0.6641	-0.4379	6.3843
S8	-0.6820	0.1084	-0.8945	-0.4695	5.8891
S9	-0.3340	0.1385	-0.6055	-0.0625	5.5118
S10	-0.3840	0.0472	-0.4765	-0.2915	6.4553
S11	0.0564	0.0671	-0.0751	0.1879	6.3104
S12	-0.9910	0.0528	-1.0945	-0.8875	6.4190
S13	-0.7230	0.0319	-0.7855	-0.6605	6.5352
S14	-0.4240	0.0289	-0.4806	-0.3674	6.5474
S15	0.0178	0.0817	-0.1423	0.1779	6.1779
S16	-0.1870	0.0203	-0.2268	-0.1472	6.5759
[Combined]					
Average	-0.3712	0.0800	-0.5280	-0.2145	

This report displays results for the log hazard ratio.

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
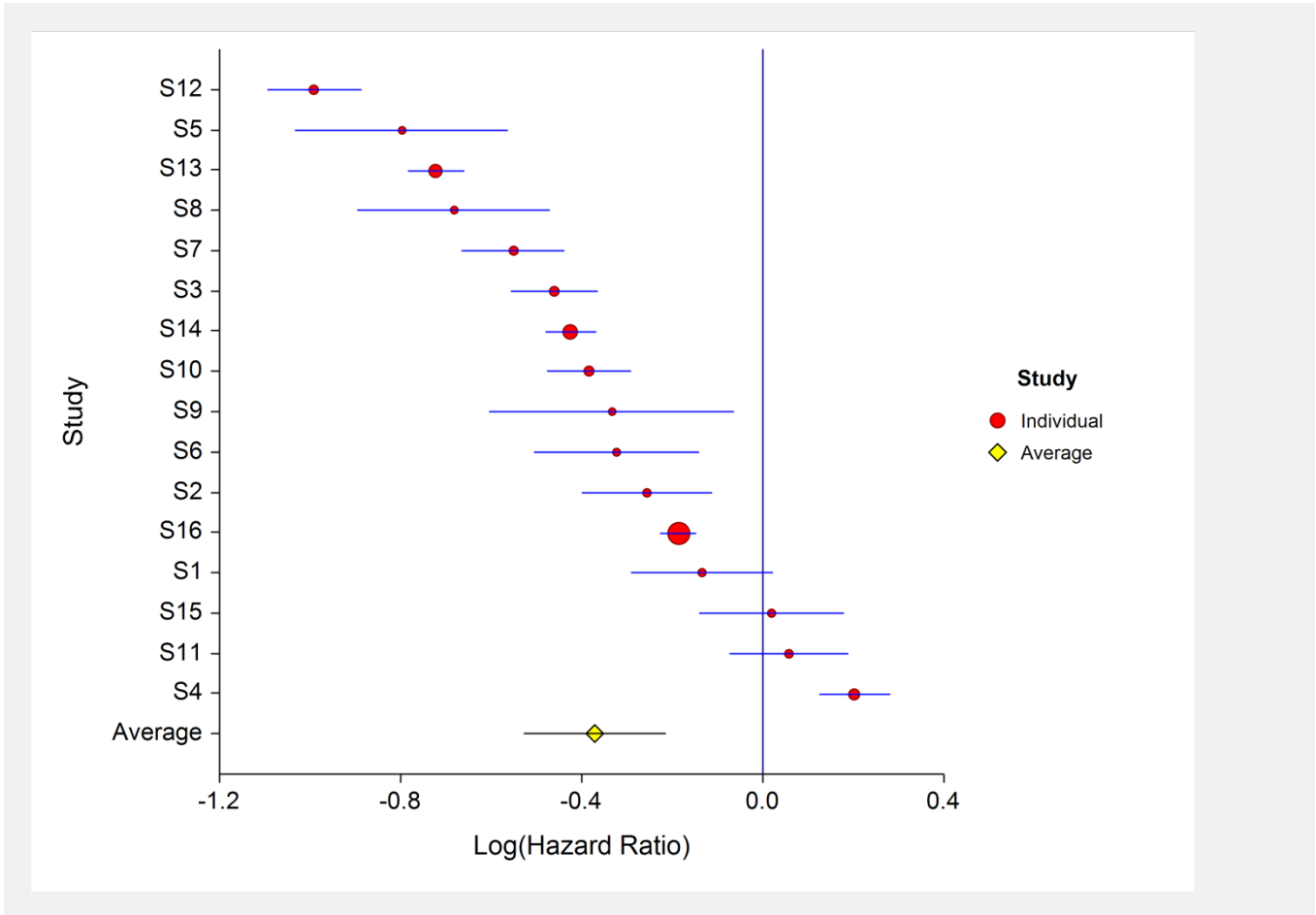
This plot presents the results for each study on one plot. The size of the plot symbol is proportional to the sample size of the study. The points on the plot are sorted by the mean difference. The lines represent the confidence intervals about the log hazard ratios. Note that the narrower the confidence 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
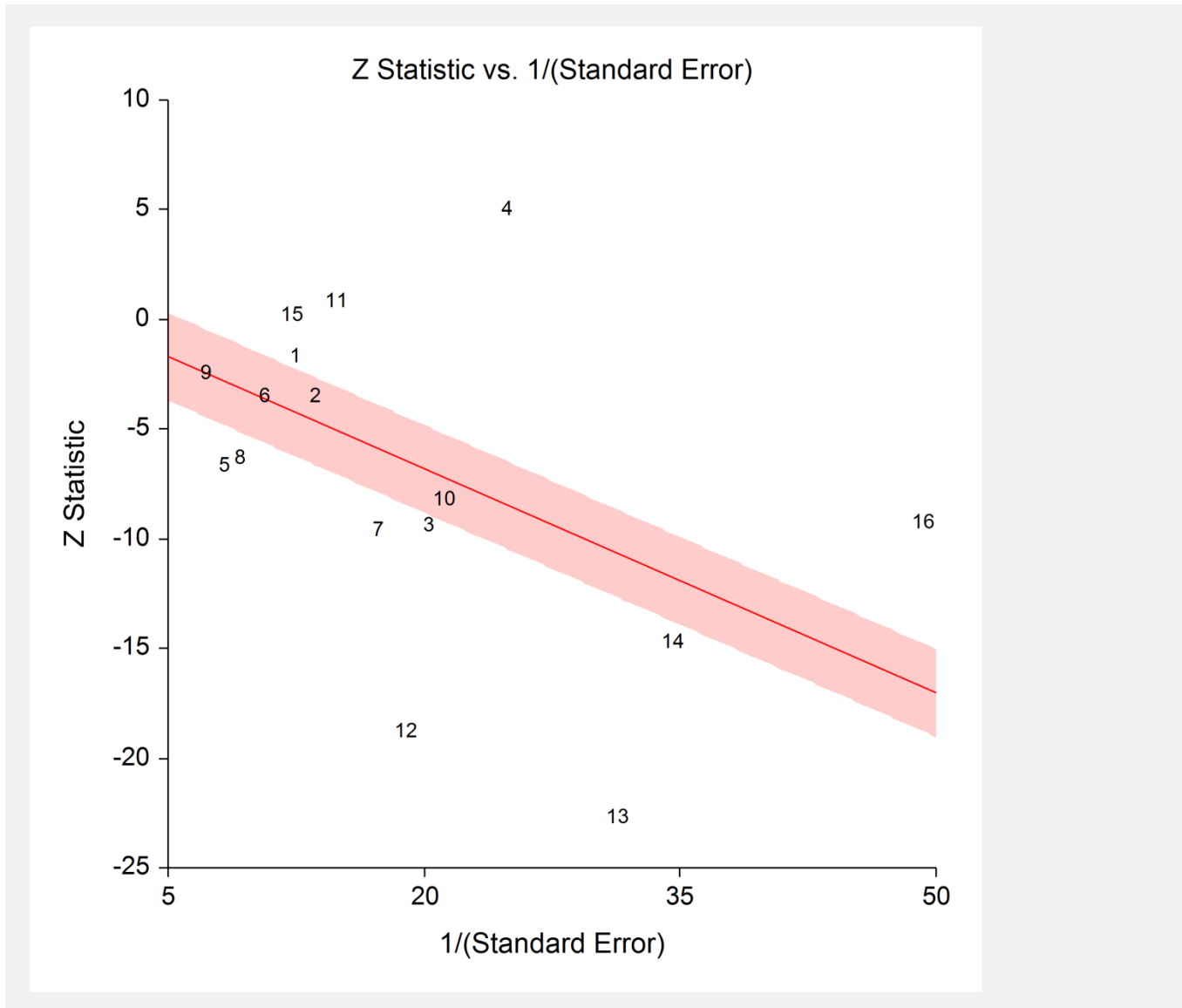
The radial (or Galbraith) plot shows the z-statistic (outcome divided by standard error) on the vertical axis and a measure of weight on the horizontal axis. Studies that have the largest weight are closest to the Y axis. Studies within the limits are interpreted as homogeneous. Studies outside the limits may be outliers.