

# Meta-Analysis

## Fixed effect vs. random effects

Michael Borenstein

Larry Hedges

Hannah Rothstein

Section: Fixed effect vs. random effects models.....	4
Overview .....	4
Fixed effect model .....	6
Definition of a combined effect.....	6
Assigning weights to the studies.....	6
Random effects model .....	11
Definition of a combined effect.....	11
Assigning weights to the studies.....	12
Fixed effect vs. random effects models .....	22
The concept.....	22
Definition of a combined effect.....	22
Computing the combined effect .....	23
Extreme effect size in large study .....	23
Extreme effect size in small study.....	25
Confidence interval width.....	26
Which model should we use? .....	29
Mistakes to avoid in selecting a model .....	30
Example 1 – Binary (2x2) Data.....	32
Start the program and enter the data .....	33
Insert column for study names.....	34
Insert columns for the effect size data .....	35
Enter the data .....	41
Show details for the computations .....	42
Customize the screen .....	47
Display weights.....	48
Compare the fixed effect and random effects models .....	49
Impact of model on study weights.....	50
Impact of model on the combined effect.....	51
Impact of model on the confidence interval .....	52
What would happen if we eliminated Manning?.....	53
Additional statistics.....	55
Test of the null hypothesis .....	56
Test of the heterogeneity .....	58
Quantifying the heterogeneity .....	60
High-resolution plots.....	62
Computational details.....	67
Computational details for the fixed effect analysis .....	69
Computational details for the random effects analysis.....	73
Example 2 – Means.....	77
Start the program and enter the data .....	78
Insert column for study names.....	79
Insert columns for the effect size data .....	80
Enter the data .....	86
Show details for the computations .....	87

Run the analysis .....	89
Customize the screen .....	91
Display weights .....	92
Compare the fixed effect and random effects models .....	93
Impact of model on study weights .....	94
Impact of model on the combined effect .....	95
Impact of model on the confidence interval .....	96
What would happen if we eliminated Manning? .....	97
Additional statistics .....	99
Test of the null hypothesis .....	100
Test of the heterogeneity .....	101
Quantifying the heterogeneity .....	102
High-resolution plots .....	104
Computational details .....	109
Computational details for the fixed effect analysis .....	109
Computational details for the random effects analysis .....	113
Example 3 – Correlational Data .....	117
Start the program and enter the data .....	118
Insert column for study names .....	119
Insert columns for the effect size data .....	120
Enter the data .....	125
What are Fisher’s Z values .....	126
Customize the screen .....	133
Display weights .....	134
Compare the fixed effect and random effects models .....	135
Impact of model on study weights .....	136
Impact of model on the combined effect .....	137
Impact of model on the confidence interval .....	138
What would happen if we eliminated Manning? .....	139
Additional statistics .....	141
Test of the null hypothesis .....	142
Test of the heterogeneity .....	144
Quantifying the heterogeneity .....	146
High-resolution plots .....	148
Computational details .....	153
Computational details for the fixed effect analysis .....	155
Computational details for the random effects analysis .....	159

# Section: Fixed effect vs. random effects models

## Overview

One goal of a meta-analysis will often be to estimate the overall, or combined effect.

If all studies in the analysis were equally precise we could simply compute the mean of the effect sizes. However, if some studies were more precise than others we would want to assign more weight to the studies that carried more information. This is what we do in a meta-analysis. Rather than compute a simple mean of the effect sizes we compute a weighted mean, with more weight given to some studies and less weight given to others.

The question that we need to address, then, is how the weights are assigned. It turns out that this depends on what we mean by a “combined effect”. There are two models used in meta-analysis, the fixed effect model and the random effects model. The two make different assumptions about the nature of the studies, and these assumptions lead to different definitions for the combined effect, and different mechanisms for assigning weights.

### Definition of the combined effect

Under the fixed effect model we assume that there is one true effect size which is shared by all the included studies. It follows that the combined effect is our estimate of this common effect size.

By contrast, under the random effects model we allow that the true effect could vary from study to study. For example, the effect size might be a little higher if the subjects are older, or more educated, or healthier; or if the study used a slightly more intensive or longer variant of the intervention; or if the effect was measured more reliably; and so on. The studies included in the meta-analysis are assumed to be a random sample of the relevant distribution of effects, and the combined effect estimates the mean effect in this distribution.

### Computing a combined effect

Under the fixed effect model all studies are estimating the same effect size, and so we can assign weights to all studies based entirely on the amount of information captured by that study. A large study would be given the lion's share of the weight, and a small study could be largely ignored.

By contrast, under the random effects model we are trying to estimate the mean of a distribution of true effects. Large studies may yield more precise estimates than small studies, but each study is estimating a different effect size, and each of these effect sizes serve as a sample from the population whose mean we want to estimate. Therefore, as compared with the fixed effect model, the weights assigned under random effects are more balanced. Large studies are less likely to dominate the analysis and small studies are less likely to be trivialized.

### Precision of the combined effect

Under the fixed effect model the only source of error in our estimate of the combined effect is the random error within studies. Therefore, with a large enough sample size the error will tend toward zero. This holds true whether the large sample size is confined to one study or distributed across many studies.

By contrast, under the random effects model there are two levels of sampling and two levels of error. First, each study is used to estimate the true effect in a specific population. Second, all of the true effects are used to estimate the mean of the true effects. Therefore, our ability to estimate the combined effect precisely will depend on both the number of subjects within studies (which addresses the first source of error) and also the total number of studies (which addresses the second). In other words, even if each study had infinite sample size there would still be uncertainty in our estimate of the mean, since these studies have been sampled from all possible studies.

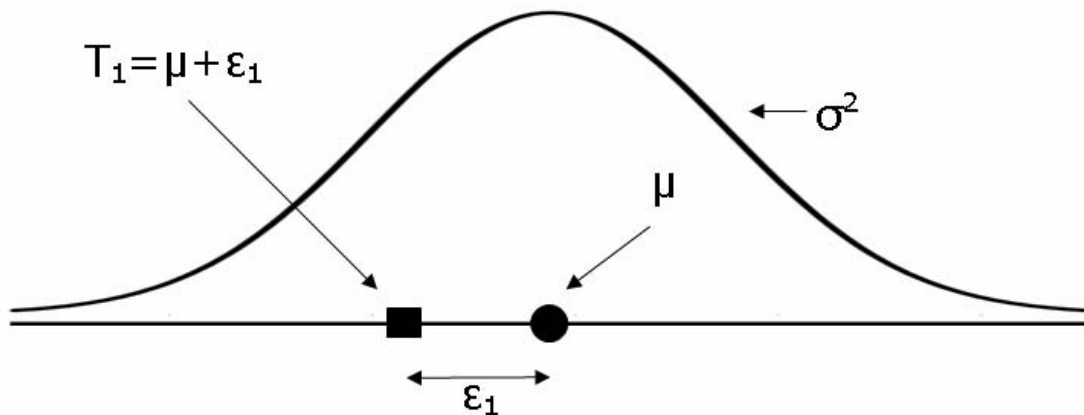
### How this section is organized

The two chapters that follow provide detail on the fixed effect model and the random effects model. These chapters include computational details and worked examples for each model. Then, a chapter highlights the differences between the two.

## Fixed effect model

### Definition of a combined effect

In a fixed effect analysis we assume that all the included studies share a common effect size,  $\mu$ . The observed effects will be distributed about  $\mu$ , with a variance  $\sigma^2$  that depends primarily on the sample size for each study.



Fixed effect model. The observed effects are sampled from a distribution with true effect  $\mu$ , and variance  $\sigma^2$ . The observed effect  $T_1$  is equal to  $\mu + \varepsilon_1$ .

In this schematic the observed effect in Study 1,  $T_1$ , is determined by the common effect  $\mu$  plus the within-study error  $\varepsilon_1$ . More generally, for any observed effect  $T_i$ ,

$$T_i = \mu + \varepsilon_i \quad (0.1)$$

### Assigning weights to the studies

In the fixed effect model there is only one level of sampling, since all studies are sampled from a population with effect size  $\mu$ . Therefore, we need to deal with only one source of sampling error – within studies ( $e$ ).

Since our goal is to assign more weight to the studies that carry more information, we might propose to weight each study by its sample size, so that a study with 1000 subjects would get 10 times the weight of a study with 100 subjects. This is basically the approach used, except that we assign weights based on the inverse of the variance rather than sample size. The inverse variance is roughly proportional to sample size, but is a more nuanced measure (see notes), and serves to minimize the variance of the combined effect.

Concretely, the weight assigned to each study is

$$w_i = \frac{1}{v_i} \quad (0.2)$$

where  $v_i$  is the within-study variance for study ( $i$ ). The weighted mean ( $\bar{T}_\bullet$ ) is then computed as

$$\bar{T}_\bullet = \frac{\sum_{i=1}^k w_i T_i}{\sum_{i=1}^k w_i} \quad (0.3)$$

that is, the sum of the products  $w_i T_i$  (effect size multiplied by weight) divided by the sum of the weights. The variance of the combined effect is defined as the reciprocal of the sum of the weights, or

$$v_\bullet = \frac{1}{\sum_{i=1}^k w_i} \quad (0.4)$$

and the standard error of the combined effect is then the square root of the variance,

$$SE(\bar{T}_\bullet) = \sqrt{v_\bullet} \quad (0.5)$$

The 95% confidence interval for the combined effect would be computed as

$$\text{Lower Limit} = \bar{T}_\bullet - 1.96 * SE(\bar{T}_\bullet) \quad (0.6)$$

$$\text{Upper Limit} = \bar{T}_\bullet + 1.96 * SE(\bar{T}_\bullet) \quad (0.7)$$

Finally, if one were so inclined, the Z-value could be computed using

$$Z = \frac{\bar{T}_i}{SE(\bar{T}_i)} \quad (0.8)$$

For a one-tailed test the p-value would be given by

$$p = 1 - \Phi(Z) \quad (0.9)$$

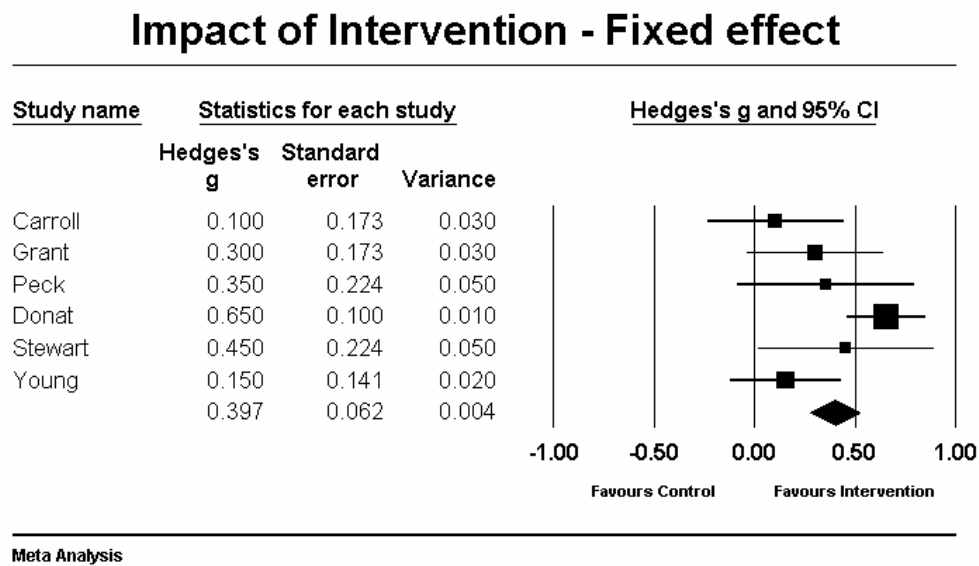
and for a two-tailed test by

$$p = 2[1 - (\Phi(|Z|))] \quad (0.10)$$

where  $\Phi(Z)$  is the standard normal cumulative distribution function.

### Illustrative example

The following figure is the forest plot of a fictional meta-analysis that looked at the impact of an intervention on reading scores in children.



In this example the Carroll study has a variance of 0.03. The weight for that study would be computed as

$$w_1 = \frac{1}{(0.03)} = 33.333$$



and so on for the other studies. Then,

$$\bar{T}_{\bullet} = \frac{101.833}{256.667} = 0.3968$$

$$v_{\bullet} = \frac{1}{256.667} = 0.0039$$

$$SE(\bar{T}_{\bullet}) = \sqrt{0.0039} = 0.0624$$

$$\text{Lower Limit} = 0.3968 - 1.96 * 0.0624 = 0.2744$$

$$\text{Upper Limit} = 0.3968 + 1.96 * 0.0624 = 0.5191$$

$$Z = \frac{0.3968}{0.0624} = 6.3563$$

$$p_{1T} = 1 - \Phi(6.3563) < .0001$$

$$p_{2T} = 2[1 - (\Phi(|6.3563|))] < .0001$$

The fixed effect computations are shown in this spreadsheet

	A	B	C	D	E	F
1	<b>(1) Data</b>			<b>(2) Fixed Effect</b>		
2	Study	ES	Variance	Variance Within	WT	ES*WT
3	Carroll	0.10	0.03	0.030	33.333	3.333
4	Grant	0.30	0.03	0.030	33.333	10.000
5	Peck	0.35	0.05	0.050	20.000	7.000
6	Donat	0.65	0.01	0.010	100.000	65.000
7	Stewart	0.45	0.05	0.050	20.000	9.000
8	Young	0.15	0.02	0.020	50.000	7.500
9	Sum				256.667	101.833
10						
11						
12				<b>(3) Fixed Effect</b>		
13				Effect size		0.3968
14				Variance		0.0039
15				Standard error		0.0624
16				95% Lower limit		0.2744
17				95% Upper limit		0.5191
18				Z-value		6.3563
19				p-value (1-tailed)		0.0000
20				p-value (2-tailed)		0.0000

Make above into table, not Excel

Incorporate formula refs directly into the table above

Column (Cell)	Label	Content	Excel Formula*	See formula
<b>(Section 1) Effect size and weights for each study</b>				
A	Study name	Entered		
B	Effect size	Entered		
C	Variance	Entered		
<b>(Section 2) Compute WT and WT*ES for each study</b>				
D	Variance within study	=C3		
E	Weight	=1/D3		(0.2)
F	ES*WT	=B3*E3		
<b>Sum the columns</b>				
E9	Sum of WT	=SUM(E3:E8)		
F9	Sum of WT*ES	=SUM(F3:F8)		
<b>(Section 3) Compute combined effect and related statistics</b>				
F13	Effect size	=F9/E9		(0.3)
F14	Variance	=1/E9		(0.4)
F15	Standard error	=SQRT(F14)		(0.5)
F16	95% lower limit	=F13-1.96*F15		(0.6)
F17	95% upper limit	=F13+1.96*F15		(0.7)
F18	Z-value	=F13/F15		(0.8)
F19	p-value (1-tailed)	=(1-(NORMDIST((F18),0,1,TRUE)))		(0.9)
F20	p-value (2-tailed)	=(1-(NORMDIST(ABS(F18),0,1,TRUE)))*2		(0.10)

## Comments

Some formulas include a “\$”. In Excel this means that the reference is to a specific column. These are not needed here, but will be needed when we expand this spreadsheet in the next chapter to allow for other computational models.

### Inverse variance vs. sample size.

As noted, weights are based on the inverse variance rather than the sample size. The inverse variance is determined primarily by the sample size, but it is a more nuanced measure. For example, the variance of a mean difference takes account not only of the total N, but also the sample size in each group. Similarly, the variance of an odds ratio is based not only on the total N but also on the number of subjects in each cell.

The combined mean computed with inverse variance weights will have the smallest possible variance.

## Random effects model

The fixed effect model, discussed above, starts with the assumption that the true effect is the same in all studies. However, this assumption may be implausible in many systematic reviews. When we decide to incorporate a group of studies in a meta-analysis we assume that the studies have enough in common that it makes sense to synthesize the information. However, there is generally no reason to assume that they are “identical” in the sense that the true effect size is exactly the same in all the studies.

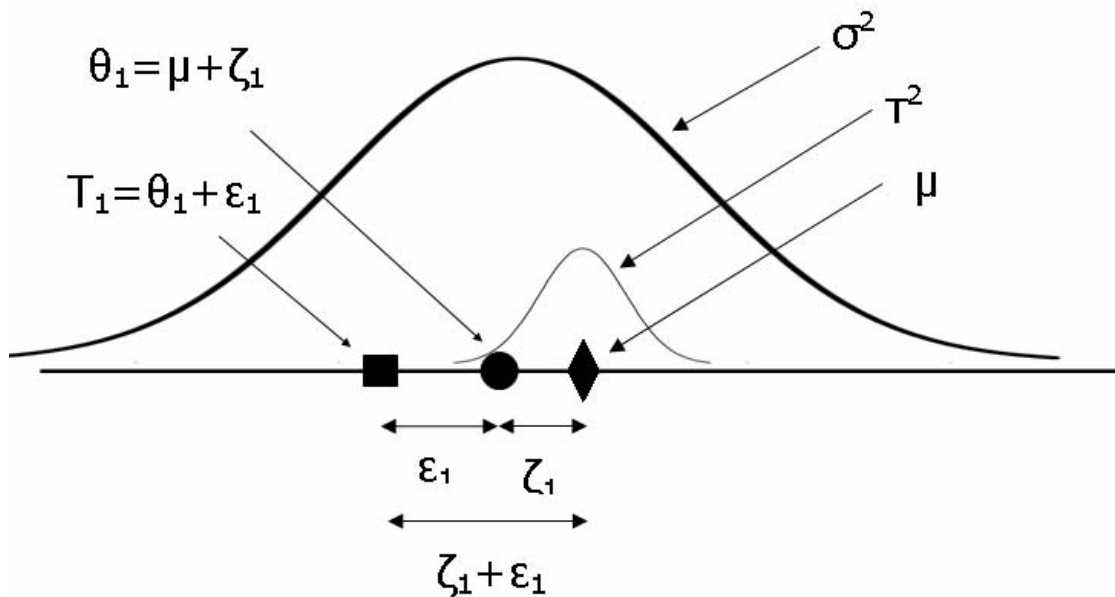
For example, assume that we are working with studies that compare the proportion of patients developing a disease in two groups (vaccination vs. placebo). If the treatment works we would expect the effect size (say, the risk ratio) to be similar but not identical across studies. The impact of the treatment might be more pronounced in studies where the patients were older, or where they had less natural immunity.

Or, assume that we are working with studies that assess the impact of an educational intervention. The magnitude of the impact might vary depending on the other resources available to the children, the class size, the age, and other factors, which are likely to vary from study to study.

We might not have assessed these covariates in each study. Indeed, we might not even know what covariates actually are related to the size of the effect. Nevertheless, experience says that such factors exist and may lead to variations in the magnitude of the effect.

### Definition of a combined effect

Rather than assume that there is one true effect, we allow that there is a distribution of true effect sizes. The combined effect therefore cannot represent the one common effect, but instead represents the mean of the population of true effects.



Random effects model. The observed effect  $T_1$  (box) is sampled from a distribution with true effect  $\theta_1$ , and variance  $\sigma^2$ . This true effect  $\theta_1$ , in turn, is sampled from a distribution with mean  $\mu$  and variance  $\tau^2$ .

In this schematic the observed effect in Study 1,  $T_1$ , is determined by the true effect  $\theta_1$  plus the within-study error  $\epsilon_1$ . In turn,  $\theta_1$ , is determined by the mean of all true effects,  $\mu$  and the between-study error  $\zeta_1$ . More generally, for any observed effect  $T_i$ ,

$$T_i = \theta_i + \epsilon_i = \mu + \zeta_i + \epsilon_i \tag{1.1}$$

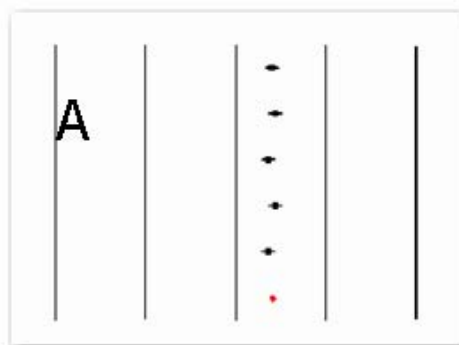
Assigning weights to the studies

Under the random effects model we need to take account of two levels of sampling, and two source of error. First, the true effect sizes  $\theta$  are distributed about  $\mu$  with a variance  $\tau^2$  that reflects the actual distribution of the true effects about their mean. Second, the observed effect  $T$  for any given  $\theta$  will be distributed about that  $\theta$  with a variance  $\sigma^2$  that depends primarily on the sample size for that study. Therefore, in assigning weights to estimate  $\mu$ , we need to deal with both sources of sampling error – within studies ( $\epsilon$ ), and between studies ( $\zeta$ ).

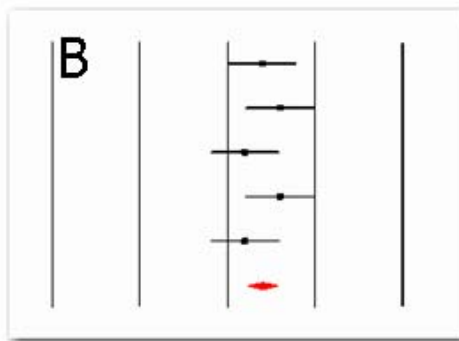
### Decomposing the variance

The approach of a random effects analysis is to decompose the observed variance into its two component parts, within-studies and between-studies, and then use both parts when assigning the weights. The goal will be to take account of both sources of imprecision.

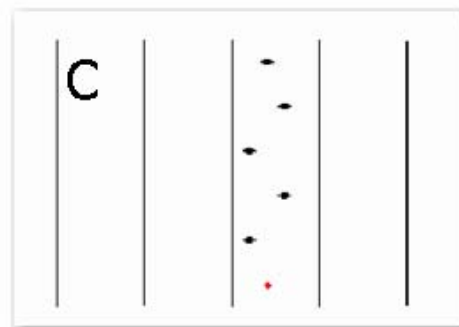
The mechanism used to decompose the variance is to compute the total variance (which is observed) and then to isolate the within-studies variance. The difference between these two values will give us the variance between-studies, which is called tau-squared ( $\tau^2$ ). Consider the three graphs in the following figure.



Between-studies variance is low because total variance is low



Between-studies variance is low because within-studies variance is high



Between-studies variance is high because total variance is high And within-studies variance is low

In (A) the studies all line up pretty much in a row. There is no variance between studies, and therefore tau-squared is low (or zero).

In (B) there is variance between studies, but it is fully explained by the variance within studies. Put another way, given the imprecision of the studies, we would expect the effect size to vary somewhat from one study to the next. Therefore, the between-studies variance is again low (or zero).

In (C) there is variance between studies. And, it cannot be fully explained by the variance within studies, since the within-study variance is minimal. The excess variation (between-studies variance), will be reflected in the value of tau-squared.

It follows that tau-squared will increase as either the variance within-studies decreases and/or the observed variance increases.

This logic is operationalized in a series of formulas. We will compute  $Q$ , which represents the total variance, and  $df$ , which represents the expected variance if all studies have the same true effect. The difference,  $Q - df$ , will give us the excess variance. Finally, this value will be transformed, to put it into the same scale as the within-study variance. This last value is called tau-squared ( $\tau^2$ ).

The  $Q$  statistic represents the total variance and is defined as

$$Q = \sum_{i=1}^k w_i (T_i - \bar{T})^2 \quad (1.2)$$

that is, the sum of the squared deviations of each study ( $T_i$ ) from the combined mean ( $\bar{T}$ ). Note the “ $w_i$ ” in the formula, which indicates that each of the squared deviations is weighted by the study’s inverse variance. A large study that falls far from the mean will have more impact on  $Q$  than would a small study in the same location. An equivalent formula, useful for computations, is

$$Q = \sum_{i=1}^k w_i T_i^2 - \frac{\left( \sum_{i=1}^k w_i T_i \right)^2}{\sum_{i=1}^k w_i} \quad (1.3)$$

Since  $Q$  reflects the total variance, it must now be broken down into its component parts. If the only source of variance was within-study error, then the expected value of  $Q$  would be the degrees of freedom ( $df$ ) for the meta-analysis where

$$df = (\text{Number Studies}) - 1 \quad (1.4)$$

This allows us to compute the between-studies variance,  $\tau^2$ , as

$$\tau^2 = \begin{cases} \frac{Q - df}{C} & \text{if } Q > df \\ 0 & \text{if } Q \leq df \end{cases} \quad (1.5)$$

where

$$C = \sum w_i - \frac{\sum w_i^2}{\sum w_i} \quad (1.6)$$

The numerator,  $Q - df$ , is the excess (observed minus expected) variance. The denominator,  $C$ , is a scaling factor that has to do with the fact that  $Q$  is a weighted sum of squares. By applying this scaling factor we ensure that tau-squared is in the same metric as the variance within-studies.

In the running example,

$$Q = 53.208 - \left( \frac{101.833^2}{256.667} \right) = 12.8056$$

$$df = (6 - 1) = 5$$

$$C = 256.667 - \left( \frac{15522.222}{256.667} \right) = 196.1905$$

$$\tau^2 = \frac{12.8056 - 5}{196.1905} = 0.0398$$

### Assigning weights under the random effects model

In the fixed effect analysis each study was weighted by the inverse of its variance. In the random effects analysis, too, each study will be weighted by the inverse of its variance. The difference is that the variance now includes the original (within-studies) variance plus the between-studies variance, tau-squared.

Note the correspondence between the formulas here and those in the previous chapter. We use the same notations, but add a (\*) to represent the random effects version. Concretely, under the random effects model the weight assigned to each study is

$$w_i^* = \frac{1}{v_i^*} \quad (1.7)$$

where  $v_i^*$  is the within-study variance for study ( $i$ ) plus the between-studies variance, tau-squared. That is,

$$v_i^* = v_i + \tau^2 .$$

The weighted mean ( $\bar{T}_{\cdot}^*$ ) is then computed as

$$\bar{T}_{\cdot}^* = \frac{\sum_{i=1}^k w_i^* T_i}{\sum_{i=1}^k w_i^*} \quad (1.8)$$

that is, the sum of the products (effect size multiplied by weight) divided by the sum of the weights.

The variance of the combined effect is defined as the reciprocal of the sum of the weights, or

$$v_{\cdot}^* = \frac{1}{\sum_{i=1}^k w_i^*} \quad (1.9)$$

and the standard error of the combined effect is then the square root of the variance,

$$SE(\bar{T}_{\cdot}^*) = \sqrt{v_{\cdot}^*} \quad (1.10)$$

The 95% confidence interval for the combined effect would be computed as

$$\text{Lower Limit}^* = \bar{T}_{\cdot}^* - 1.96 * SE(\bar{T}_{\cdot}^*) \quad (1.11)$$

$$\text{Upper Limit}^* = \bar{T}_{\cdot}^* + 1.96 * SE(\bar{T}_{\cdot}^*) \quad (1.12)$$

Finally, if one were so inclined, the Z-value could be computed using

$$Z^* = \frac{\bar{T}_{\cdot}^*}{SE(\bar{T}_{\cdot}^*)} \quad (1.13)$$



The one-tailed  $p$ -value (assuming an effect in the hypothesized direction) is given by

$$p^* = 1 - \Phi(Z^*) \quad (1.14)$$

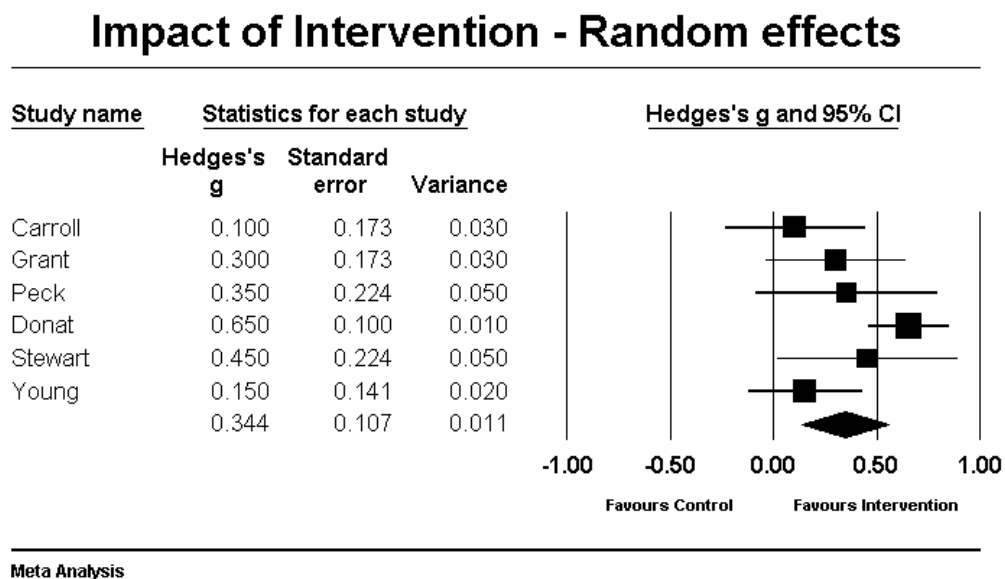
and the two-tailed  $p$ -value by

$$p^* = 2[1 - \Phi(|Z^*|)] \quad (1.15)$$

where  $\Phi(Z)$  is the standard normal cumulative distribution function.

### Illustrative example

The following figure is based on the same studies we used for the fixed effect example.



Note the differences from the fixed effect model.

- The weights are more balanced. The boxes for the large studies such as Donat have decreased in area while those for the small studies such as Peck have increase in area.
- The combined effect has moved toward the left, from 0.40 to 0.34. This reflects the fact that the impact of Donat (on the right) has been reduced.
- The confidence interval for the combined effect has increased in width.

In the running example the weight for the Carroll study would be computed as

$$w_i^* = \frac{1}{(0.030 + 0.040)} = \frac{1}{(0.070)} = 14.330$$

and so on for the other studies. Then,

$$\bar{T}^* = \frac{30.207}{87.747} = 0.3442$$

$$v.^* = \frac{1}{87.747} = 0.0114$$

$$SE(\bar{T}^*) = \sqrt{0.0114} = 0.1068$$

$$\text{Lower Limit}^* = 0.3442 - 1.96 * 0.1068 = 0.1350$$

$$\text{Upper Limit}^* = 0.3968 + 1.96 * 0.1068 = 0.5535$$

$$Z^* = \frac{0.3442}{0.1068} = 3.2247$$

$$P_{1T} = 1 - \Phi(3.2247) = 0.0006$$

$$P_{2T} = [1 - (\Phi(\text{ABS}(3.2247)))] * 2 = 0.0013$$

These formulas are incorporated in the following spreadsheet

	A	B	C	D	E	F		H	I	J	K	L	M
1	<b>(1) Data</b>			<b>(2) Fixed Effect</b>			<b>(4) Compute Tau^2</b>		<b>(6) Random effects</b>				
2	Study	ES	Variance	Variance Within	WT	ES*WT	ES^2*WT	WT^2	Variance Within	Variance Between	Variance Total	WT	ES*WT
3	Carroll	0.10	0.03	0.030	33.333	3.333	0.333	1111.111	0.030	0.040	0.070	14.330	1.433
4	Grant	0.30	0.03	0.030	33.333	10.000	3.000	1111.111	0.030	0.040	0.070	14.330	4.299
5	Peck	0.35	0.05	0.050	20.000	7.000	2.450	400.000	0.050	0.040	0.090	11.138	3.898
6	Donat	0.65	0.01	0.010	100.000	65.000	42.250	10000.000	0.010	0.040	0.050	20.086	13.056
7	Stewart	0.45	0.05	0.050	20.000	9.000	4.050	400.000	0.050	0.040	0.090	11.138	5.012
8	Young	0.15	0.02	0.020	50.000	7.500	1.125	2500.000	0.020	0.040	0.060	16.726	2.509
9	Sum				256.667	101.833	53.208	15522.222				87.747	30.207
10													
11													
12				<b>(3) Fixed Effect</b>			<b>(5) Compute Tau^2</b>				<b>(7) Random Effects</b>		
13				Effect size	0.3968	Q	12.8056			Effect size	0.3442		
14				Variance	0.0039	df	5.0000			Variance	0.0114		
15				Standard error	0.0624	Numerato	7.8056			Standard error	0.1068		
16				95% Lower limit	0.2744	C	196.1905			95% Lower limit	0.1350		
17				95% Upper limit	0.5191	Tau-sq	0.0398			95% Upper limit	0.5535		
18				Z-value	6.3563					Z-value	3.2247		
19				p-value (1-tailed)	0.0000					p-value (1-tailed)	0.0006		
20				p-value (2-tailed)	0.0000					p-value (2-tailed)	0.0013		

This spreadsheet builds on the spreadsheet for a fixed effect analysis. Columns A-F are identical to those in that spreadsheet. Here, we add columns for tau-squared (columns G-H) and random effects analysis (columns I-M).

Note that the formulas for fixed effect and random effects analyses are identical, the only difference being the definition of the variance. For the fixed effect analysis the variance (Column D) is defined as the variance within-studies (for example D3=\$C3). For the random effects analysis the variance is defined as the variance within-studies plus the variance between-studies (for example, K3=I3+J3).

Column (Cell)	Label	Content	Excel Formula*	See formula
<b>(Section 1) Effect size and weights for each study</b>				
A	Study name	Entered		
B	Effect size	Entered		
C	Variance	Entered		
<b>(Section 2) Compute fixed effect WT and WT*ES for each study</b>				
D	Variance within study	=C3		
E	Weight	=1/D3		(0.2)
F	ES*WT	=B3*E3		
<b>Sum the columns</b>				
E9	Sum of WT	=SUM(E3:E8)		
F9	Sum of WT*ES	=SUM(F3:F8)		
<b>(Section 3) Compute combined effect and related statistics for fixed effect model</b>				
F13	Effect size	=F9/E9		(0.3)
F14	Variance	=1/E9		(0.4)
F15	Standard error	=SQRT(F14)		(0.5)
F16	95% lower limit	=F13-1.96*F15		(0.6)
F17	95% upper limit	=F13+1.96*F15		(0.7)
F18	Z-value	=F13/F15		(0.8)
F19	p-value (1-tailed)	=(1-(NORMDIST((F18),0,1,TRUE)))		(0.9)
F20	p-value (2-tailed)	=(1-(NORMDIST(ABS(F18),0,1,TRUE))) <sup>2</sup>		(0.10)
<b>(Section 4) Compute values needed for tau-squared</b>				
G3	ES <sup>2</sup> *WT	=B3 <sup>2</sup> *E3		
H3	WT <sup>2</sup>	=E3 <sup>2</sup>		
<b>Sum the columns</b>				
G9	Sum of ES <sup>2</sup> *WT	=SUM(G3:G8)		
H9	Sum of WT <sup>2</sup>	=SUM(H3:H8)		
<b>(Section 5) Compute tau-squared</b>				
H13	Q	=G9-F9 <sup>2</sup> /E9		(1.3)
H14	Df	=COUNT(B3:B8)-1		(1.4)
H15	Numerator	=MAX(H13-H14,0)		
H16	C	=E9-H9/E9		(1.6)
H17	tau-sq	=H15/H16		(1.5)
<b>(Section 6) Compute random effects WT and WT*ES for each study</b>				
I3	Variance within	=C3		
J3	Variance between	=H\$17		
K3	Variance total	=I3+J3		
L3	WT	=1/K3		(1.7)
M3	ES*WT	=B3*L3		
<b>Sum the columns</b>				
L9	Sum of WT	=SUM(L3:L8)		
M9	Sum of ES*WT	=SUM(M3:M8)		
<b>(Section 7) Compute combined effect and related statistics for random effects model</b>				
M13	Effect size	=M9/L9		(1.8)
M14	Variance	=1/L9		(1.9)
M15	Standard error	=SQRT(M14)		(1.10)
M16	95% lower limit	=M13-1.96*M15		(1.11)

M17	95% upper limit	=M13+1.96*M15	(1.12)
M18	Z-value	=M13/M15	(1.13)
M19	p-value (1-tailed)	=(1-(NORMDIST((M18),0,1,TRUE)))	(1.14)
M20	p-value (2-tailed)	=(1-(NORMDIST(ABS(M18),0,1,TRUE))) <sup>2</sup>	(1.15)

## Fixed effect vs. random effects models

In the previous two chapters we outlined the two basic approaches to meta-analysis – the Fixed effect model and the Random effects model. This chapter will discuss the differences between the two.

### The concept

The fixed effect and random effects models represent two conceptually different approaches.

#### Fixed effect

The fixed effect model assumes that all studies in the meta-analysis share a common true effect size. Put another way, all factors which could influence the effect size are the same in all the study populations, and therefore the effect size is the same in all the study populations. It follows that the observed effect size varies from one study to the next only because of the random error inherent in each study.

#### Random effects

By contrast, the random effects model assumes that the studies were drawn from populations that differ from each other in ways that could impact on the treatment effect. For example, the intensity of the intervention or the age of the subjects may have varied from one study to the next. It follows that the effect size will vary from one study to the next for two reasons. The first is random error within studies, as in the fixed effect model. The second is true variation in effect size from one study to the next.

### Definition of a combined effect

The meaning of the “combined effect” is different for fixed effect vs. random effects analyses.

#### Fixed effect

Under the fixed effect model there is one true effect size. It follows that the combined effect is our estimate of this value.

#### Random effects

Under the random effects model there is not one true effect size, but a distribution of effect sizes. It follows that the combined estimate is not an estimate of one value, but rather is meant to be the average of a distribution of values.

### Computing the combined effect

These differences in the definition of the combined effect lead to differences in the way the combined effect is computed.

#### Fixed effect

Under the fixed effect model we assume that the true effect size for all studies is identical, and the only reason the effect size varies between studies is random error. Therefore, when assigning weights to the different studies we can largely ignore the information in the smaller studies since we have better information about the same effect size in the larger studies.

#### Random effects

By contrast, under the random effects model the goal is not to estimate one true effect, but to estimate the mean of a distribution of effects. Since each study provides information about an effect size in a different population, we want to be sure that all the populations captured by the various studies are represented in the combined estimate.

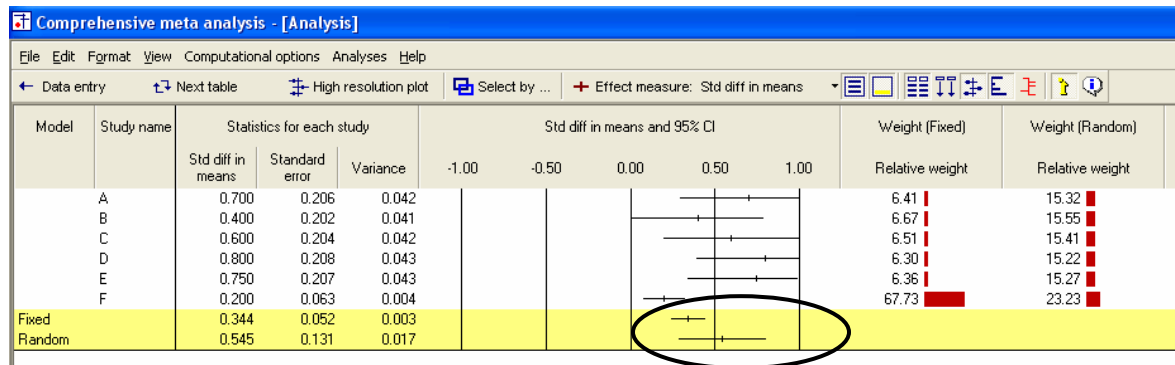
This means that we cannot discount a small study by giving it a very small weight (the way we would in a fixed effect analysis). The estimate provided by that study may be imprecise, but it is information about a population that no other study has captured. By the same logic we cannot give too much weight to a very large study (the way we might in a fixed effect analysis). Our goal is to estimate the effects in a range of populations, and we do not want that overall estimate to be overly influenced by any one population.

### Extreme effect size in large study

How will the selection of a model influence the overall effect size? Consider the case where there is an extreme effect in a large study. Here, we have five small studies (Studies A-E, with 100 subjects per study) and one large study (Study F, with 1000 subjects). The confidence interval for each of the studies A-E is wide, reflecting relatively poor precision, while the confidence interval for Study F is narrow, indicating greater precision. In this example the small studies all have relatively large effects (in the range of 0.40 to 0.80) while the large study has a relatively small effect (0.20).

## Fixed effect

Under the fixed effect model these studies are all estimating the same effect size, and the large study (F) provides a more precise estimate of this effect size. Therefore, this study is assigned 68% of the weight in the combined effect, with each of the remaining studies being assigned about 6% of the weight (see the column labeled “Relative weight” under fixed effects).



Because Study F is assigned so much of the weight it “pulls” the combined estimate toward itself. Study F had a smaller effect than the other studies and so it pulls the combined estimate toward the left. On the graph, note the point estimate for the large study (Study F, with  $d=.2$ ), and how it has “pulled” the fixed effect estimate down to 0.34 (see the shaded row marked “Fixed” at the bottom of the plot).

## Random effects

By contrast, under the random effects model these studies are drawn from a range of populations in which the effect size varies and our goal is to summarize this range of effects. Each study is estimating an effect size for its unique population, and so each must be given appropriate weight in the analysis. Now, Study F is assigned only 23% of the weight (rather than 68%), and each of the small studies is given about 15% of the weight (rather than 6%) (see the column labeled “Relative weights” under random effects).

What happens to our estimate of the combined effect when we weight the studies this way? The overall effect is still being pulled by the large study, but not as much as before. In the plot, the bottom two lines reflect the fixed effect and random effect estimates, respectively. Compare the point estimate for “Random” (the last line) with the one for “Fixed” just above it. The overall effect is now 0.55 (which is much closer to the range of the small studies) rather than 0.34 (as it was for the fixed effect model). The impact of the large study is now less pronounced.

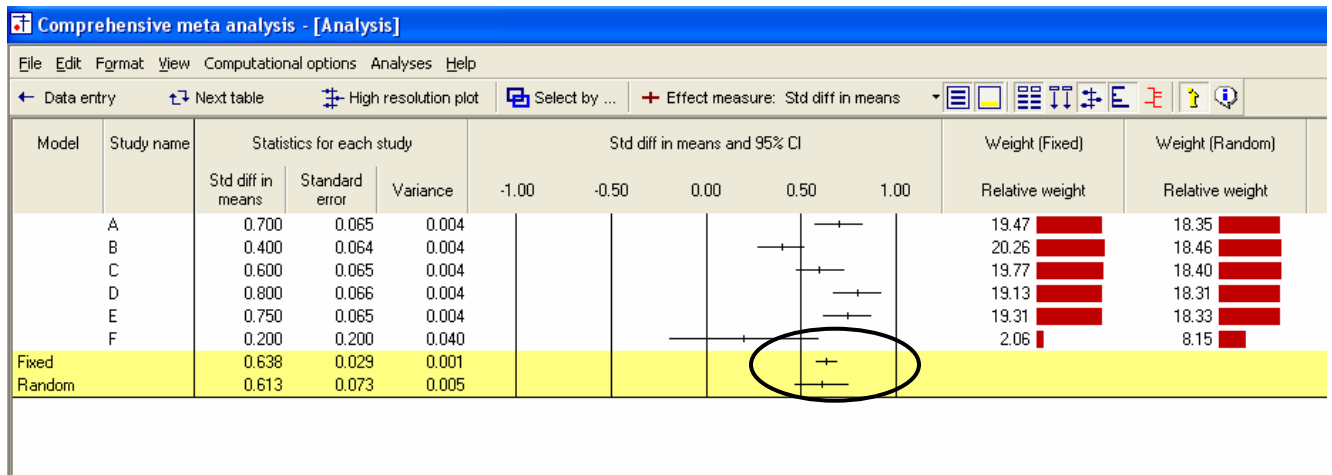


## Extreme effect size in small study

Now, let's consider the reverse situation: The effect sizes for each study are the same as in the prior example, but this time the first 5 studies are large while the sixth study is small. Concretely, we have five large studies (A-E, with 1000 subjects per study) and one small study (F, with 100 subjects). On the graphic, the confidence intervals for studies A-E are each relatively narrow, indicating high precision, while that for Study F is relatively wide, indicating less precision. The large studies all have relatively large effects (in the range of 0.40 to 0.80) while the small study has a relatively small effect (0.20).

### Fixed effect

Under the fixed effect model the large studies (A-E) are each assigned about 20% of the weight, while the small study (F) is assigned only about 2% of the weight (see column labeled "Relative weights" under Fixed effect). This follows from the logic of the fixed effect model. The larger studies provide a good estimate of the common effect, and the small study offers a less reliable estimate of that same effect, so it is assigned a small (in this case trivial) weight. With only 2% of the weight, Study F has little impact on the combined value, which is computed as 0.64.



### Random effects

By contrast, under the random effects model each study is estimating an effect size for its unique population, and so each must be assigned appropriate weight in the analysis. As shown in the column "Relative weights" under random effects each of the large studies (A-E) is now assigned about 18% of the weight (rather than 20%) while the small study (F) receives 8% of the weight (rather than 2%).

What happens to our estimate of the combined effect when we weight the studies this way? Where the small study has almost no impact under the fixed effect model, it now has a substantially larger impact. Concretely, it gets 8% of the weight, which is nearly half the weight assigned to any of the larger studies (18%).

The small study therefore has more of an impact now than it did under the fixed effect model. Where it was assigned only 2% of the weight before, it is now assigned 8% of the weight. This is 50% of the weight assigned to studies A-E, and as such is no longer a trivial amount. Compare the two lines labeled “Fixed” and “Random” at the bottom of the plot. The overall effect is now 0.61, which is .03 points closer to study F than it had been under the fixed effect model (0.64).

## Summary

The operating premise, as illustrated in these examples, is that the relative weights assigned under random effects will be more balanced than those assigned under fixed effects. As we move from fixed effect to random effects, extreme studies will lose influence if they are large, and will gain influence if they are small.

In these two examples we included a single study with an extreme size and an extreme effect, to highlight the difference between the two weighting schemes. In most analyses, of course, there will be a range of sample sizes within studies and the larger (or smaller) studies could fall anywhere in this range. Nevertheless, the same principle will hold.

## Confidence interval width

Above, we considered the impact of the model (fixed vs. random effects) on the combined effect size. Now, let’s consider the impact on the width of the confidence interval.

Recall that the fixed effect model defines “variance” as the variance within a study, while the random effects model defines it as variance within a study plus variance between studies. To understand how this difference will affect the width of the confidence interval, let’s consider what would happen if all studies in the meta-analysis were of infinite size, which means that the within-study error is effectively zero.

### Fixed effect

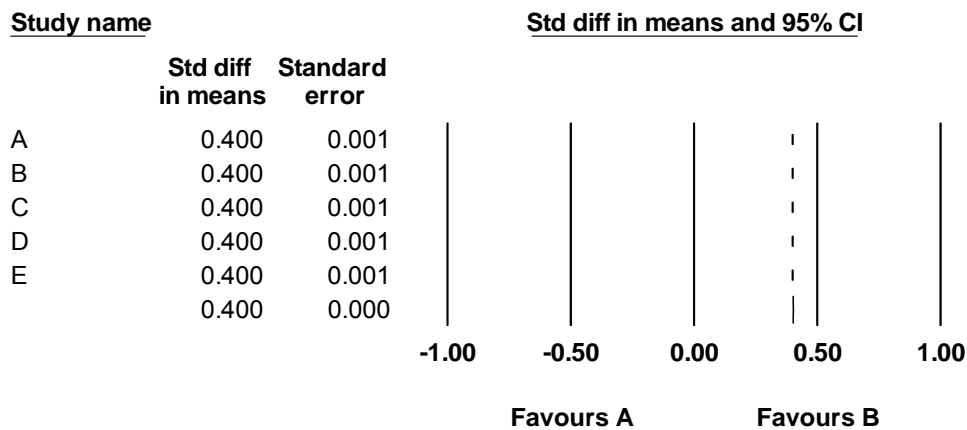
Since we’ve started with the assumption that all variation is due to random error, and this error has now been removed, it follows that

- The observed effects would all be identical.

- The combined effect would be exactly the same as each of the individual studies.
- The width of the confidence interval for the combined effect would approach zero.

All of these points can be seen in the figure. In particular, note that the diamond representing the combined effect has a width of zero, since the width of the confidence interval is zero.

### Fixed effect model with huge N



Meta Analysis

Generally, we are concerned with the precision of the combined effect rather than the precision of the individual studies. For this purpose it doesn't matter whether the sample is concentrated in one study or dispersed among many studies. In either case, as the total N approaches infinity the errors will cancel out and the standard error will approach zero.

#### Random effects

Under the random effects model the effect size for each study would still be known precisely. However, the effects would not line up in a row since the true treatment effect is assumed to vary from one study to the next. It follows that –

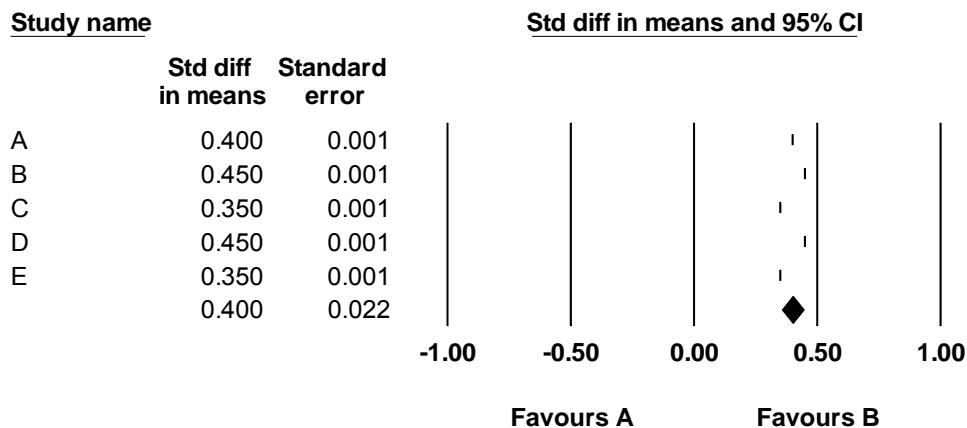
- The within-study error would approach zero, and the width of the confidence interval for each study would approach zero.

- Since the studies are all drawn from different populations, even though the effects are now being estimated without error, the observed effects would not be identical to each other.
- The width of the confidence interval for the combined effect would not approach zero unless the number of studies approached infinity.

Generally, we are concerned with the precision of the combined effect rather than the precision of the individual studies. Under the random effects model we need an infinite number of studies in order for the standard error in estimating  $\mu$  to approach zero. In our example we know the value of the five effects precisely, but these are only a random sample of all possible effects, and so there remains substantial error in our estimate of the mean effect.

Note. While the distribution of the  $\theta_i$  about  $\mu$  represents a real distribution of effect sizes, we nevertheless refer to this as “error” since it introduces error into our estimate of the mean effect. If the studies that we do observe tend to cluster closely together and/or our meta-analysis includes a large number of studies, this source of error will tend to be small. If the studies that we do observe show much dispersion and/or we have only a small sample of studies, then this source of error will tend to be large.

## Random effects model with huge N



Meta Analysis

## Summary

Since the variation under random effects incorporates the same error as fixed effects plus an additional component, it cannot be less than the variation under the fixed effect model. As long as the between-studies variation is non-zero, the variance, standard error, and confidence interval will always be larger under random effects.

The standard error of the combined effect in both models is inversely proportional to the number of studies. Therefore, in both models, the width of the confidence interval tends toward zero as the number of studies increases. In the case of the fixed effect model the standard error and the width of the confidence interval can tend toward zero even with a finite number of studies if any of the studies is sufficiently large. By contrast, for the random effects model, the confidence interval can tend toward zero only with an infinite number of studies (unless the between-study variation is zero).

### Which model should we use?

The selection of a computational model should be based on the nature of the studies and our goals.

#### Fixed effect

The fixed effect model makes sense if (a) there is reason to believe that all the studies are functionally identical, and (b) our goal is to compute the common effect size, which would then be generalized to other examples of this same population.

For example, assume that a drug company has run five studies to assess the effect of a drug. All studies recruited patients in the same way, used the same researchers, dose, and so on, so all are expected to have the identical effect (as though this were one large study, conducted with a series of cohorts). Also, the regulatory agency wants to see if the drug works in this one population. In this example, a fixed effect model makes sense.

#### Random effects

By contrast, when the researcher is accumulating data from a series of studies that had been performed by other people, it would be unlikely that all the studies were functionally equivalent. Typically, the subjects or interventions in these studies would have differed in ways that would have impacted on the results, and therefore we should not assume a common effect size. Therefore, in these

cases the random effects model is more easily justified than the fixed effect model.

Additionally, the goal of this analysis is usually to generalize to a range of populations. Therefore, if one did make the argument that all the studies used an identical, narrowly defined population, then it would not be possible to extrapolate from this population to others, and the utility of the analysis would be limited.

#### Note

If the number of studies is very small, then it may be impossible to estimate the between-studies variance ( $\tau^2$ ) with any precision. In this case, the fixed effect model may be the only viable option. In effect, we would then be treating the included studies as the only studies of interest.

#### Mistakes to avoid in selecting a model

Some have adopted the practice of starting with the fixed effect model and then moving to a random effects model if  $Q$  is statistically significant. This practice should be discouraged for the following reasons.

- If the logic of the analysis says that the study effect sizes have been sampled from a distribution of effect sizes then the random effects formula, which reflects this idea, is the logical one to use.
- If the actual dispersion turns out to be trivial (that is, less than expected under the hypothesis of homogeneity), then the random effects model will reduce to the fixed effect model. Therefore, there is no “cost” to using the random effects model in this case.
- If the actual dispersion turns out to be non-trivial, then this dispersion should be incorporated in the analysis, which the random effects model does, and the fixed effect model does not. That the  $Q$  statistic meets or does not meet a criterion for significance is simply not relevant.

The last statement above would be true even if the  $Q$  test did a good job of identifying dispersion. In fact, though, if the number of studies is small and the within-studies variance is large, the test based on the  $Q$  statistic may have low power even if the between-study variance is substantial. In this case, using the  $Q$  test as a criterion for selecting the model is problematic not only from a conceptual perspective, but could also lead to the use of a fixed effect analysis in cases with substantial dispersion.





## Example 1 – Binary (2x2) Data

This appendix shows how to use Comprehensive Meta-Analysis (CMA) to perform a meta-analysis for odds ratios using fixed and random effects models.

To download a free trial copy of CMA go to [www.Meta-Analysis.com](http://www.Meta-Analysis.com)

If your trial copy has expired, simply ask for a free extension

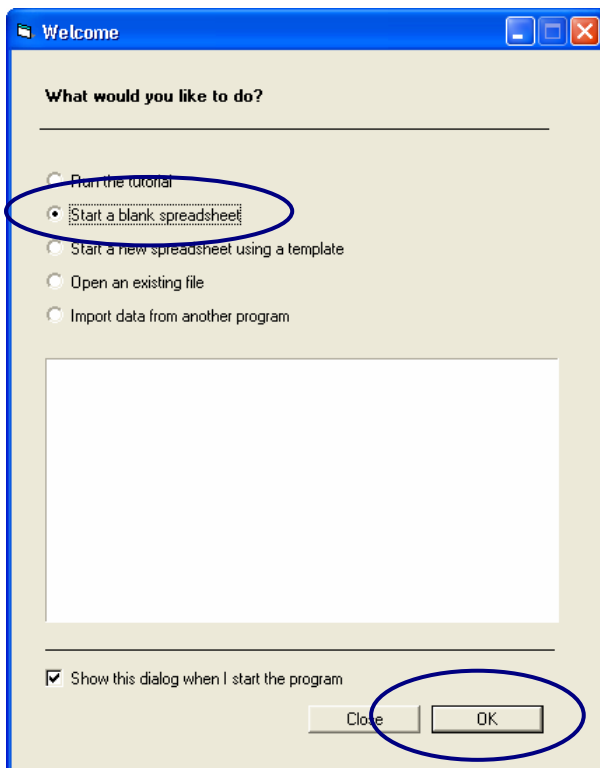
- Start the program
- Click “I want to get an unlock code”, to create an e-mail
- In the e-mail, ask for an extension
- We’ll send a code by return e-mail



## Start the program and enter the data

→ Start CMA

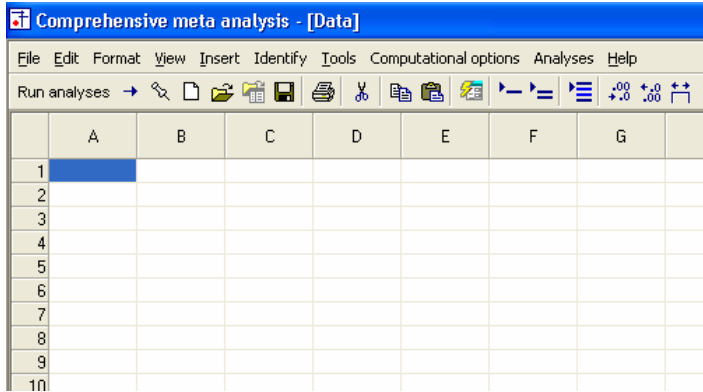
The program shows this dialog box.



→ Select START A BLANK SPREADSHEET

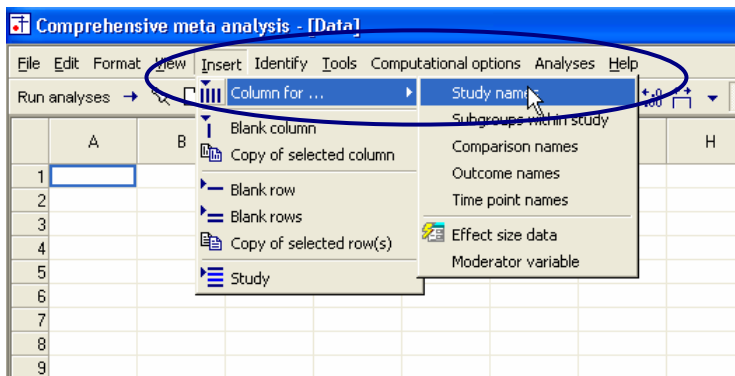
→ Click OK

The program displays this screen.

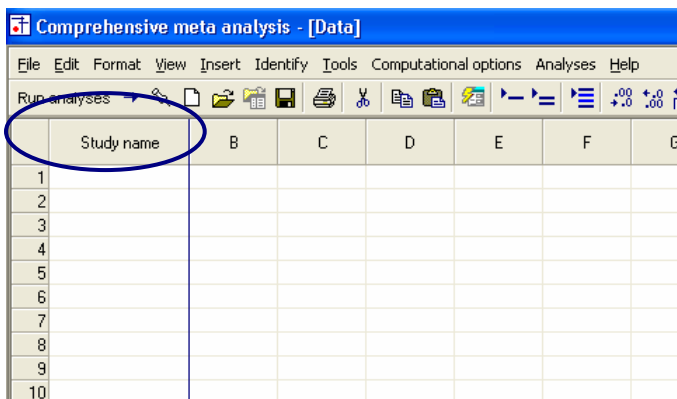


### Insert column for study names

→ Click INSERT > COLUMN FOR > STUDY NAMES



The program has added a column for Study names.

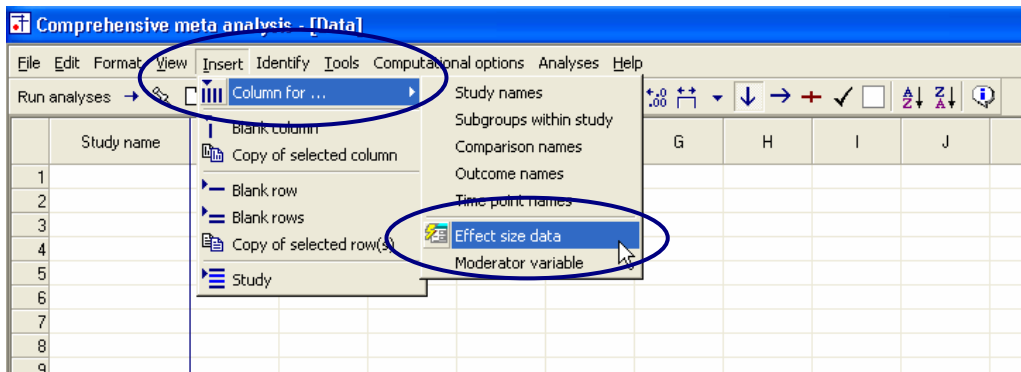


## Insert columns for the effect size data

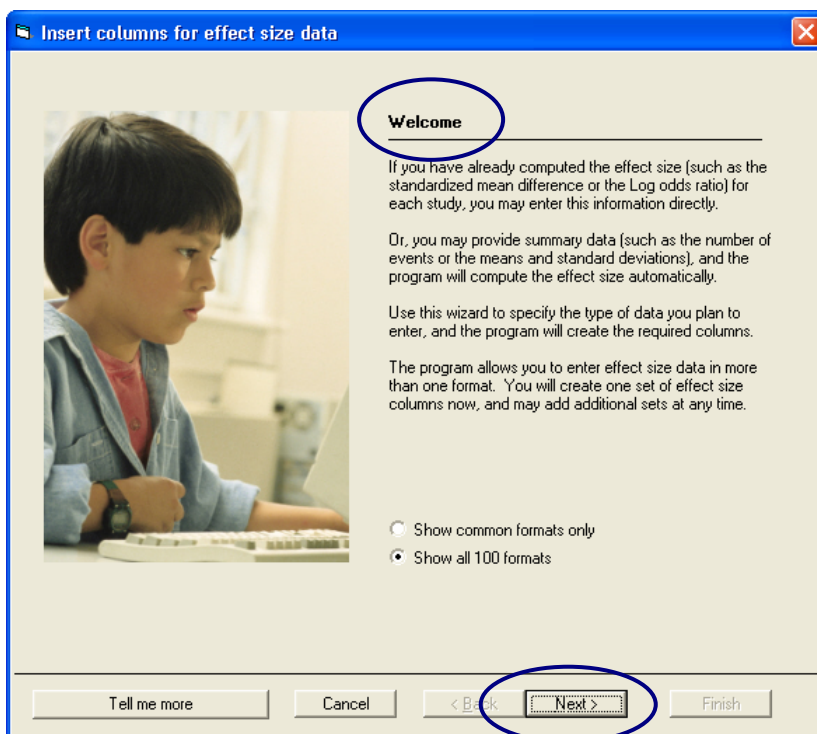
Since CMA will accept data in more than 100 formats, you need to tell the program what format you want to use.

You do have the option to use a different format for each study, but for now we'll start with one format.

→ Click INSERT > COLUMN FOR > EFFECT SIZE DATA

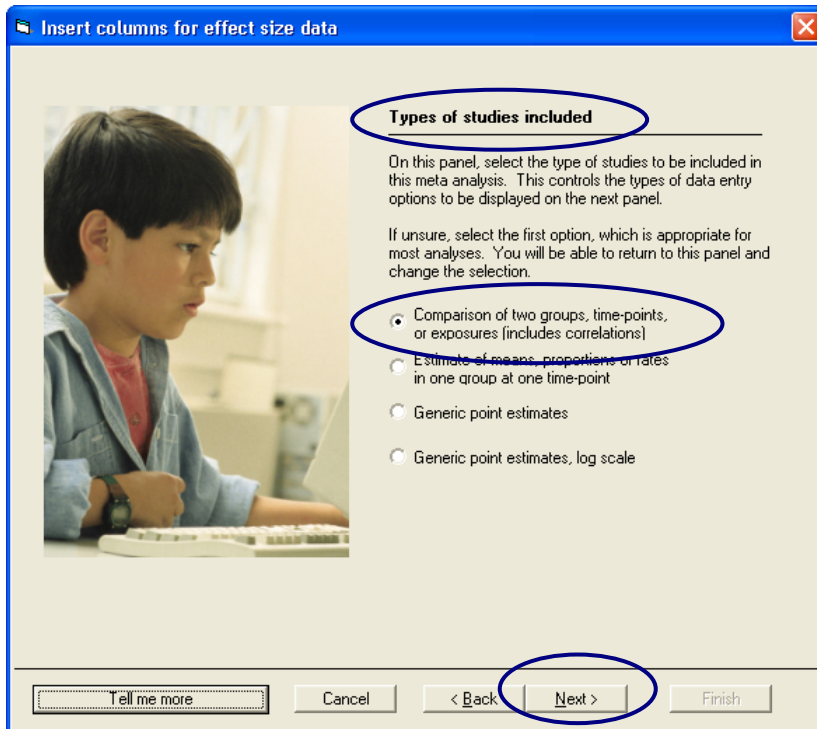


The program shows this dialog box.



→ Click NEXT

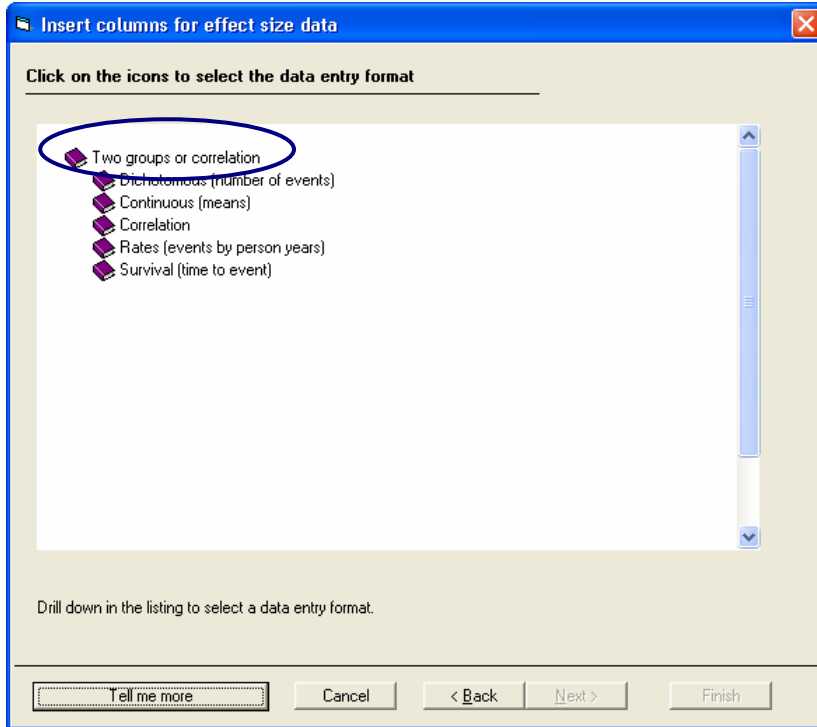
The dialog box lists four sets of effect sizes.



→ Select COMPARISON OF TWO GROUPS, TIME-POINTS, OR EXPOSURES (INCLUDES CORRELATIONS)

→ Click NEXT

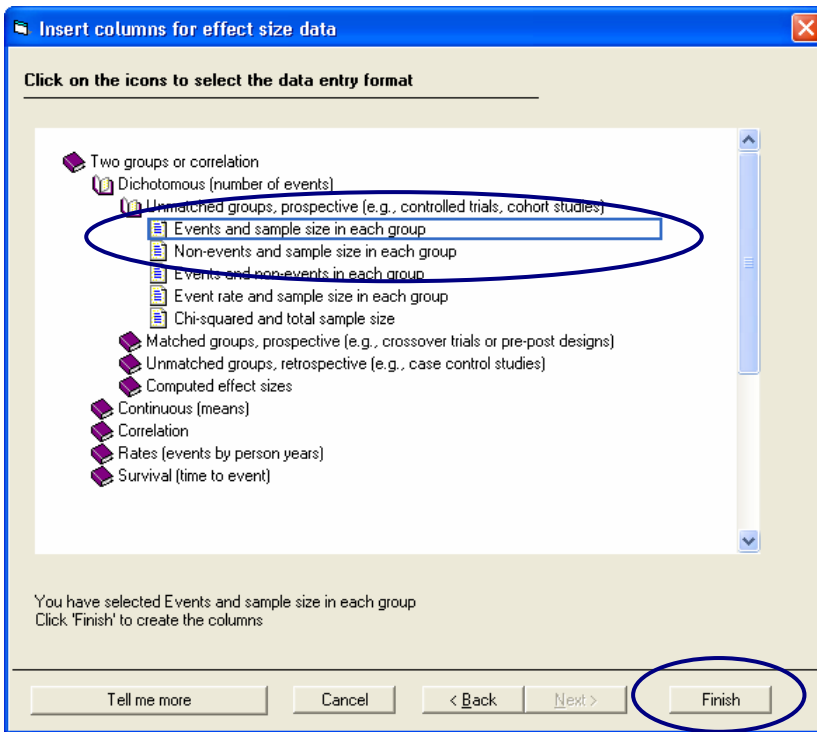
The program displays this dialog box.



→ Drill down to

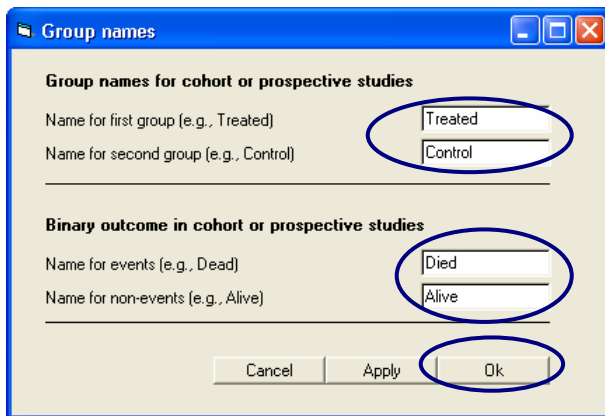
- DICHOTOMOUS (NUMBER OF EVENTS)
- UNMATCHED GROUPS, PROSPECTIVE (E.G. CONTROLLED TRIALS, COHORT STUDIES)
- EVENTS AND SAMPLE SIZE IN EACH GROUP

→ Click FINISH



The program will return to the main data-entry screen.

The program displays a dialog box that you can use to name the groups.



- Enter the names TREATED and CONTROL for the group
- Enter DIED and ALIVE for the outcome
- Click OK

The program displays the columns needed for the selected format (Treated Died, Treated Total N, Control Died, Control Total N).

You will enter data into the white columns (at left). The program will compute the effect size for each study and display that effect size in the yellow columns (at right).

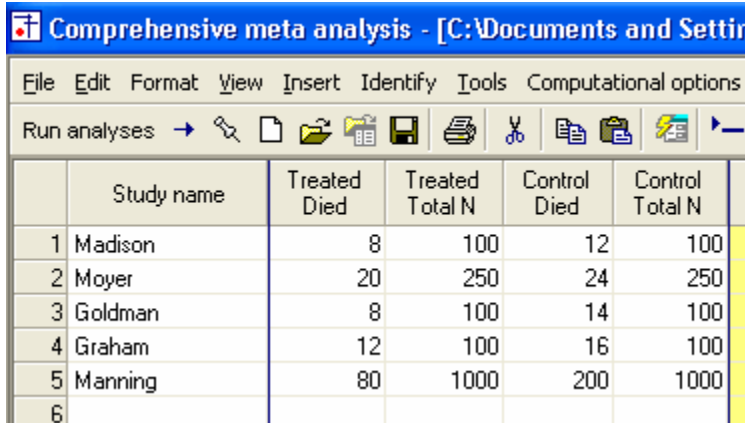
Since you elected to enter events and sample size, the program initially displays columns for the odds ratio and the log odds ratio. You can add other indices as well.

	Study name	Treated Died	Treated Total N	Control Died	Control Total N	Odds ratio	Log odds ratio	Std Err	Variance	J
1										
2										
3										
4										
5										
6										
7										



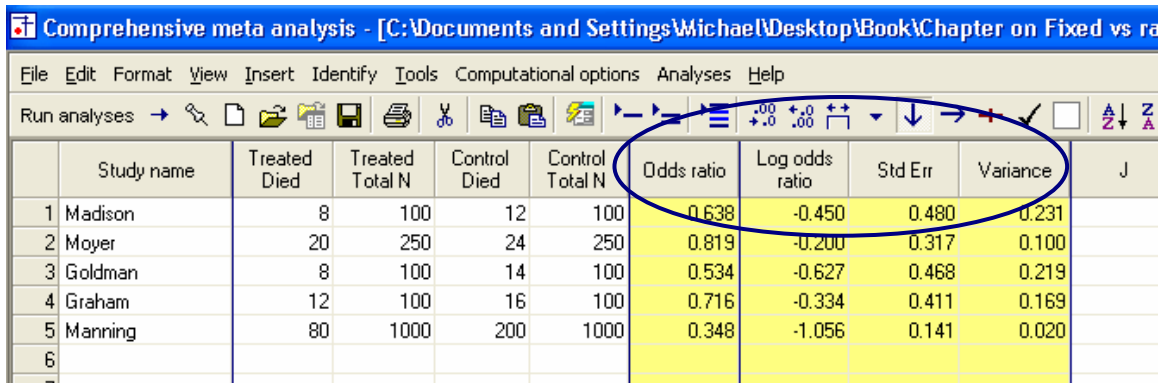
## Enter the data

→ Enter the events and total N for each group as shown here



	Study name	Treated Died	Treated Total N	Control Died	Control Total N
1	Madison	8	100	12	100
2	Moyer	20	250	24	250
3	Goldman	8	100	14	100
4	Graham	12	100	16	100
5	Manning	80	1000	200	1000
6					

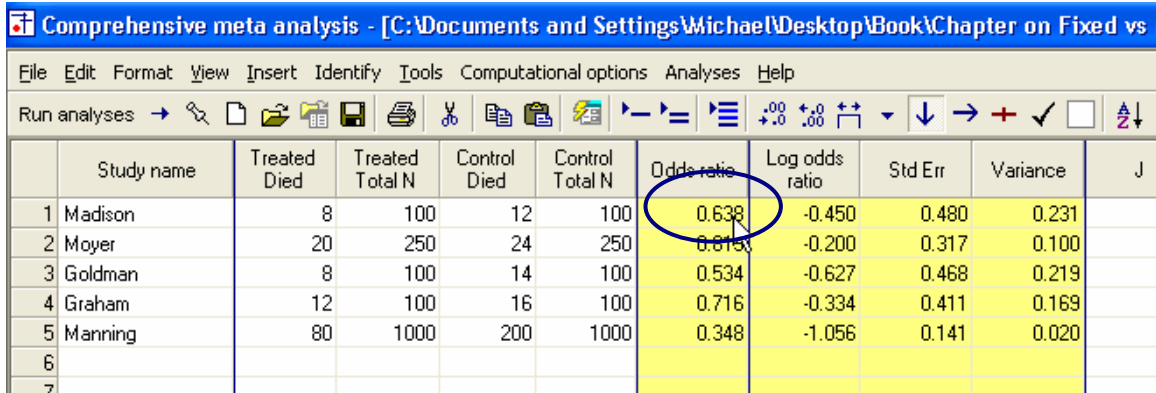
The program will automatically compute the effects as shown here in the yellow columns.



	Study name	Treated Died	Treated Total N	Control Died	Control Total N	Odds ratio	Log odds ratio	Std Err	Variance	J
1	Madison	8	100	12	100	0.638	-0.450	0.480	0.231	
2	Moyer	20	250	24	250	0.819	-0.200	0.317	0.100	
3	Goldman	8	100	14	100	0.534	-0.627	0.468	0.219	
4	Graham	12	100	16	100	0.716	-0.334	0.411	0.169	
5	Manning	80	1000	200	1000	0.348	-1.056	0.141	0.020	
6										
7										

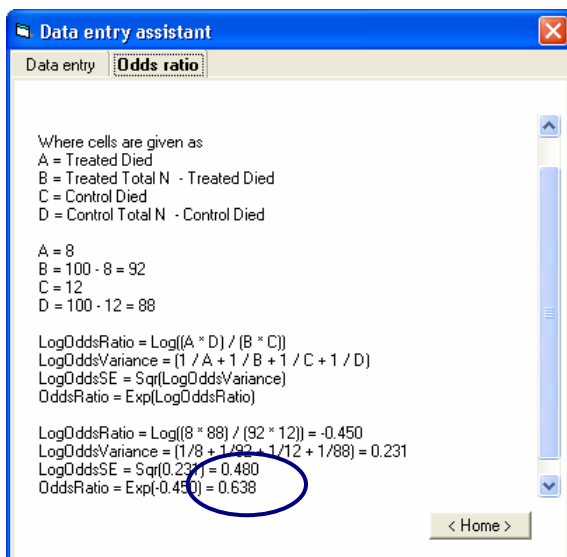
## Show details for the computations

→ Double click on the value 0.638



	Study name	Treated Died	Treated Total N	Control Died	Control Total N	Odds ratio	Log odds ratio	Std Err	Variance	J
1	Madison	8	100	12	100	0.638	-0.450	0.480	0.231	
2	Moyer	20	250	24	250	0.813	-0.200	0.317	0.100	
3	Goldman	8	100	14	100	0.534	-0.627	0.468	0.219	
4	Graham	12	100	16	100	0.716	-0.334	0.411	0.169	
5	Manning	80	1000	200	1000	0.348	-1.056	0.141	0.020	
6										
7										

The program shows how this value was computed.



Data entry assistant

Data entry Odds ratio

Where cells are given as  
A = Treated Died  
B = Treated Total N - Treated Died  
C = Control Died  
D = Control Total N - Control Died

A = 8  
B = 100 - 8 = 92  
C = 12  
D = 100 - 12 = 88

$\text{LogOddsRatio} = \text{Log}[(A * D) / (B * C)]$   
 $\text{LogOddsVariance} = (1 / A + 1 / B + 1 / C + 1 / D)$   
 $\text{LogOddsSE} = \text{Sqr}(\text{LogOddsVariance})$   
 $\text{OddsRatio} = \text{Exp}(\text{LogOddsRatio})$

$\text{LogOddsRatio} = \text{Log}[(8 * 88) / (92 * 12)] = -0.450$   
 $\text{LogOddsVariance} = (1 / 8 + 1 / 92 + 1 / 12 + 1 / 88) = 0.231$   
 $\text{LogOddsSE} = \text{Sqr}(0.231) = 0.480$   
 $\text{OddsRatio} = \text{Exp}(-0.450) = 0.638$

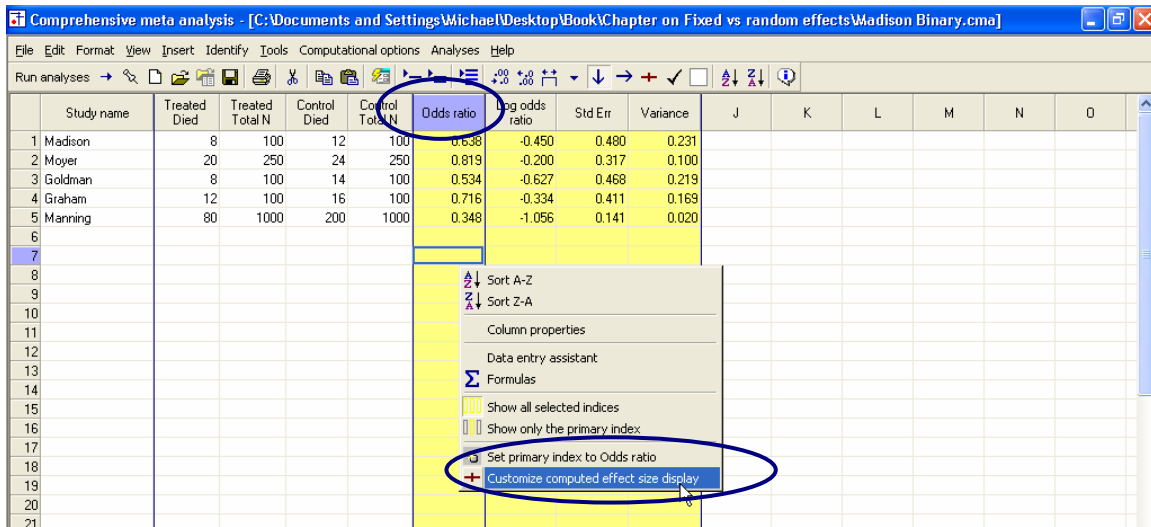
< Home >

## Set the default index

At this point, the program has displayed the odds ratio and the log odds ratio, which we'll be using in this example.

You have the option of adding additional indices, and/or specifying which index should be used as the "Default" index when you run the analysis.

- Right-click on any of the yellow columns
- Select CUSTOMIZE COMPUTED EFFECT SIZE DISPLAY

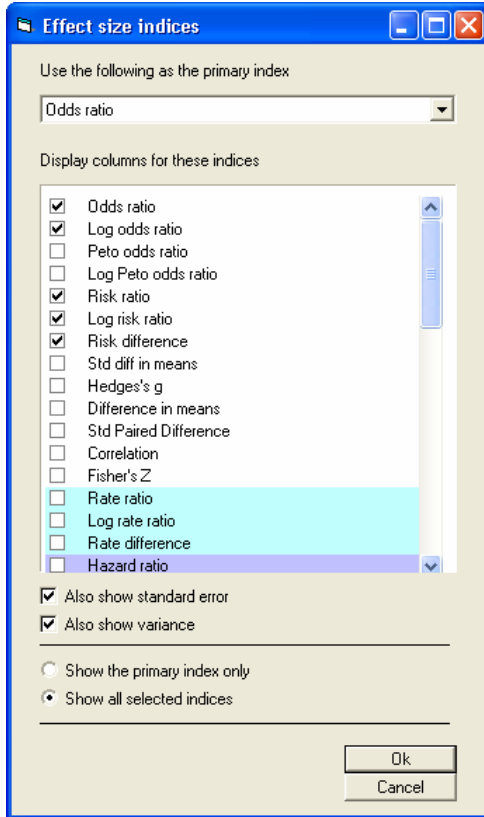


The screenshot shows the 'Comprehensive meta analysis' software interface. The main window displays a table with the following data:

Study name	Treated Died	Treated Total N	Control Died	Control Total N	Odds ratio	Log odds ratio	Std Err	Variance	J	K	L	M	N	O
1 Madison	8	100	12	100	0.638	-0.450	0.480	0.231						
2 Moyer	20	250	24	250	0.819	-0.200	0.317	0.100						
3 Goldman	8	100	14	100	0.534	-0.627	0.468	0.219						
4 Graham	12	100	16	100	0.716	-0.334	0.411	0.169						
5 Manning	80	1000	200	1000	0.348	-1.056	0.141	0.020						

A context menu is open over the 'Odds ratio' column. The menu items are: Sort A-Z, Sort Z-A, Column properties, Data entry assistant, Formulas, Show all selected indices, Show only the primary index, Set primary index to Odds ratio, and Customize computed effect size display. The 'Customize computed effect size display' option is highlighted with a blue oval.

The program displays this dialog box.



- Check Risk ratio, Log risk ratio, Risk difference
- Click OK

	Study name	Treated Died	Treated Total N	Control Died	Control Total N	Odds ratio	Log odds ratio	Std Err	Variance	Risk ratio	Log risk ratio	Std Err	Variance	Risk difference	Std Err
1	Madison	8	100	12	100	0.638	-0.450	0.480	0.231	0.667	-0.405	0.434	0.188	-0.040	0.042
2	Moyer	20	250	24	250	0.819	-0.200	0.317	0.100	0.833	-0.182	0.289	0.084	-0.016	0.025
3	Goldman	8	100	14	100	0.534	-0.627	0.468	0.219	0.571	-0.560	0.420	0.176	-0.060	0.044
4	Graham	12	100	16	100	0.716	-0.334	0.411	0.169	0.750	-0.288	0.355	0.126	-0.040	0.049
5	Manning	80	1000	200	1000	0.348	-1.056	0.141	0.020	0.400	-0.916	0.124	0.016	-0.120	0.015
6															
7															
8															
9															
10															
11															

- The program has added columns for these indices

## Run the analysis

→ Click RUN ANALYSES

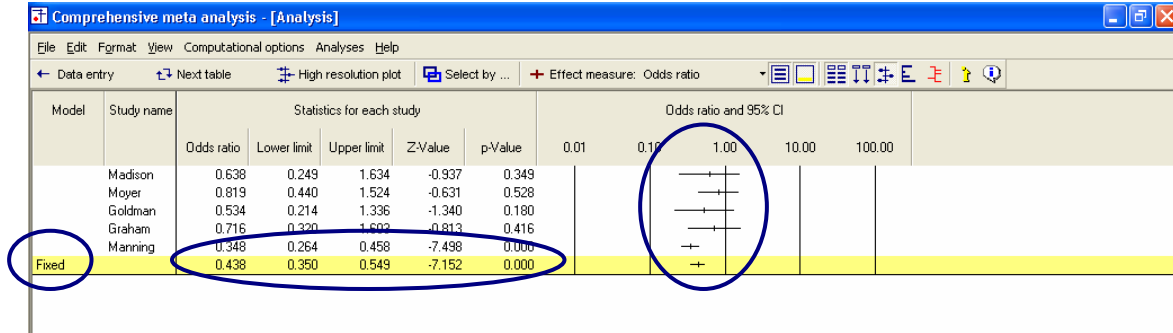
Study name	Treated Died	Treated Total N	Control Died	Control Total N	Odds ratio	Log odds ratio	Std Err	Variance	Risk ratio	Log risk ratio	Std Err	Variance	Risk difference	Std Err
1 Madison	8	100	12	100	0.638	-0.450	0.480	0.231	0.667	-0.405	0.434	0.188	-0.040	0.042
2 Moyer	20	250	24	250	0.819	-0.200	0.317	0.100	0.833	-0.182	0.289	0.084	-0.016	0.025
3 Goldman	8	100	14	100	0.534	-0.627	0.468	0.219	0.571	-0.560	0.420	0.176	-0.060	0.044
4 Graham	12	100	16	100	0.716	-0.334	0.411	0.169	0.750	-0.288	0.355	0.126	-0.040	0.049
5 Manning	80	1000	200	1000	0.348	-1.056	0.141	0.020	0.400	-0.916	0.124	0.016	-0.120	0.015

The program displays this screen.

- The default effect size is the odds ratio
- The default model is fixed effect

Model	Study name	Statistics for each study					Odds ratio and 95% CI				
		Odds ratio	Lower limit	Upper limit	Z-Value	p-Value	0.01	0.10	1.00	10.00	100.00
	Madison	0.638	0.249	1.634	-0.937	0.349					
	Moyer	0.819	0.440	1.524	-0.631	0.528					
	Goldman	0.534	0.214	1.336	-1.340	0.180					
	Graham	0.716	0.320	1.603	-0.813	0.416					
	Manning	0.348	0.264	0.458	-7.498	0.000					
Fixed		0.438	0.350	0.549	-7.152	0.000					

The screen should look like this.



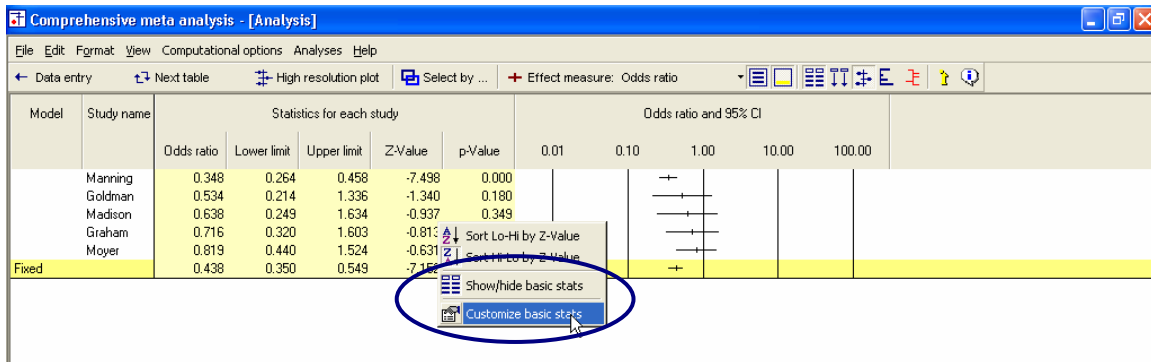
We can immediately get a sense of the studies and the combined effect. For example,

- All the effects fall below 1.0, in the range of 0.350 to 0.820. The treated group did better than the control group in all studies
- Some studies are clearly more precise than others. The confidence interval for Madison is substantially wider than the one for Manning, with the other three studies falling somewhere in between
- The combined effect is 0.438 with a 95% confidence interval of 0.350 to 0.549

## Customize the screen

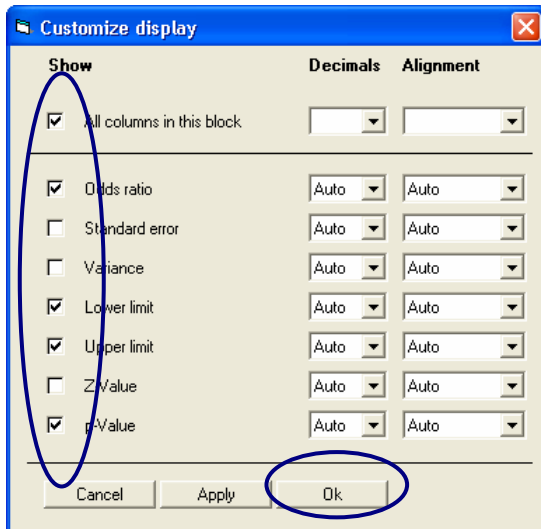
We want to hide the column for the z-value.

- Right-click on one of the “Statistics” columns
- Select CUSTOMIZE BASIC STATS



Model	Study name	Statistics for each study					Odds ratio and 95% CI				
		Odds ratio	Lower limit	Upper limit	Z-Value	p-Value	0.01	0.10	1.00	10.00	100.00
	Manning	0.348	0.264	0.458	-7.498	0.000			+		
	Goldman	0.534	0.214	1.336	-1.340	0.180			+		
	Madison	0.638	0.249	1.634	-0.937	0.349			+		
	Graham	0.716	0.320	1.603	-0.813				+		
	Moyer	0.819	0.440	1.524	-0.631				+		
Fixed		0.438	0.350	0.549	-7.157				+		

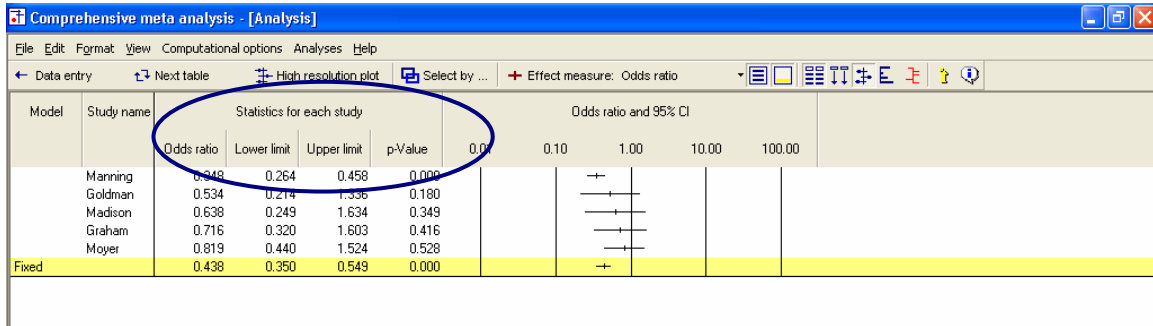
- Assign check-marks as shown here
- Click OK



Show	Decimals	Alignment
<input checked="" type="checkbox"/> All columns in this block		
<input checked="" type="checkbox"/> Odds ratio	Auto	Auto
<input type="checkbox"/> Standard error	Auto	Auto
<input type="checkbox"/> Variance	Auto	Auto
<input checked="" type="checkbox"/> Lower limit	Auto	Auto
<input checked="" type="checkbox"/> Upper limit	Auto	Auto
<input type="checkbox"/> Z-Value	Auto	Auto
<input checked="" type="checkbox"/> p-Value	Auto	Auto

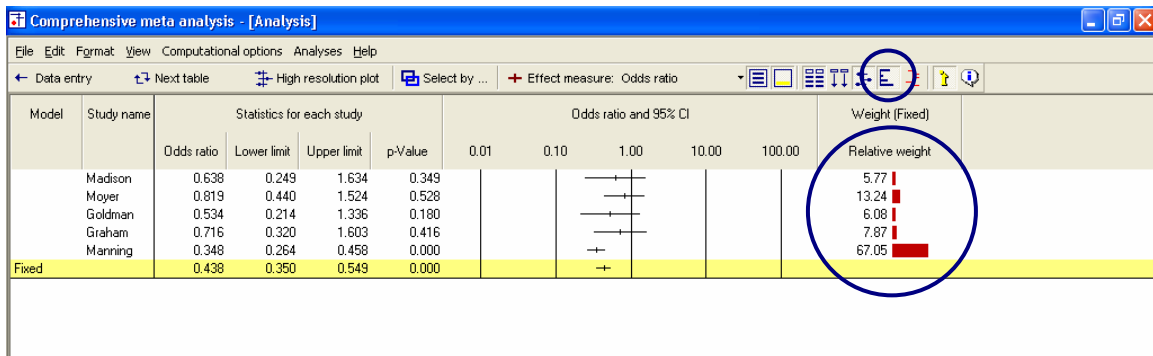
Note – the standard error and variance are never displayed for the odds ratio. They are displayed when the corresponding boxes are checked and Log odds ratio is selected as the index.

The program has hidden some of the columns, leaving us more room to work with on the display.



## Display weights

→ Click the tool for SHOW WEIGHTS



The program now shows the relative weight assigned to each study for the fixed effect analysis. By “relative weight” we mean the weights as a percentage of the total weights, with all relative weights summing to 100%.

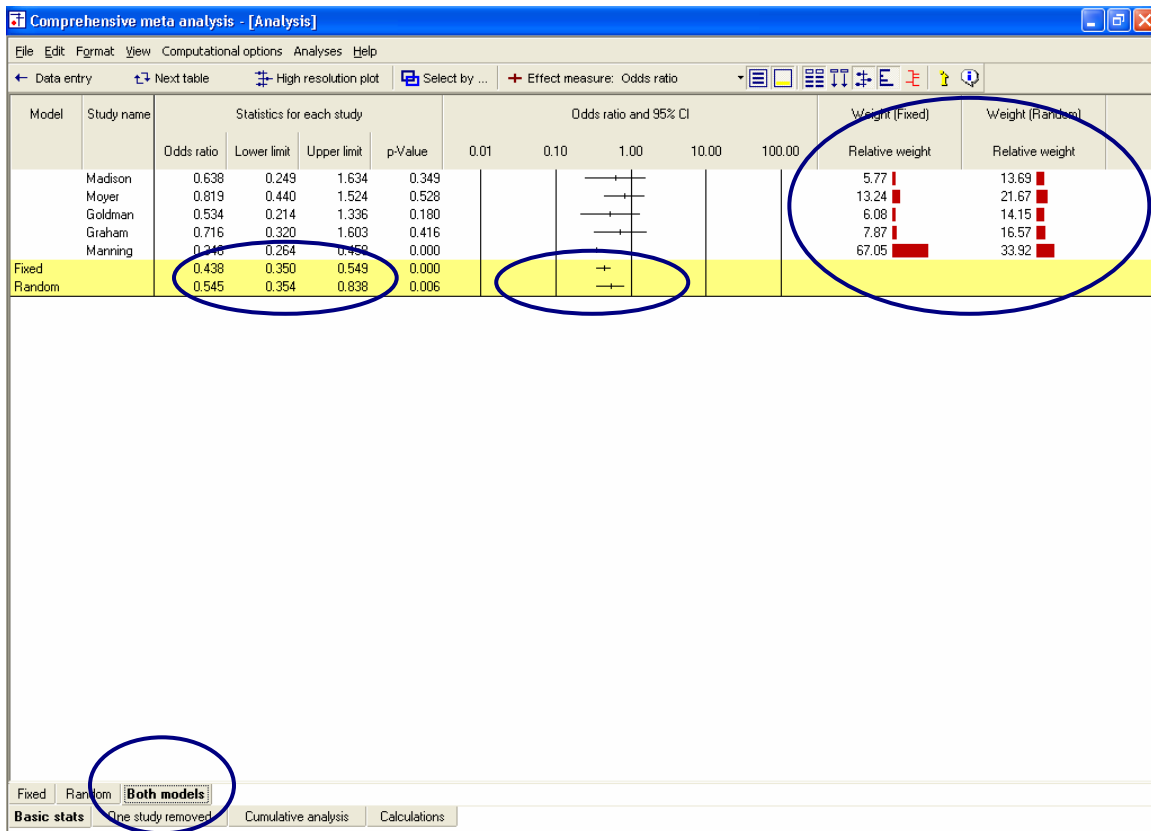
For example, Madison was assigned a relative weight of 5.77% while Manning was assigned a relative weight of 67.05%.



## Compare the fixed effect and random effects models

→ At the bottom of the screen, select BOTH MODELS

- The program shows the combined effect and confidence limits for both fixed and random effects models
- The program shows weights for both the fixed effect and the random effects models



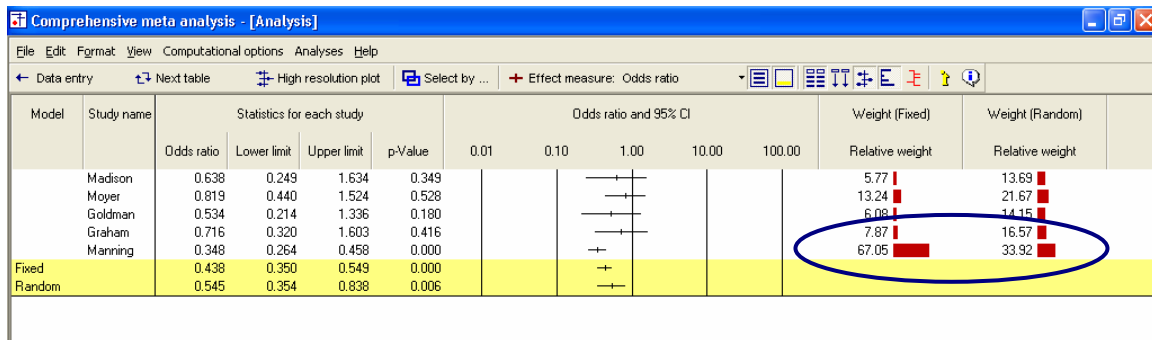
## Impact of model on study weights

The Manning study, with a large sample size (N=1000 per group) is assigned 67% of the weight under the fixed effect model but only 34% of the weight under the random effects model.

This follows from the logic of fixed and random effects models explained earlier.

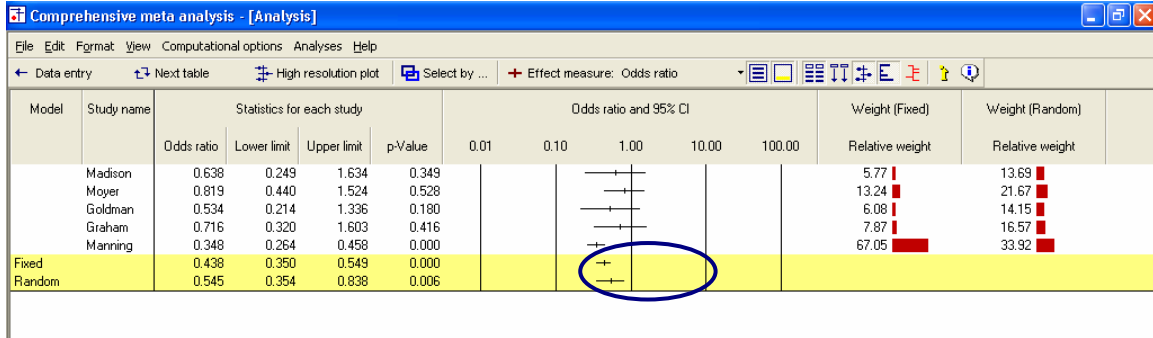
Under the fixed effect model we assume that all studies are estimating the same value and this study yields a better estimate than the others, so we take advantage of that.

Under the random effects model we assume that each study is estimating a unique effect. The Manning study yields a precise estimate of its population, but that population is only one of many, and we don't want it to dominate the analysis. Therefore, we assign it 34% of the weight. This is more than the other studies, but not the dominant weight that we gave it under fixed effects.



## Impact of model on the combined effect

As it happens, the Manning study has a powerful effect size (that is, an odds ratio of 0.34), which represents a very substantial impact, roughly a 66% drop in risk. Under the fixed effect model, where this study dominates the weights, it pulls the effect size to the left to 0.44 (that is, to a more substantial benefit). Under the random effects model, it still pulls the effect size to the left, but only to 0.55.

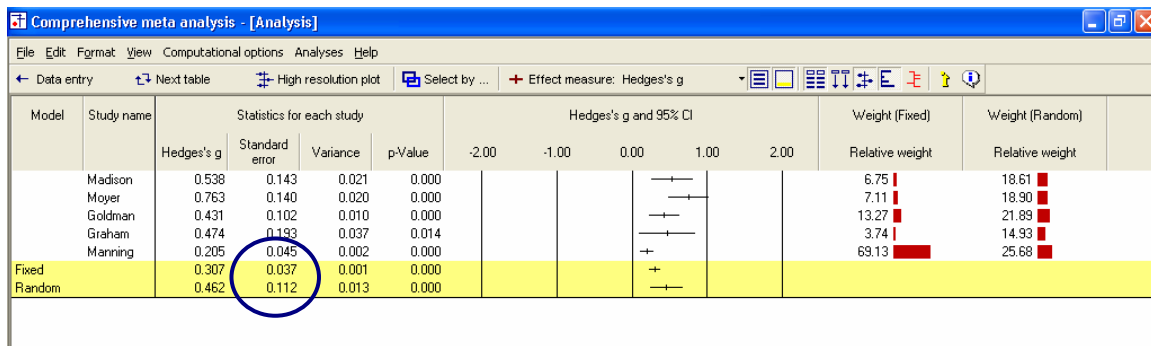


## Impact of model on the confidence interval

Under fixed effect, we “set” the between-studies dispersion to zero. Therefore, for the purpose of estimating the mean effect, the only source of uncertainty is within-study error. With a combined total near 1500 subjects per group the within-study error is small, so we have a precise estimate of the combined effect. The confidence interval is relatively narrow, extending from 0.35 to 0.55.

Under random effects, dispersion between studies is considered a real source of uncertainty. And, there is a lot of it. The fact that these five studies vary so much one from the other tells us that the effect will vary depending on details that vary randomly from study to study. If the persons who performed these studies happened to use older subjects, or a shorter duration, for example, the effect size would have changed.

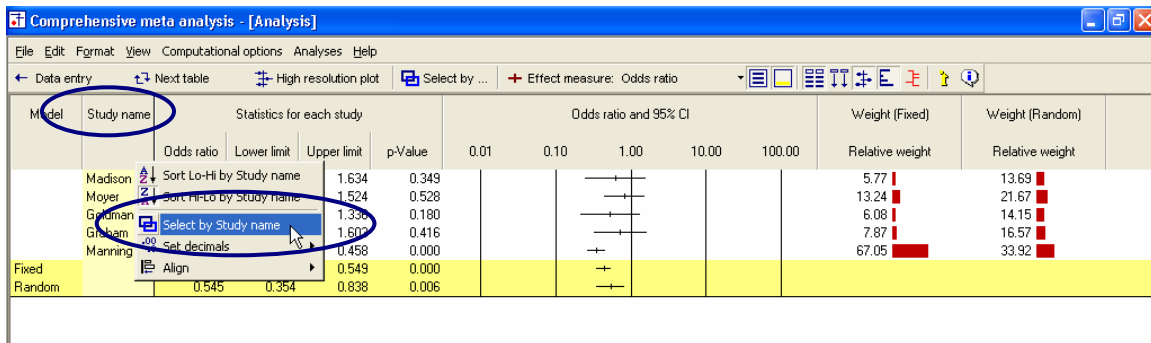
While this dispersion is “real” in the sense that it is caused by real differences among the studies, it nevertheless represents error if our goal is to estimate the mean effect. For computational purposes, the variance due to between-study differences is included in the error term. In our example we have only five studies, and the effect sizes do vary. Therefore, our estimate of the mean effect is not terribly precise, as reflected in the width of the confidence interval, 0.35 to 0.84, substantially wider than that for the fixed effect model.



## What would happen if we eliminated Manning?

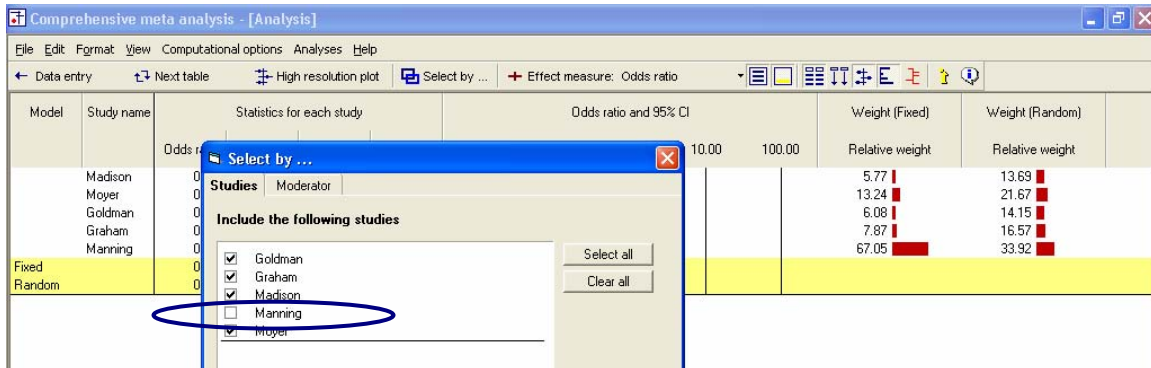
Manning was the largest study, and also the study with the most powerful (left-most) effect size. To better understand the impact of this study under the two models, let's see what would happen if we were to remove this study from the analysis.

- Right-click on STUDY NAME
- Select SELECT BY STUDY NAME

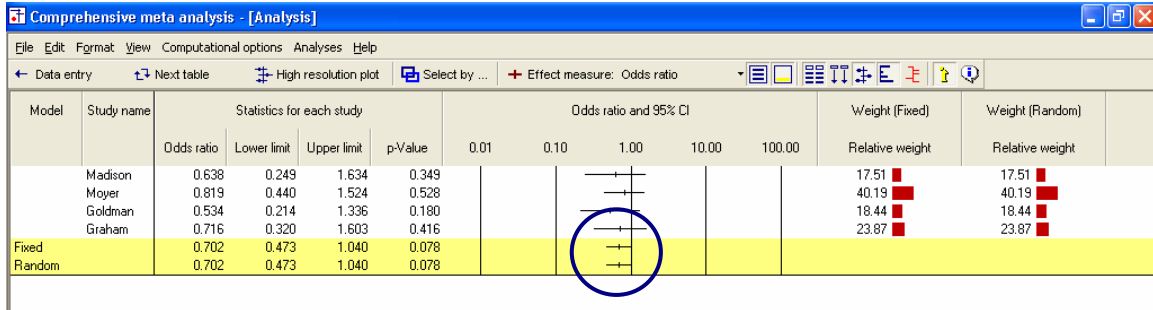


The program opens a dialog box with the names of all studies.

- Remove the check from Manning
- Click OK



The analysis now looks like this.

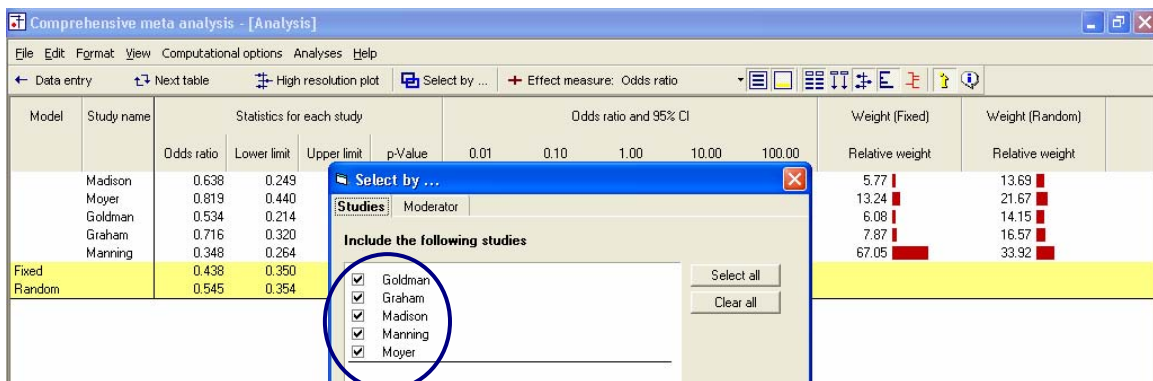


For both the fixed effect and random effects models, the combined effect is now close to 0.70.

- Under fixed effects Manning had pulled the effect down to 0.44
- Under random effects Manning had pulled the effect down to 0.55
- Thus, this study had a substantial impact under either model, but more so under fixed than random effects

→ Right-click on STUDY NAME

→ Add a check for Manning so the analysis again has five studies



## Additional statistics

→ Click NEXT TABLE on the toolbar

Model	Study name	Statistics for each study				Odds ratio and 95% CI					Weight (Fixed)		Weight (Random)	
		Odds ratio	Lower limit	Upper limit	p-Value	0.01	0.10	1.00	10.00	100.00	Relative weight	Relative weight		
	Madison	0.638	0.249	1.634	0.349						5.77	13.69		
	Moyer	0.819	0.440	1.524	0.528						13.24	21.67		
	Goldman	0.534	0.214	1.336	0.180						6.08	14.15		
	Graham	0.716	0.320	1.603	0.416						7.87	16.57		
	Manning	0.348	0.264	0.458	0.000						67.05	33.92		
Fixed		0.438	0.350	0.549	0.000									
Random		0.545	0.354	0.838	0.006									

The program switches to this screen.

The program shows the point estimate and confidence interval. These are the same values that had been shown on the forest plot.

Model	Effect size and 95% interval				Test of null (2-Tail)		Heterogeneity				Tau-squared	
	Number Studies	Point estimate	Lower limit	Upper limit	Z-value	P-value	Q-value	df (Q)	P-value	I-squared	Tau Squared	Standard Error
Fixed	5	0.438	0.350	0.549	-7.152	0.000	8.796	4	0.066	54.523	0.123	0.1
Random	5	0.545	0.354	0.838	-2.761	0.006						

- Under fixed effect the combined effect is 0.438 with 95% confidence interval of 0.350 to 0.549
- Under random effects the combined effect is 0.545 with 95% confidence interval of 0.354 to 0.838

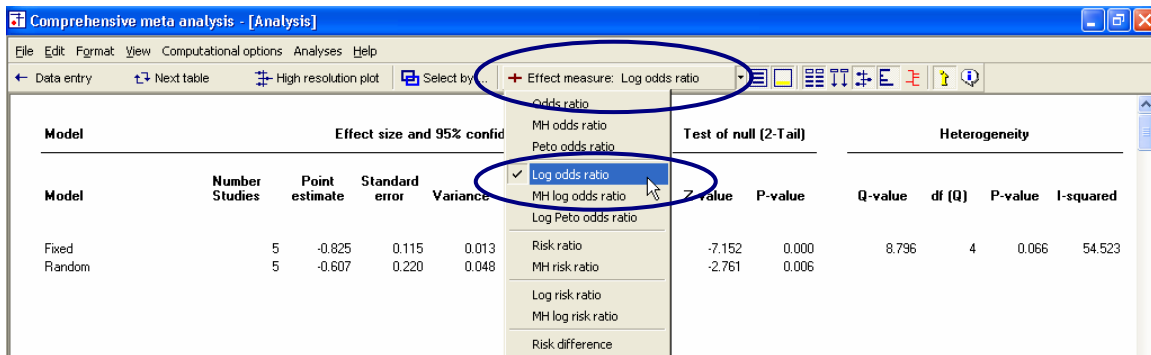
## Test of the null hypothesis

Under the fixed effect model the null hypothesis is that the common effect is zero. Under the random effects model the null hypothesis is that the mean of the true effects is zero.

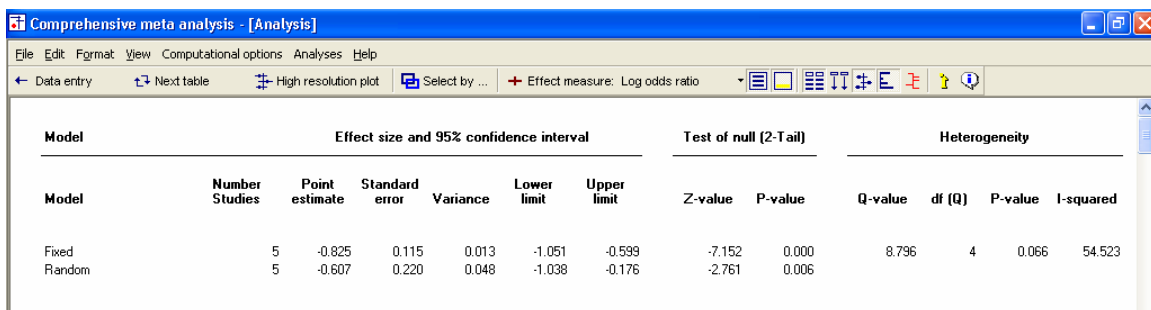
In either case, the null hypothesis is tested by the z-value, which is computed as Log odds ratio/SE for the corresponding model.

To this point we've been displaying the odds ratio. The z-value is correct as displayed (since it is always based on the log), but to understand the computation we need to switch the display to show log values.

Select LOG ODDS RATIO from the drop down box.



The screen should look like this.



Note that all values are now in log units.

- The point estimate for the fixed effect and random effects models are now -0.825 and -0.607, which are the natural logs of 0.438 and 0.545
- The program now displays the standard error and variance, which can be displayed for the log odds ratio but not for the odds ratio



For the fixed effect model

$$Z = \frac{-0.825}{0.115} = -7.152$$

For the random effects model

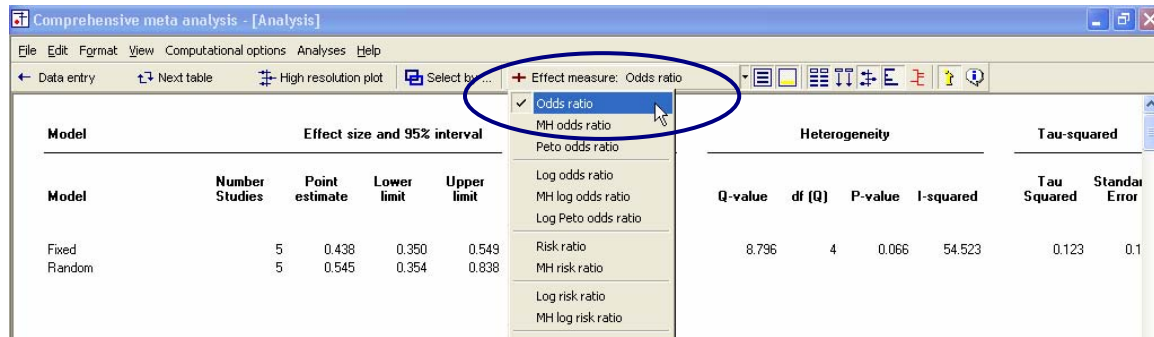
$$Z^* = \frac{-0.607}{0.220} = -2.761$$

With two-tailed p-values < 0.001 and 0.006 respectively.

## Test of the heterogeneity

Switch the display back to Odds ratio.

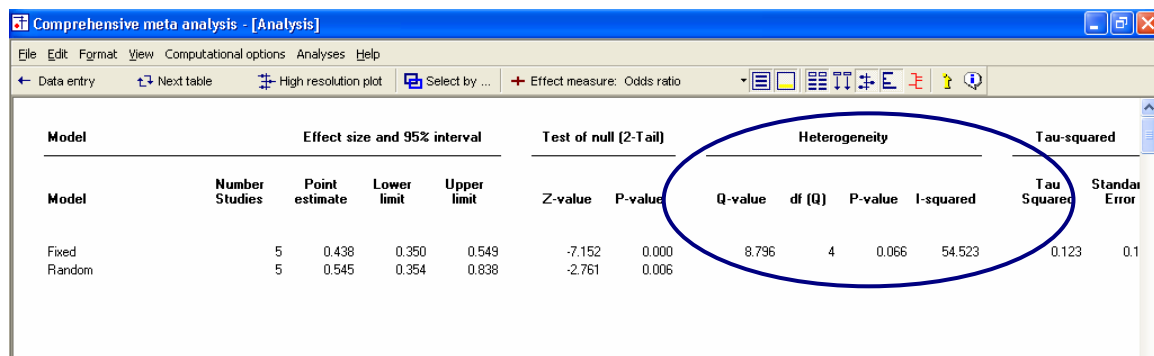
→ Select ODDS RATIO from the drop-down box



Note, however, that the statistics addressed in this section are always computed using log value, regardless of whether Odds ratio or Log odds ratio has been selected as the index.

The null hypothesis for heterogeneity is that the studies share a common effect size.

The statistics in this section address the question of whether the observed dispersion among effects exceeds the amount that would be expected by chance.



The Q statistic reflects the observed dispersion. Under the null hypothesis that all studies share a common effect size, the expected value of Q is equal to the degrees of freedom (the number of studies minus 1), and is distributed as Chi-square with  $df = k - 1$  (where k is the number of studies).

- The Q statistic is 8.796, as compared with an expected value of 4

- The p-value is 0.066

If we elect to set alpha at 0.10, then this p-value meets the criterion for statistical significance. If we elect to set alpha at 0.05, then this p-value just misses the criterion. This, of course, is one of the hazards of significance tests.

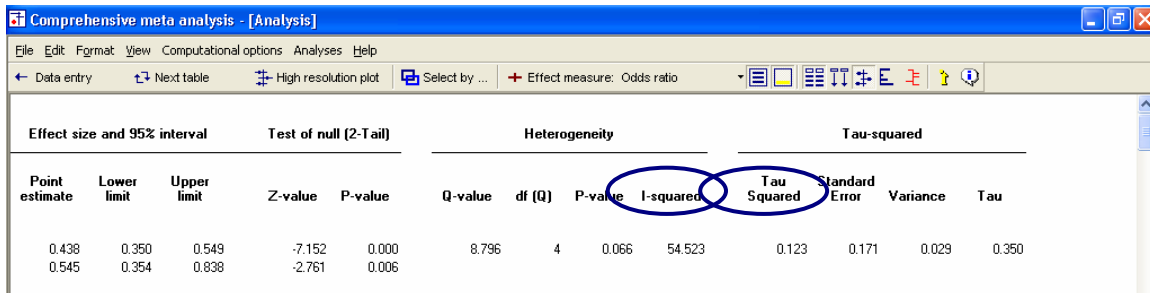
It seems clear that there is substantial dispersion, and probably more than we would expect based on random differences. There probably is real variance among the effects.

As discussed in the text, the decision to use a random effects model should be based on our understanding of how the studies were acquired, and should not depend on a statistically significant p-value for heterogeneity. In any event, this p-value does suggest that a fixed effect model does not fit the data.

## Quantifying the heterogeneity

While Q is meant to test the null hypothesis that there is no dispersion across effect sizes, we want also to quantify this dispersion. For this purpose we would turn to I-squared and tau-squared.

→ To see these statistics, scroll the screen toward the right



The screenshot shows the 'Comprehensive meta-analysis - [Analysis]' window. The 'Effect measure: Odds ratio' is selected. The table displays the following data:

Effect size and 95% interval			Test of null (2-Tail)		Heterogeneity			Tau-squared				
Point estimate	Lower limit	Upper limit	Z-value	P-value	Q-value	df (Q)	P-value	I-squared	Tau Squared	Standard Error	Variance	Tau
0.438	0.350	0.549	-7.152	0.000	8.796	4	0.066	54.523	0.123	0.171	0.029	0.350
0.545	0.354	0.838	-2.761	0.006								

- I-squared is 54.5, which means that 55% of the observed variance between studies is due to real differences in the effect size. Only about 45% of the observed variance would have been expected based on random error.
- Tau-squared is 0.123. This is the “Between studies” variance that was used in computing weights.

The Q statistic and tau-squared are reported on the fixed effect line, and not on the random effects line.

These values are displayed on the fixed effect line because they are computed using fixed effect weights. These values are used in both fixed and random effects analyses, but for different purposes.

For the fixed effect analysis Q addresses the question of whether or not the fixed effect model fits the data (is it reasonable to assume that tau-squared is actually zero). However, tau-squared is actually set to zero for the purpose of assigning weights.

For the random effects analysis, these values are actually used to assign the weights.

Return to the main analysis screen

Click NEXT TABLE again to get back to the other screen

Model	Effect size and 95% interval				Test of null (2-Tail)		Heterogeneity				Tau-squared	
	Number Studies	Point estimate	Lower limit	Upper limit	Z-value	P-value	Q-value	df (Q)	P-value	I-squared	Tau Squared	Standard Error
Fixed	5	0.438	0.350	0.549	-7.152	0.000	8.796	4	0.066	54.523	0.123	0.1
Random	5	0.545	0.354	0.838	-2.761	0.006						

Your screen should look like this.

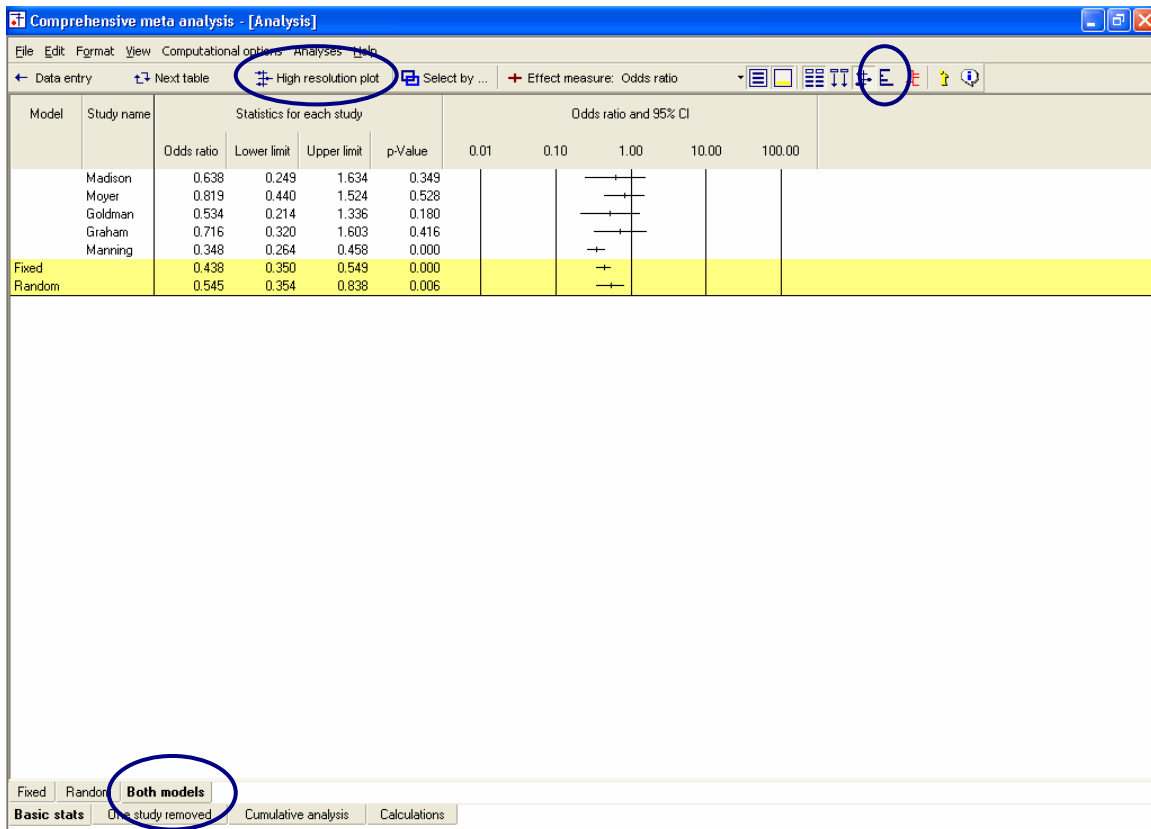
Model	Study name	Statistics for each study				Odds ratio and 95% CI					Weight (Fixed)		Weight (Random)	
		Odds ratio	Lower limit	Upper limit	p-Value	0.01	0.10	1.00	10.00	100.00	Relative weight	Relative weight		
	Madison	0.638	0.249	1.634	0.349			+			5.77	13.69		
	Moyer	0.819	0.440	1.524	0.528			+			13.24	21.67		
	Goldman	0.534	0.214	1.336	0.180			+			6.08	14.15		
	Graham	0.716	0.320	1.603	0.416			+			7.87	16.57		
	Manning	0.348	0.264	0.458	0.000			+			67.05	33.92		
Fixed		0.438	0.350	0.549	0.000			+						
Random		0.545	0.354	0.838	0.006			+						

## High-resolution plots

To this point we've used bar graphs to show the weight assigned to each study.

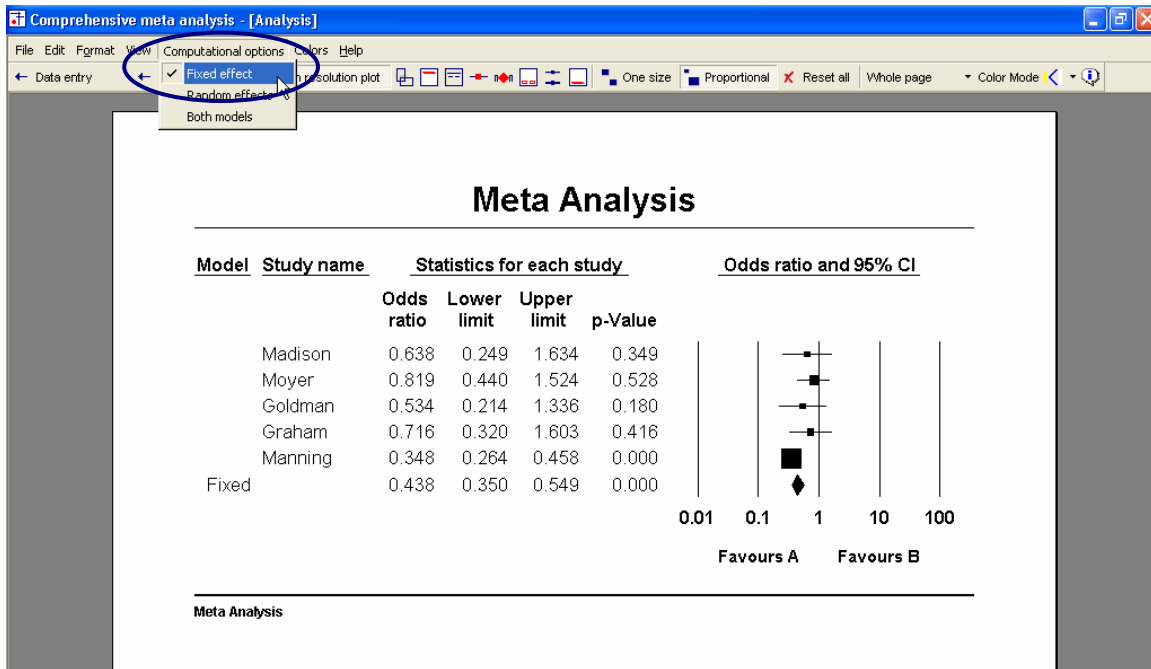
Now, we'll switch to a high-resolution plot, where the weight assigned to each study will determine the size of the symbol representing that study.

- Select BOTH MODELS at the bottom of the screen
- Unclick the SHOW WEIGHTS button on the toolbar
- Click HIGH RESOLUTION PLOT on the toolbar

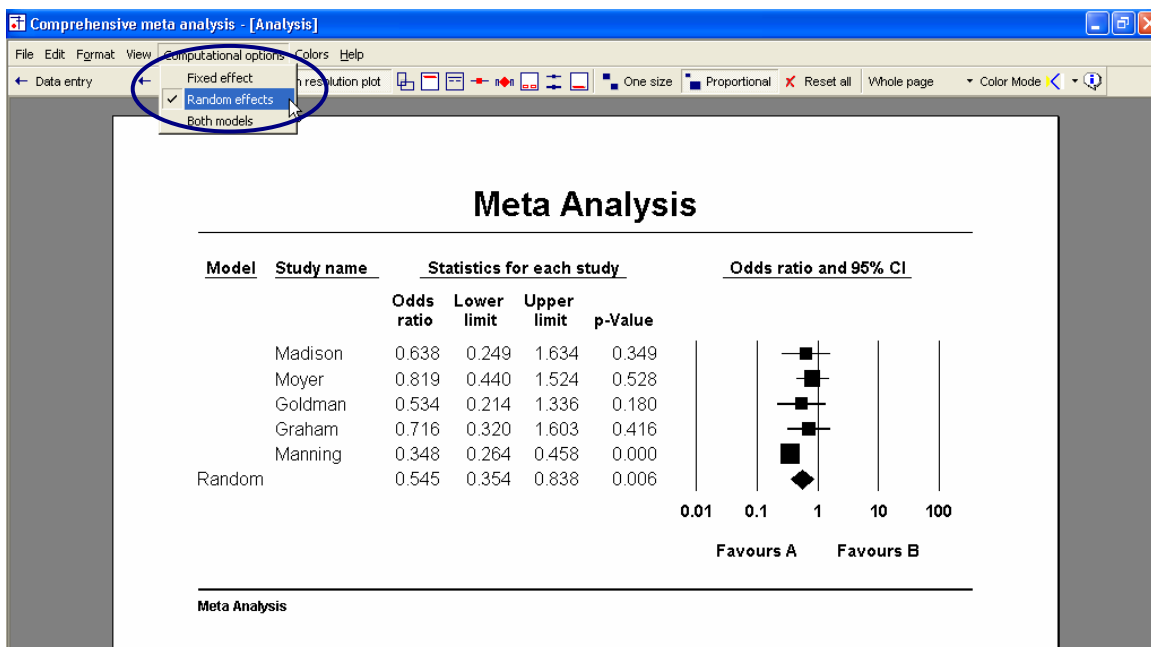


The program shows this screen.

→ Select COMPUTATIONAL OPTIONS > FIXED EFFECT



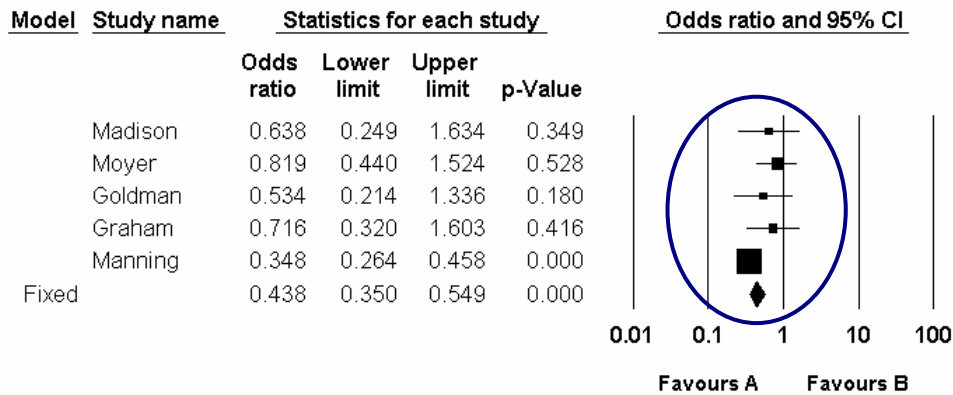
→ Select COMPUTATIONAL OPTIONS > RANDOM EFFECTS



Compare the weights.

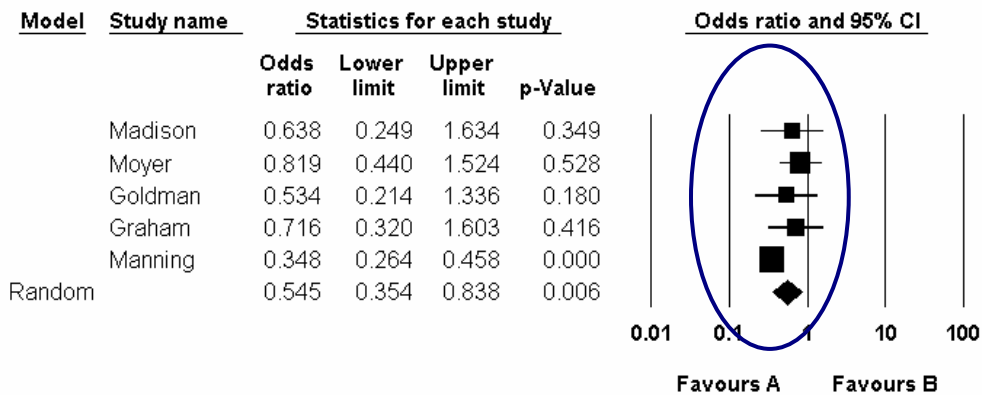
In the first plot, using fixed effect weights, the area of the Manning box is about 10 times that of Madison. In the second, using random effects weights, the area of the Manning box is only about 30% larger than Madison.

## Meta Analysis



Meta Analysis

## Meta Analysis

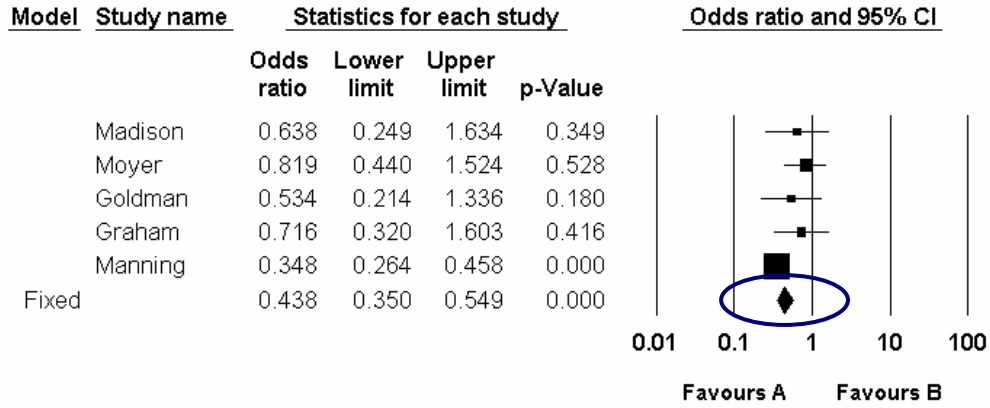


Meta Analysis



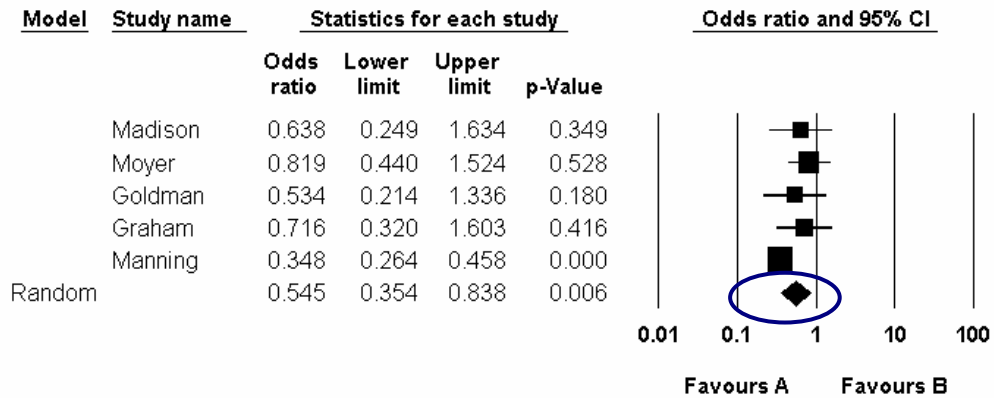
Compare the combined effects and standard error.

## Meta Analysis



Meta Analysis

## Meta Analysis

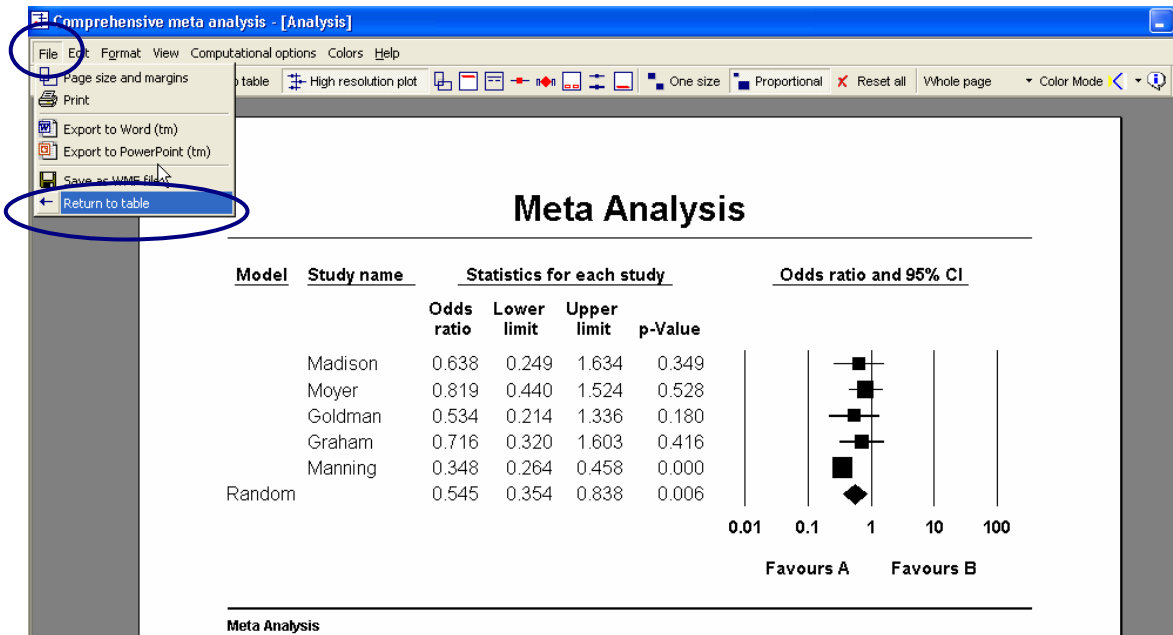


Meta Analysis

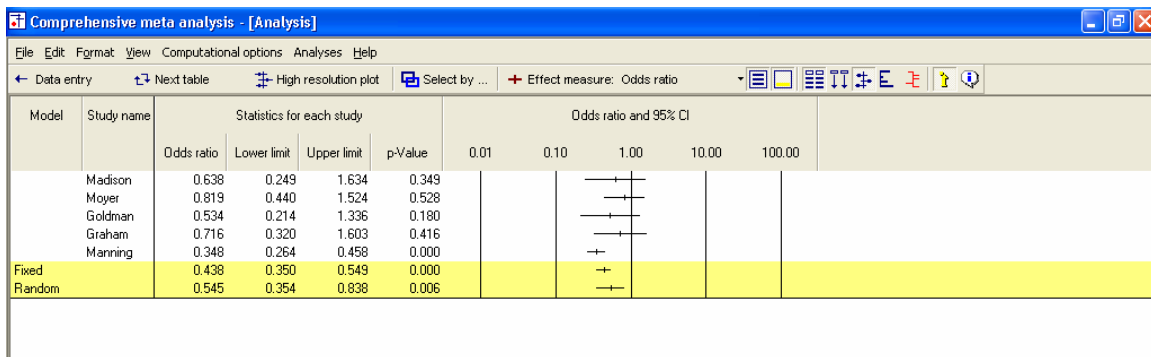
In the first case (fixed), Manning is given substantial weight and pulls the combined effect left to 0.438. In the second case (random), Manning is given less weight, and the combined effect is 0.545.

In the first case (fixed) the only source of error is the error within studies and the confidence interval about the combined effect is relatively narrow. In the second case (random) the fact that the true effect varies from study to study introduces another level of uncertainty to our estimate. The confidence interval about the combined effect is substantially wider than it is for the fixed effect analysis.

→ Click FILE > RETURN TO TABLE



The program returns to the main analysis screen.

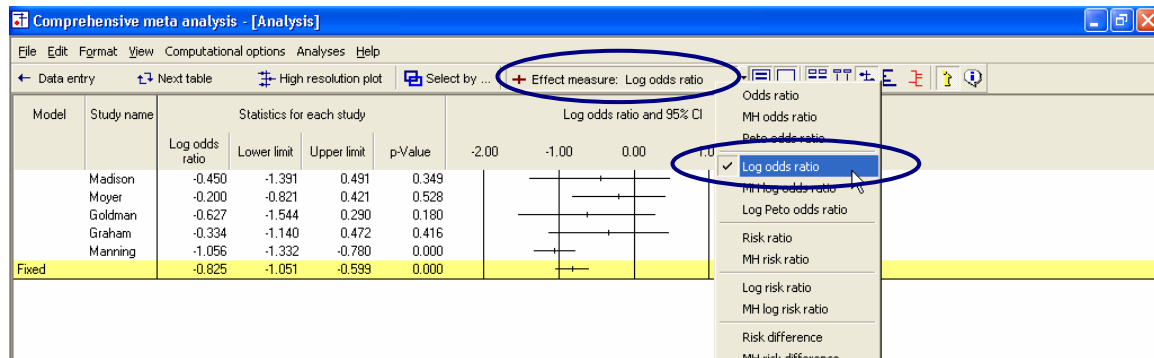


## Computational details

The program allows you to view details of the computations.

Since all calculations are performed using log values, they are easier to follow if we switch the screen to use the log odds ratio as the effect size index.

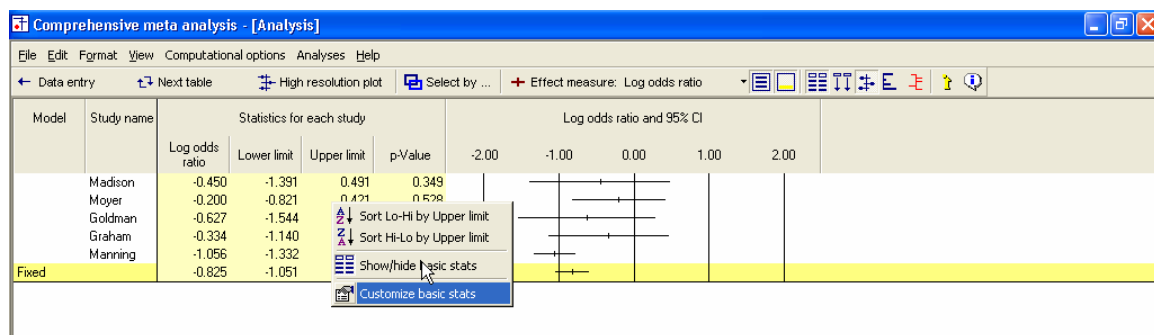
→ Select LOG ODDS RATIO from the drop-down box

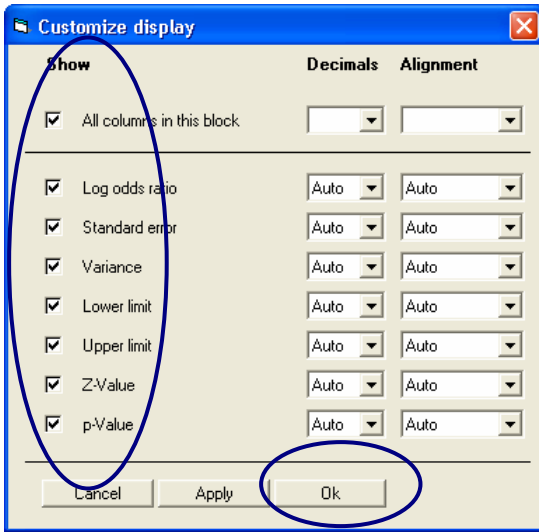


- The program is now showing the effect for each study and the combined effect using log values

We want to add columns to display the standard error and variance.

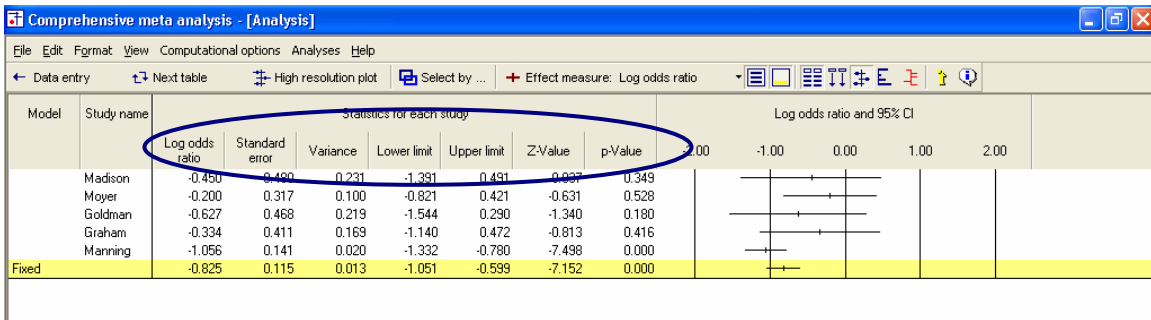
- Right-click on any of the columns in the STATISTICS FOR EACH STUDY section
- Select CUSTOMIZE BASIC STATS





- Check the box next to each statistic
- Click OK

The screen should look like this.



- Note that we now have columns for the standard error and variance, and all values are in log units

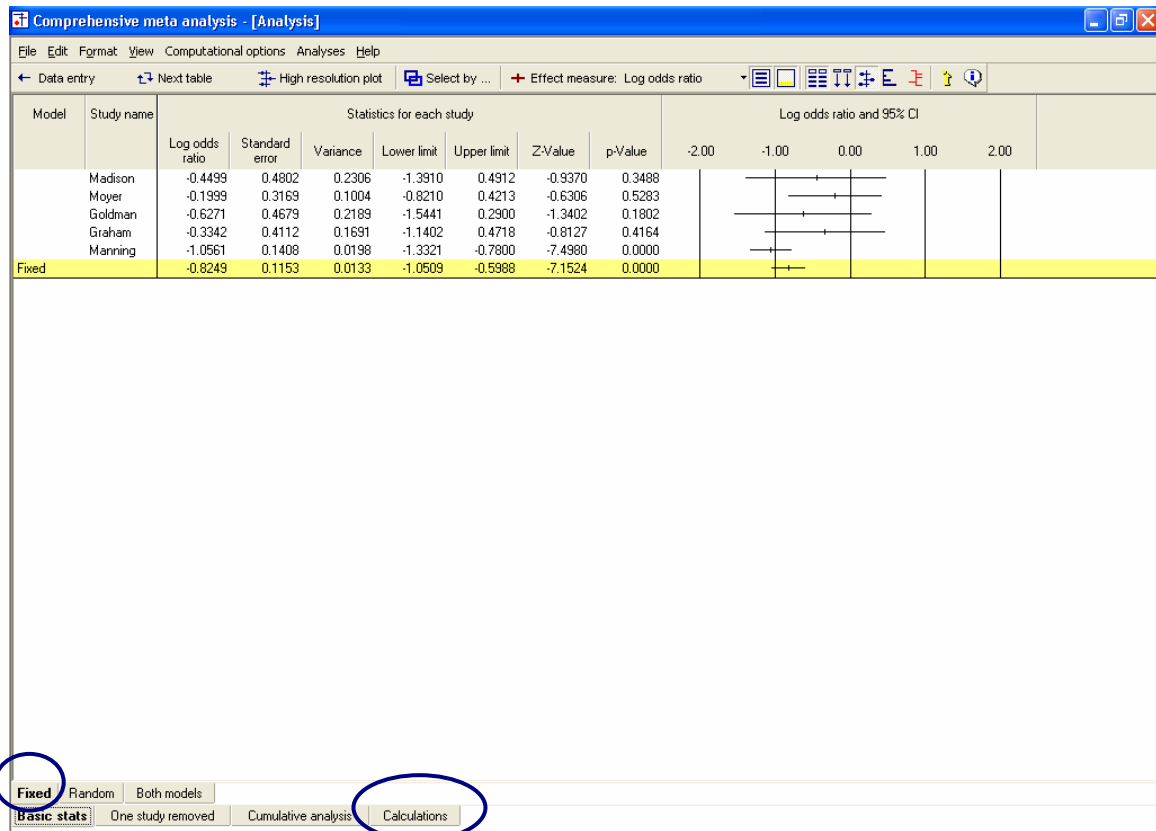
## Computational details for the fixed effect analysis

→ Select **FORMAT > INCREASE DECIMALS** on the menu

This has no effect on the computations, which are always performed using all significant digits, but it makes the example easier to follow.

→ On the bottom of the screen select **FIXED**

→ On the bottom of the screen select **CALCULATIONS**



The program switches to this display.

Model	Study name	Calculations (Fixed)							
		Point	Study Variance	Tau <sup>2</sup> Within	Tau <sup>2</sup> Between	Total Variance	IV-Weight	W	T*W
	Madison	-0.4499	0.2306	0.0000	0.0000	0.2306	4.3371	4.3371	-1.9514
	Moyer	-0.1999	0.1004	0.0000	0.0000	0.1004	9.9563	9.9563	-1.9899
	Goldman	-0.6271	0.2189	0.0000	0.0000	0.2189	4.5678	4.5678	-2.8642
	Graham	-0.3342	0.1691	0.0000	0.0000	0.1691	5.9136	5.9136	-1.9763
	Manning	-1.0561	0.0198	0.0000	0.0000	0.0198	50.4110	50.4110	-53.2366
		-2.6671	0.7389	0.0000	0.0000	0.7389	75.1857	75.1857	-62.0185

For the first study, Madison

$$\text{LogOddsRatio} = \ln\left(\frac{8 * 88}{92 * 12}\right) = -0.4499$$

$$\text{LogOddsVariance} = \left(\frac{1}{8} + \frac{1}{92} + \frac{1}{12} + \frac{1}{88}\right) = 0.2306$$

The weight is computed as

$$w_1 = \frac{1}{0.2306 + 0.0000} = 4.3371$$

Where the second term in the denominator represents tau-squared, which has been set to zero for the fixed effect analysis.

$$T_1w_1 = (-.4499)(4.3371) = -1.9514$$

and so on for the other studies. Then, working with the sums (in the last row)

$$\bar{T}_\bullet = \frac{-62.0185}{75.1857} = -0.8249$$

$$v_\bullet = \frac{1}{75.1857} = 0.0133$$

$$SE(\bar{T}_\bullet) = \sqrt{0.0133} = 0.1153$$

$$\text{Lower Limit} = (-0.8249) - 1.96 * 0.1153 = -1.0509$$

$$\text{Upper Limit} = (-0.8249) + 1.96 * 0.1153 = -0.5988$$

$$Z = \frac{-0.8249}{0.1153} = -7.1524$$

$$p_{2T} = 2[1 - (\Phi(|7.1524|))] < 0.0001$$

→ To switch back to the main analysis screen, click BASIC STATS at the bottom

On this screen, the values presented are the same as those computed above.

Model	Study name	Statistics for each study							Log odds ratio and 95% CI				
		Log odds ratio	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value	-2.00	-1.00	0.00	1.00	2.00
	Madison	-0.4499	0.4802	0.2306	-1.3910	0.4912	-0.9370	0.3488					
	Moyer	-0.1999	0.3169	0.1004	-0.8210	0.4213	-0.6306	0.5283					
	Goldman	-0.6271	0.4679	0.2189	-1.5441	0.2900	-1.3402	0.1802					
	Graham	-0.3342	0.4112	0.1691	-1.1402	0.4718	-0.8127	0.4154					
	Manning	-1.0561	0.1408	0.0198	-1.3321	-0.7800	-7.4980	0.0000					
Fixed		-0.8249	0.1153	0.0133	-1.0509	-0.5988	-7.1524	0.0000					

Finally, if we select Odds ratio as the index, the program takes the effect size and confidence interval, and displays them as odds ratios.

Model	Study name	Statistics for each study					Odds ratio and 95% CI				
		Odds ratio	Lower limit	Upper limit	Z-Value	p-Value	0.01	0.10	1.00	10.00	100.00
	Madison	0.6377	0.2488	1.6343	-0.9370	0.3488			+		
	Moyer	0.8188	0.4400	1.5239	-0.6306	0.5283			+		
	Goldman	0.5342	0.2135	1.3364	-1.3402	0.1802			+		
	Graham	0.7159	0.3198	1.6028	-0.8127	0.4164			+		
	Manning	0.3478	0.2639	0.4584	-7.4980	0.0000			+		
Fixed		0.4383	0.3496	0.5495	-7.1524	0.0000			+		

Concretely,

$$\bar{T}_\bullet = \exp(-0.8249) = 0.4383$$

$$Lower\ Limit = \exp(-1.0509) = 0.3493$$

$$Upper\ Limit = \exp(-0.5988) = 0.5495$$

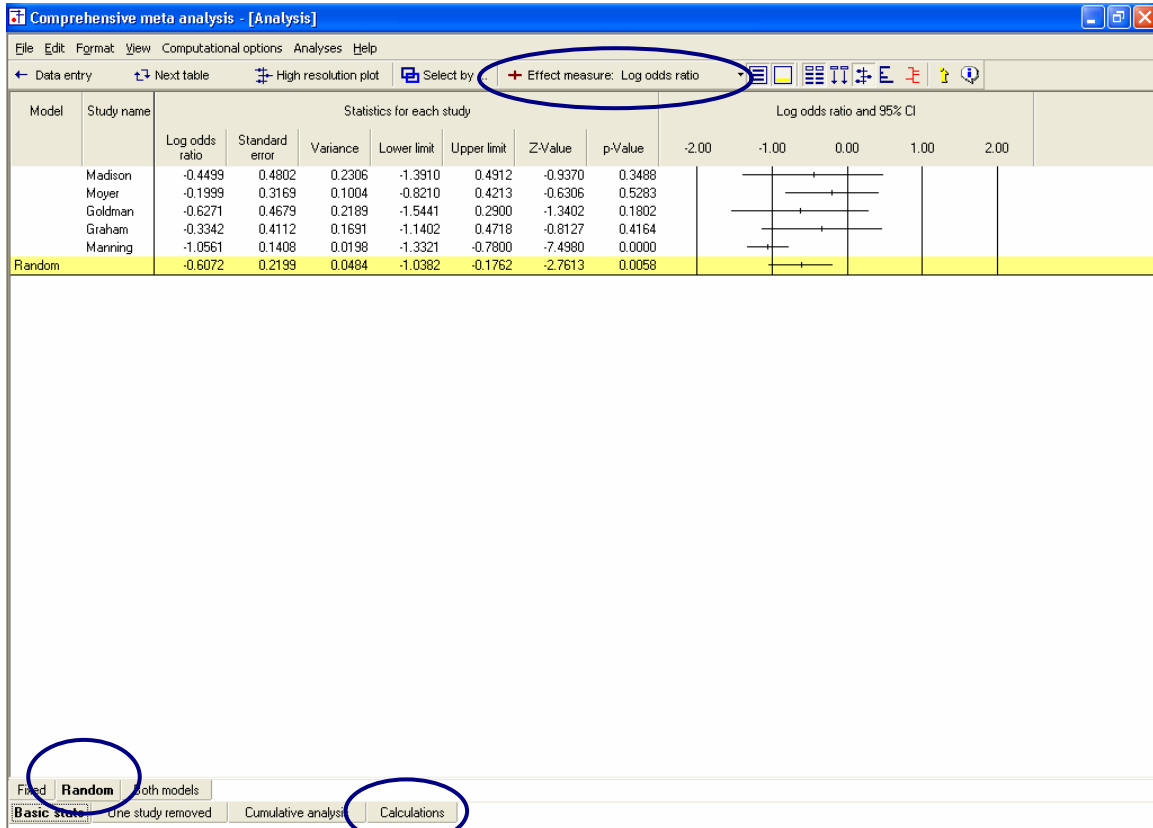
The columns for variance and standard error are hidden. The z-value and p-value that had been computed using log values apply here as well, and are displayed without modification.



## Computational details for the random effects analysis

Now, we can repeat the exercise for random effects

- Select Log odds ratio as the index
- On the bottom of the screen select RANDOM
- On the bottom of the screen select CALCULATIONS



The program switches to this display.

Model	Study name	Calculations (Random)							
		Point	Study Variance	Tau <sup>2</sup> Within	Tau <sup>2</sup> Between	Total Variance	IV-Weight	W	T*W
	Madison	-0.4499	0.2306	0.1227	0.0000	0.3533	2.8305	2.8305	-1.2735
	Moyer	-0.1999	0.1004	0.1227	0.0000	0.2232	4.4809	4.4809	-0.8956
	Goldman	-0.6271	0.2189	0.1227	0.0000	0.3417	2.9269	2.9269	-1.8353
	Graham	-0.3342	0.1691	0.1227	0.0000	0.2918	3.4266	3.4266	-1.1452
	Manning	-1.0561	0.0198	0.1227	0.0000	0.1426	7.0142	7.0142	-7.4074
		-2.6671	0.7389	0.6137	0.0000	1.3525	20.6791	20.6791	-12.5569

For the first study, Madison

$$\text{LogOddsRatio} = \ln\left(\frac{8 * 88}{92 * 12}\right) = -0.4499$$

$$\text{LogOddsVariance} = \left(\frac{1}{8} + \frac{1}{92} + \frac{1}{12} + \frac{1}{88}\right) = 0.2306$$

The weight is computed as

$$w_1^* = \frac{1}{0.2306 + 0.1227} = \frac{1}{0.3533} = 2.8305$$

Where the (\*) indicates that we are using random effects weights, and the second term in the denominator represents tau-squared.

$$T_1^* w_1^* = (-0.4499)(2.8305) = -1.2735$$

and so on for the other studies. Then, working with the sums (in the last row)

$$\bar{T}^* = \frac{-12.5569}{20.6791} = -0.6072$$

$$v^* = \frac{1}{20.6791} = 0.0484$$

$$SE(\bar{T}^*) = \sqrt{0.0484} = 0.2199$$

$$\text{Lower Limit}^* = (-0.6072) - 1.96 * 0.2199 = -1.0382$$

$$\text{Upper Limit}^* = (-0.6072) + 1.96 * 0.2199 = -0.1762$$

$$Z^* = \frac{-0.6072}{0.2199} = -2.7613$$

$$p^*_{2T} = 2 \left[ 1 - (\Phi(|2.7613|)) \right] < 0.0058$$

(Note – The column labeled TAU-SQUARED WITHIN is actually tau-square between studies, and the column labeled TAU-SQUARED BETWEEN is reserved for a fully random effects analysis, where we are performing an analysis of variance).

→ To switch back to the main analysis screen, click BASIC STATS at the bottom

On this screen, the values presented are the same as those computed above.

Model	Study name	Statistics for each study							Log odds ratio and 95% CI				
		Log odds ratio	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value	-2.00	-1.00	0.00	1.00	2.00
	Madison	-0.4499	0.4802	0.2306	-1.3910	0.4912	-0.9370	0.3488					
	Moyer	-0.1999	0.3169	0.1004	-0.8210	0.4213	-0.6306	0.5283					
	Goldman	-0.6271	0.4679	0.2189	-1.5441	0.2900	-1.3402	0.1802					
	Graham	-0.3342	0.4112	0.1691	-1.1402	0.4718	-0.8127	0.4164					
	Manning	-1.0561	0.1408	0.0198	-1.3321	-0.7800	-7.4980	0.0000					
Random		-0.6072	0.2199	0.0484	-1.0382	-0.1762	-2.7613	0.0058					

Finally, if we select Odds ratio as the index, the program takes the effect size and confidence interval, and displays them as odds ratios. The columns for variance and standard error are then hidden.

Model	Study name	Statistics for each study					Odds ratio and 95% CI				
		Odds ratio	Lower limit	Upper limit	Z-Value	p-Value	0.01	0.10	1.00	10.00	100.00
	Madison	0.6377	0.2488	1.6343	-0.9370	0.3488			+		
	Moyer	0.8188	0.4400	1.5239	-0.6306	0.5283			+		
	Goldman	0.5342	0.2135	1.3364	-1.3402	0.1802			+		
	Graham	0.7159	0.3198	1.6028	-0.8127	0.4164			+		
	Manning	0.3478	0.2639	0.4584	-7.4980	0.0000			+		
Random		0.5449	0.3541	0.8384	-2.7613	0.0058			+		

Concretely,

$$\bar{T}_*^* = \exp(-0.6072) = 0.5449$$

$$Lower\ Limit^* = \exp(-1.0382) = .3541$$

$$Lower\ Limit^* = \exp(-0.1762) = 0.8384$$

The columns for variance and standard error are hidden. The z-value and p-value that had been computed using log values apply here as well, and are displayed without modification.

This is the conclusion of the exercise for binary data.



## Example 2 – Means

This appendix shows how to use Comprehensive Meta-Analysis (CMA) to perform a meta-analysis for means using fixed and random effects models.

To download a free trial copy of CMA go to [www.Meta-Analysis.com](http://www.Meta-Analysis.com)

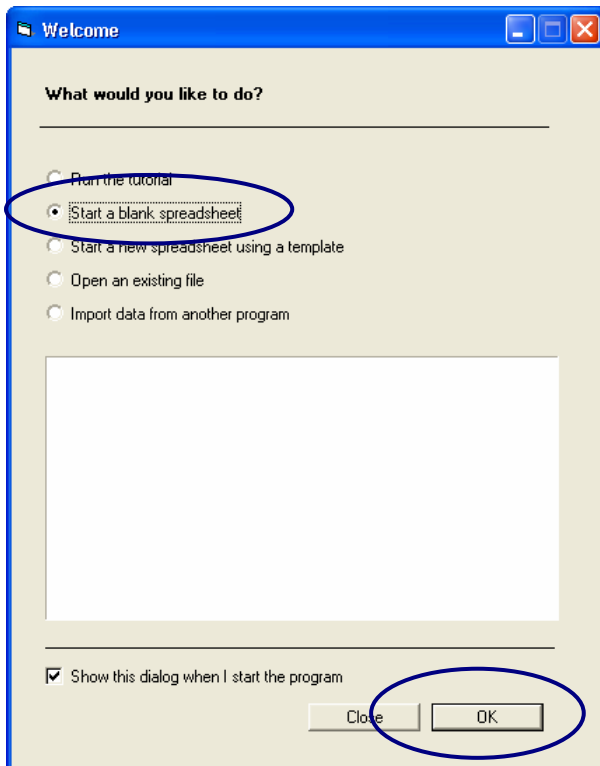
If your trial copy has expired, simply ask for a free extension

- Start the program
- Click “I want to get an unlock code”, to create an e-mail
- In the e-mail, ask for an extension
- We’ll send a code by return e-mail

## Start the program and enter the data

→ Start CMA

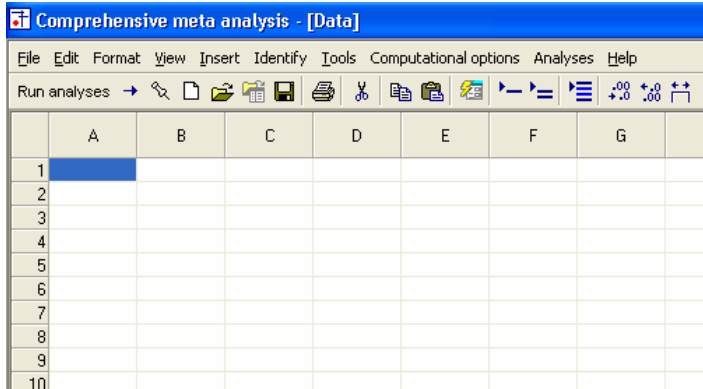
The program shows this dialog box.



→ Select START A BLANK SPREADSHEET

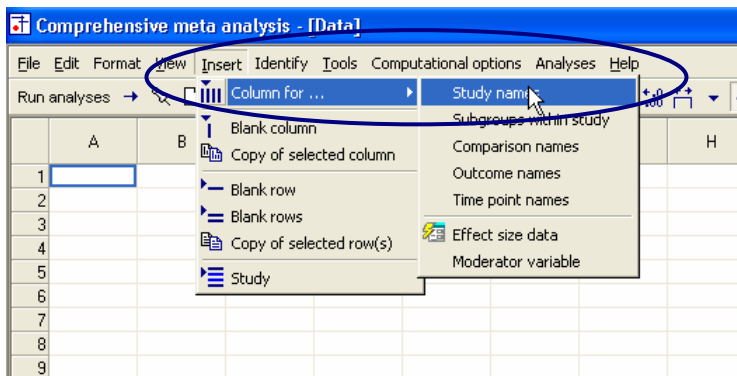
→ Click OK

The program displays this screen.

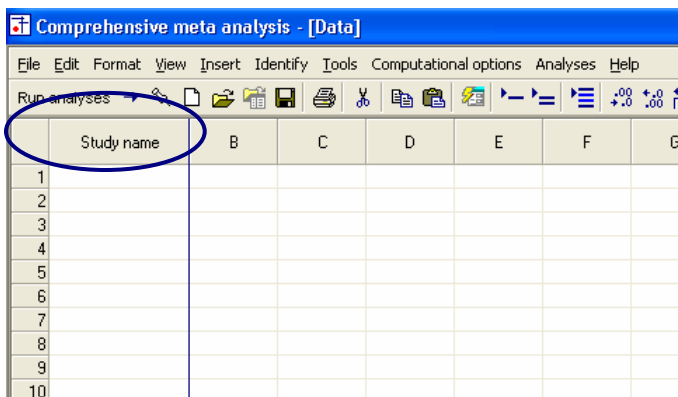


### Insert column for study names

→ Click INSERT > COLUMN FOR > STUDY NAMES



The program has added a column for Study names.

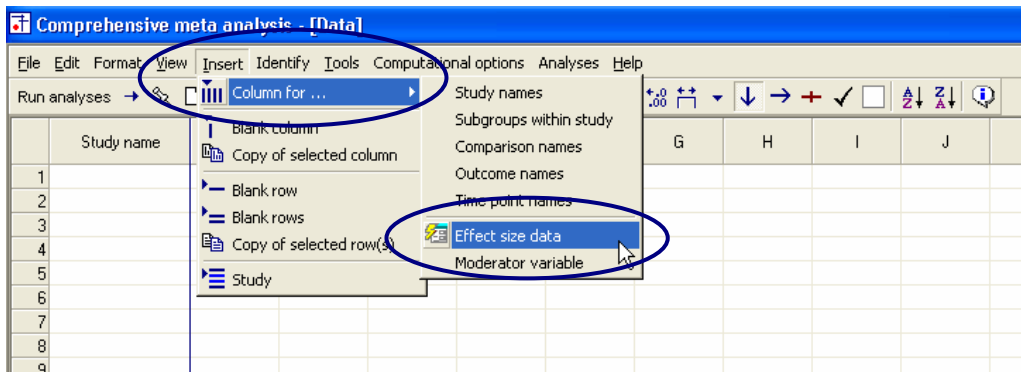


## Insert columns for the effect size data

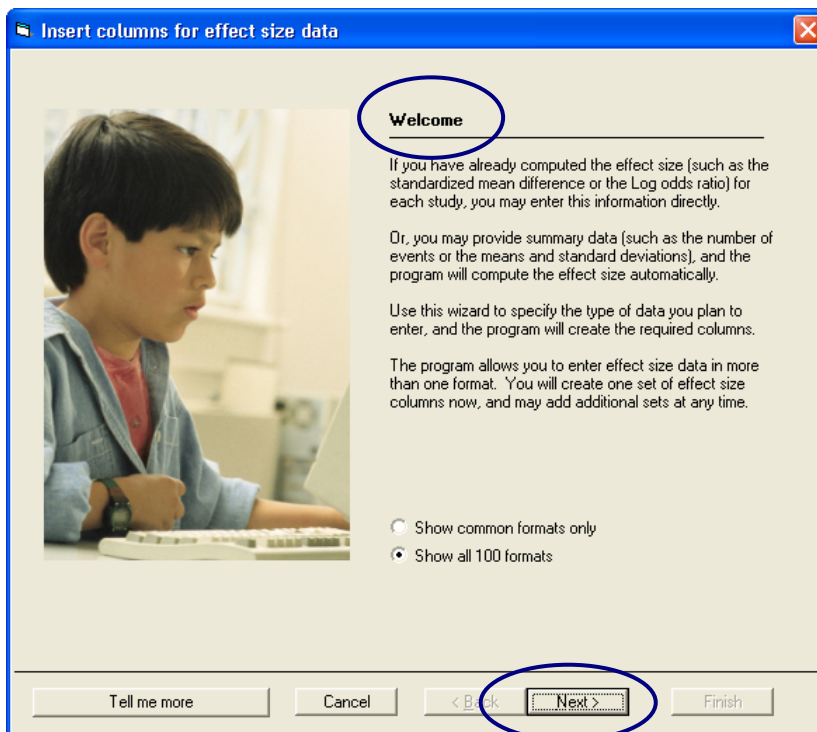
Since CMA will accept data in more than 100 formats, you need to tell the program what format you want to use.

You do have the option to use a different format for each study, but for now we'll start with one format.

→ Click INSERT > COLUMN FOR > EFFECT SIZE DATA



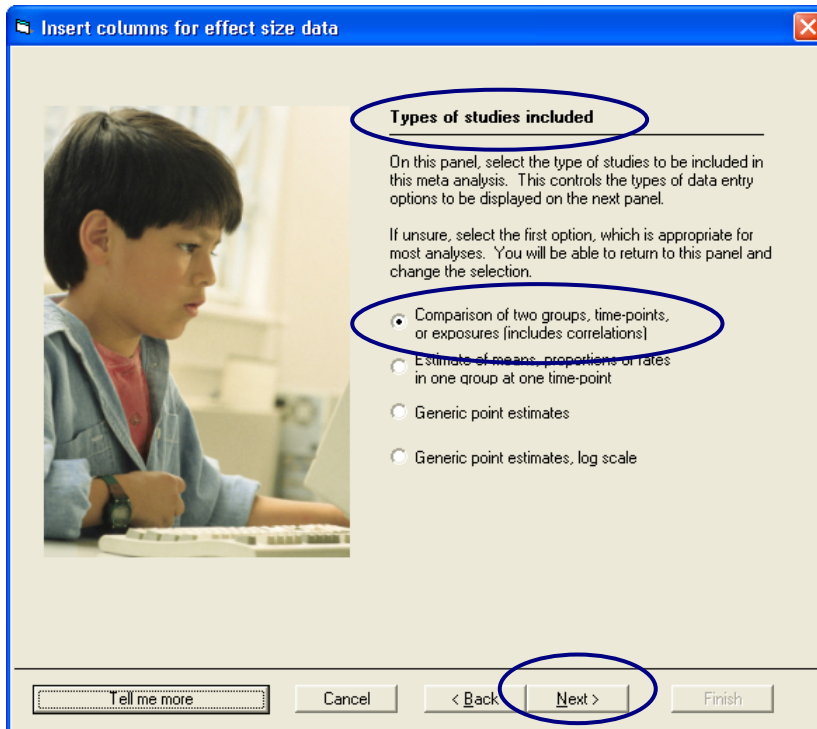
The program shows this dialog box.



→ Click NEXT



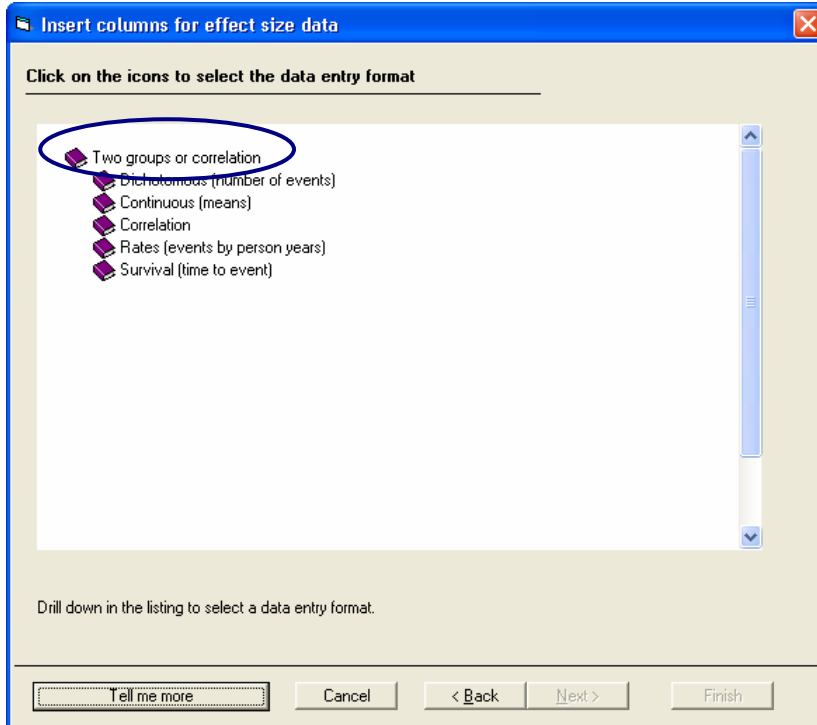
The dialog box lists four sets of effect sizes.



→ Select COMPARISON OF TWO GROUPS, TIME-POINTS, OR EXPOSURES (INCLUDES CORRELATIONS)

→ Click NEXT

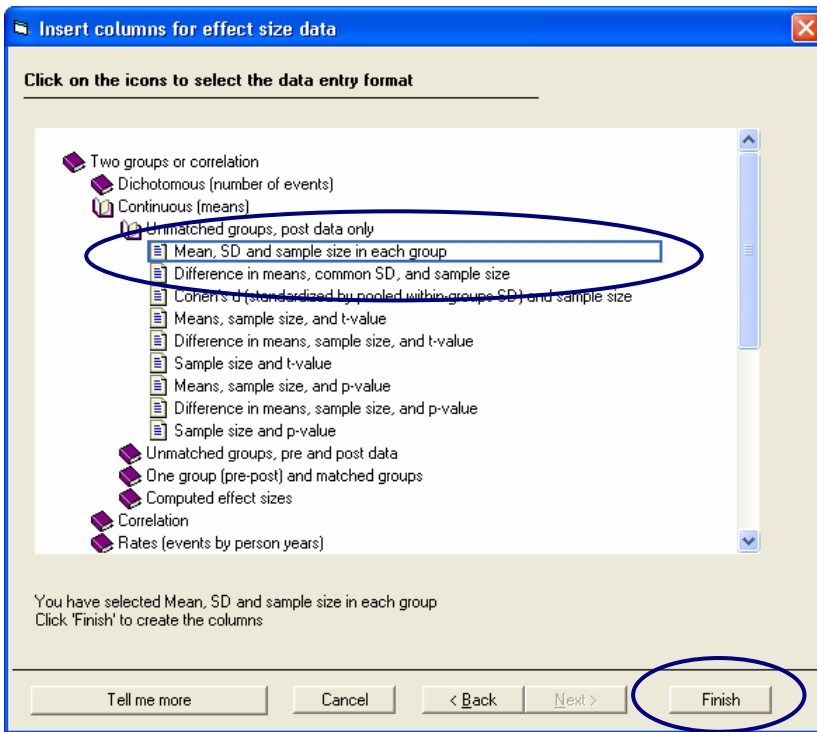
The program displays this dialog box.



→ Drill down to

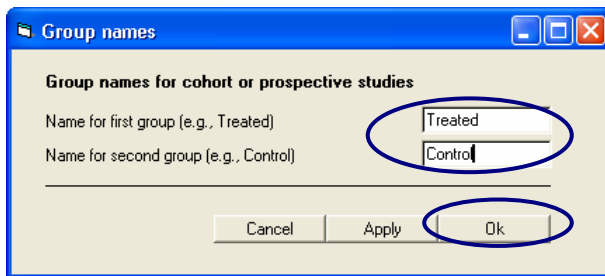
- CONTINUOUS (MEANS)
- UNMATCHED GROUPS, POST DATA ONLY
- MEAN, SD AND SAMPLE SIZE IN EACH GROUP

→ Click FINISH



The program will return to the main data-entry screen.

The program displays a dialog box that you can use to name the groups.



- Enter the names TREATED and CONTROL
- Click OK

The program displays the columns needed for the selected format (Means and standard deviations). It also labels these columns using the names Treated and Control.

You will enter data into the white columns (at left). The program will compute the effect size for each study and display that effect size in the yellow columns (at right).

Since you elected to enter means and SDs the program initially displays columns for the standardized mean difference (Cohen's  $d$ ), the bias-corrected standardized mean difference (Hedges's  $G$ ) and the raw mean difference. You can add other indices as well.

	Study name	Treated Mean	Treated Std-Dev	Treated Sample size	Control Mean	Control Std-Dev	Control Sample size	Effect direction	Std diff in means	Std Err	Hedges's g	Std Err	Difference in means	Std Err
1														
2														
3														
4														
5														
6														
7														
8														
9														

## Enter the data

- Enter the mean, SD, and N for each group as shown here
- In the column labeled EFFECT DIRECTION use the drop-down box to select AUTO

	Study name	Treated Mean	Treated Std-Dev	Treated Sample size	Control Mean	Control Std-Dev	Control Sample size	Effect direction
1	Madison	410.000	90.000	100	360.000	95.000	100	Auto
2	Moyer	480.000	80.000	100	410.000	100.000	120	Auto
3	Goldman	440.000	95.000	200	400.000	90.000	190	Auto
4	Graham	550.000	110.000	50	500.000	100.000	60	Auto
5	Manning	510.000	100.000	1000	490.000	95.000	1000	Auto
6								
7								

“Auto” means that the program will compute the mean difference as Treated minus Control. You also have the option of selecting “Positive”, in which case the program will assign a plus sign to the difference, or “Negative” in which case it will assign a minus sign to the difference.)

The program will automatically compute the effects as shown here in the yellow columns.

	Study name	Treated Mean	Treated Std-Dev	Treated Sample size	Control Mean	Control Std-Dev	Control Sample size	Effect direction	Std diff in means	Std Err	Hedges's g	Std Err	Difference in means	Std Err
1	Madison	410.000	90.000	100	360.000	95.000	100	Auto	0.540	0.144	0.538	0.143	50.000	13.086
2	Moyer	480.000	80.000	100	410.000	100.000	120	Auto	0.765	0.140	0.763	0.140	70.000	12.384
3	Goldman	440.000	95.000	200	400.000	90.000	190	Auto	0.432	0.102	0.431	0.102	40.000	9.330
4	Graham	550.000	110.000	50	500.000	100.000	60	Auto	0.478	0.194	0.474	0.193	50.000	20.040
5	Manning	510.000	100.000	1000	490.000	95.000	1000	Auto	0.205	0.045	0.205	0.045	20.000	4.362
6														
7														

## Show details for the computations

→ Double click on the value 0.540

Study name	Treated Mean	Treated Std-Dev	Treated Sample size	Control Mean	Control Std-Dev	Control Sample size	Effect direction	Std diff in means	Std Err	Hedges's g	Std Err	Difference in means	Std Err
1 Madison	410.000	90.000	100	360.000	95.000	100	Auto	0.540	0.144	0.538	0.143	50.000	13.086
2 Moyer	480.000	80.000	100	410.000	100.000	120	Auto	0.765	0.140	0.763	0.140	70.000	12.384
3 Goldman	440.000	95.000	200	400.000	90.000	190	Auto	0.432	0.102	0.431	0.102	40.000	9.381
4 Graham	550.000	110.000	50	500.000	100.000	60	Auto	0.478	0.194	0.474	0.193	50.000	20.040
5 Manning	510.000	100.000	1000	490.000	95.000	1000	Auto	0.205	0.045	0.205	0.045	20.000	4.362
6													
7													

The program shows how this value was computed.

**Data entry assistant**

Data entry | **Std diff in means** | Hedges's g | Difference in means

Starting with  
Mean, SD, N in each group

Raw difference in means

RawDiff = Mean1 - Mean2  
 $SDP = \text{Sqr}(\frac{(N1 - 1) * SD1^2 + (N2 - 1) * SD2^2}{(N1 + N2 - 2)})$   
 \*\* Option for pooled variance \*\*  
 $RawDiffSE = \text{Sqr}(1 / N1 + 1 / N2) * SDP$

RawDiff = 410.000 - 360.000 = 50.000  
 $SDP = \text{Sqr}(\frac{((100 - 1) * 90.000^2 + (100 - 1) * 95.000^2)}{(100 + 100 - 2)}) = 92.534$   
 \*\* Option for pooled variance \*\*  
 $RawDiffSE = \text{Sqr}(1 / 100 + 1 / 100) * 92.534 = 13.086$

Standardized difference in means

$StdDiff = RawDiff / SDP$   
 $StdDiffSE = \text{Sqr}(1 / N1 + 1 / N2 + StdDiff^2 / (2 * (N1 + N2)))$

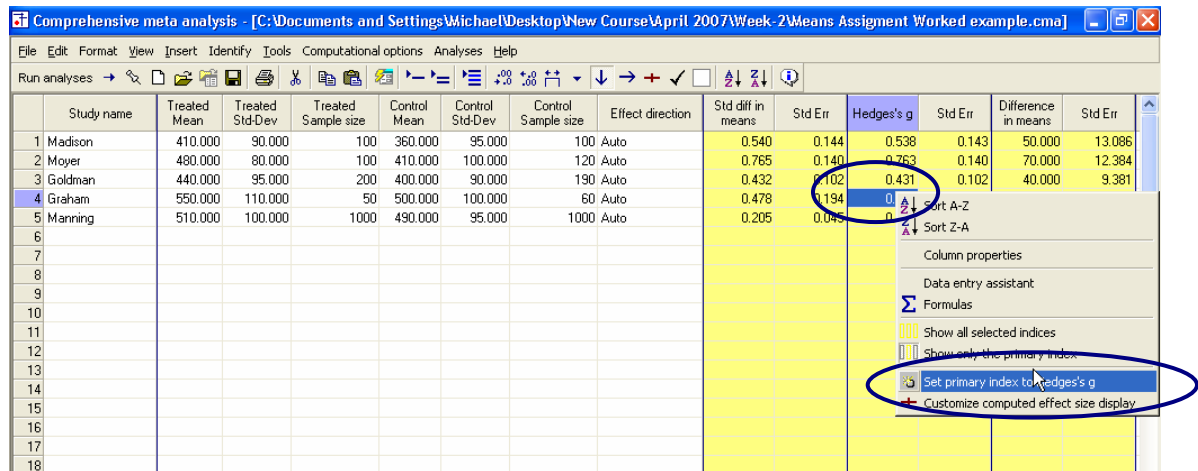
$StdDiff = 50.000 / 92.534 = 0.540$   
 $StdDiffSE = \text{Sqr}(1 / 100 + 1 / 100 + 0.540^2 / (2 * (100 + 100))) = 0.144$

< Home >

## Set the default index to Hedges's G

At this point, the program has displayed three indices of effect size – Cohen's d, Hedges's G, and the raw mean difference.

You have the option of adding additional indices, and/or specifying which index should be used as the "Default" index when you run the analysis.



The screenshot shows the 'Comprehensive meta analysis' software interface. The main window displays a table with columns for Study name, Treated Mean, Treated Std-Dev, Treated Sample size, Control Mean, Control Std-Dev, Control Sample size, Effect direction, Std diff in means, Std Err, Hedges's g, Std Err, Difference in means, and Std Err. The 'Hedges's g' column is highlighted in blue, and a context menu is open over it. The menu options include 'Sort A-Z', 'Sort Z-A', 'Column properties', 'Data entry assistant', 'Formulas', 'Show all selected indices', 'Show only the primary index', 'Set primary index to Hedges's g', and 'Customize computed effect size display'. The 'Set primary index to Hedges's g' option is circled in red.

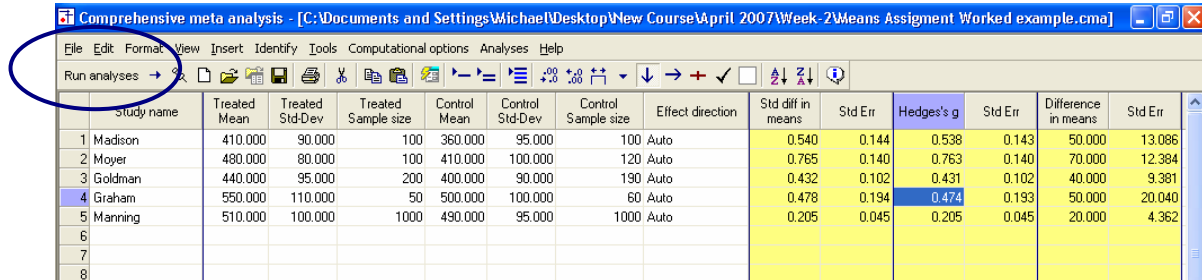
Study name	Treated Mean	Treated Std-Dev	Treated Sample size	Control Mean	Control Std-Dev	Control Sample size	Effect direction	Std diff in means	Std Err	Hedges's g	Std Err	Difference in means	Std Err
1 Madison	410.000	90.000	100	360.000	95.000	100	Auto	0.540	0.144	0.538	0.143	50.000	13.086
2 Moyer	480.000	80.000	100	410.000	100.000	120	Auto	0.765	0.140	0.762	0.140	70.000	12.384
3 Goldman	440.000	95.000	200	400.000	90.000	190	Auto	0.432	0.102	0.431	0.102	40.000	9.381
4 Graham	550.000	110.000	50	500.000	100.000	60	Auto	0.478	0.194	0.478	0.194		
5 Manning	510.000	100.000	1000	490.000	95.000	1000	Auto	0.205	0.045	0.205	0.045		

- Right-click on the column for Hedges's G
- Click SET PRIMARY INDEX TO HEDGES'S G



## Run the analysis

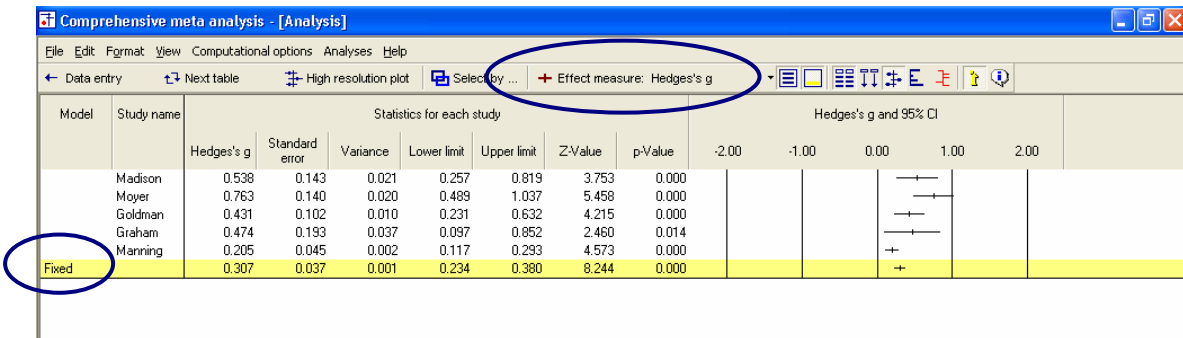
→ Click RUN ANALYSES



Study name	Treated Mean	Treated Std-Dev	Treated Sample size	Control Mean	Control Std-Dev	Control Sample size	Effect direction	Std diff in means	Std Err	Hedges's g	Std Err	Difference in means	Std Err
1 Madison	410.000	90.000	100	360.000	95.000	100	Auto	0.540	0.144	0.538	0.143	50.000	13.086
2 Moyer	480.000	80.000	100	410.000	100.000	120	Auto	0.765	0.140	0.763	0.140	70.000	12.384
3 Goldman	440.000	95.000	200	400.000	90.000	190	Auto	0.432	0.102	0.431	0.102	40.000	9.381
4 Graham	550.000	110.000	50	500.000	100.000	60	Auto	0.478	0.194	0.474	0.193	50.000	20.040
5 Manning	510.000	100.000	1000	490.000	95.000	1000	Auto	0.205	0.045	0.205	0.045	20.000	4.362
6													
7													
8													

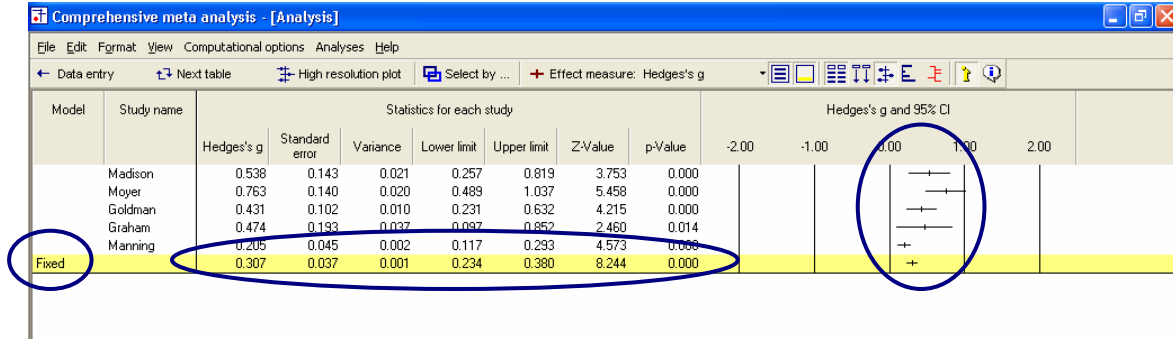
The program displays this screen.

- The default effect size is Hedges's G
- The default model is fixed effect



Model	Study name	Statistics for each study							Hedges's g and 95% CI				
		Hedges's g	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value	-2.00	-1.00	0.00	1.00	2.00
	Madison	0.538	0.143	0.021	0.257	0.819	3.753	0.000					
	Moyer	0.763	0.140	0.020	0.489	1.037	5.458	0.000					
	Goldman	0.431	0.102	0.010	0.231	0.632	4.215	0.000					
	Graham	0.474	0.193	0.037	0.097	0.852	2.460	0.014					
	Manning	0.205	0.045	0.002	0.117	0.293	4.573	0.000					
Fixed		0.307	0.037	0.001	0.234	0.380	8.244	0.000					

The screen should look like this.



We can immediately get a sense of the studies and the combined effect. For example,

- All the effects fall on the positive side of zero, in the range of 0.205 to 0.763
- Some studies are clearly more precise than others. The confidence interval for Graham is about four times as wide as the one for Manning, with the other three studies falling somewhere in between
- The combined effect is 0.307 with a 95% confidence interval of 0.234 to 0.380

## Customize the screen

We want to hide the columns for the lower and upper limit and z-value.

- Right-click on one of the “Statistics” columns
- Select CUSTOMIZE BASIC STATS

The screenshot shows the 'Comprehensive meta analysis - [Analysis]' window. The main data table is as follows:

Model	Study name	Statistics for each study							Hedges's g and 95% CI				
		Hedges's g	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value	-2.00	-1.00	0.00	1.00	2.00
	Madison	0.538	0.143	0.021	0.257	0.819	3.753	0.000					
	Moyer	0.763	0.140	0.020	0.489			0.000					
	Goldman	0.431	0.102	0.010	0.231			0.000					
	Graham	0.474	0.193	0.037	0.097			0.014					
	Manning	0.205	0.045	0.002	0.117			0.000					
Fixed		0.307	0.037	0.001	0.234			0.000					

A context menu is open over the 'Lower limit' column of the 'Fixed' row, with the 'Customize basic stats' option selected. Other menu options include 'Sort Lo-Hi by Upper limit', 'Sort Hi-Lo by Upper limit', and 'Show/hide basic stats'.

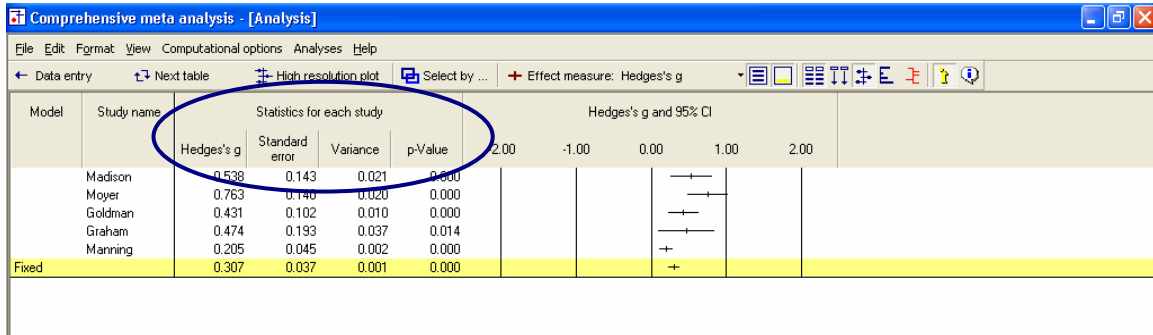
- Assign check-marks as shown here
- Click OK

The 'Customize display' dialog box is shown with the following settings:

Show	Decimals	Alignment
<input checked="" type="checkbox"/> All columns in this block		
<input checked="" type="checkbox"/> Hedges's g	Auto	Auto
<input checked="" type="checkbox"/> Standard error	Auto	Auto
<input checked="" type="checkbox"/> Variance	Auto	Auto
<input type="checkbox"/> Lower limit	Auto	Auto
<input type="checkbox"/> Upper limit	Auto	Auto
<input type="checkbox"/> Z-Value	Auto	Auto
<input checked="" type="checkbox"/> p-Value	Auto	Auto

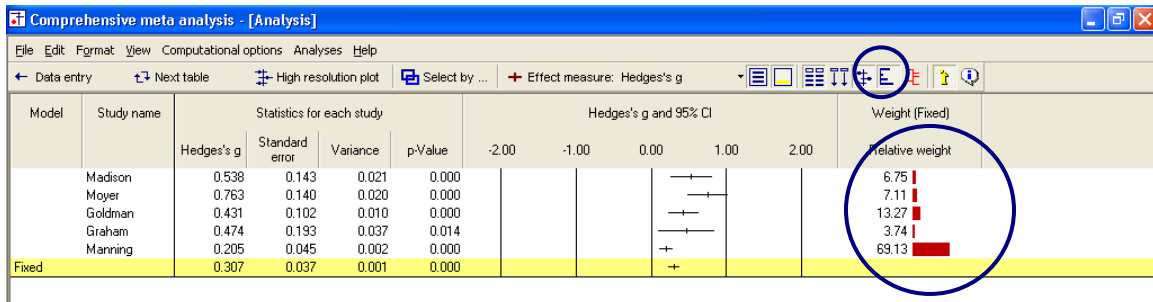
The 'Show' column is circled in blue, and the 'Ok' button is also circled in blue.

The program has hidden some of the columns, leaving us more room to work with on the display.



## Display weights

→ Click the tool for SHOW WEIGHTS



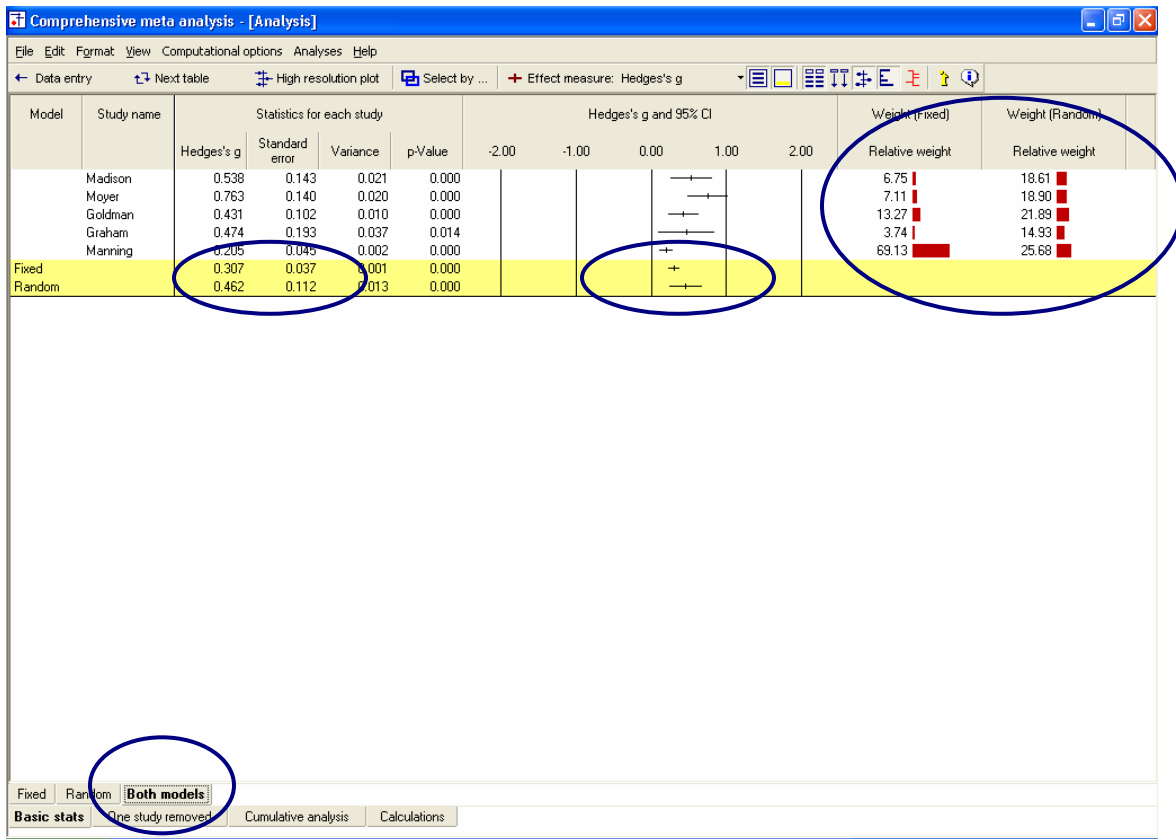
The program now shows the relative weight assigned to each study for the fixed effect analysis. By “relative weight” we mean the weights as a percentage of the total weights, with all relative weights summing to 100%.

For example, Madison was assigned a relative weight of 6.75% while Manning was assigned a relative weight of 69.13%.

## Compare the fixed effect and random effects models

→ At the bottom of the screen, select BOTH MODELS

- The program shows the combined effect and standard error for both fixed and random effects models
- The program shows weights for both the fixed effect and the random effects models



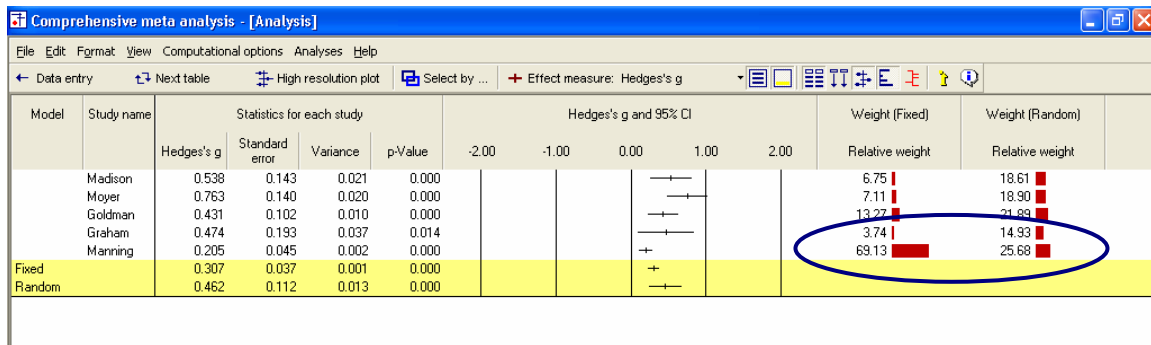
## Impact of model on study weights

The Manning study, with a large sample size (N=1000 per group) is assigned 69% of the weight under the fixed effect model but only 26% of the weight under the random effects model.

This follows from the logic of fixed and random effects models explained earlier.

Under the fixed effect model we assume that all studies are estimating the same value and this study yields a better estimate than the others, so we take advantage of that.

Under the random effects model we assume that each study is estimating a unique effect. The Manning study yields a precise estimate of its population, but that population is only one of many, and we don't want it to dominate the analysis. Therefore, we assign it 26% of the weight. This is more than the other studies, but not the dominant weight that we gave it under fixed effects.



## Impact of model on the combined effect

As it happens, the Manning study has a low effect size. Under the fixed effect model, where this study dominates the weights, it pulls the effect size down to 0.31. Under the random effects model, it still pulls the effect size down, but only to 0.46.

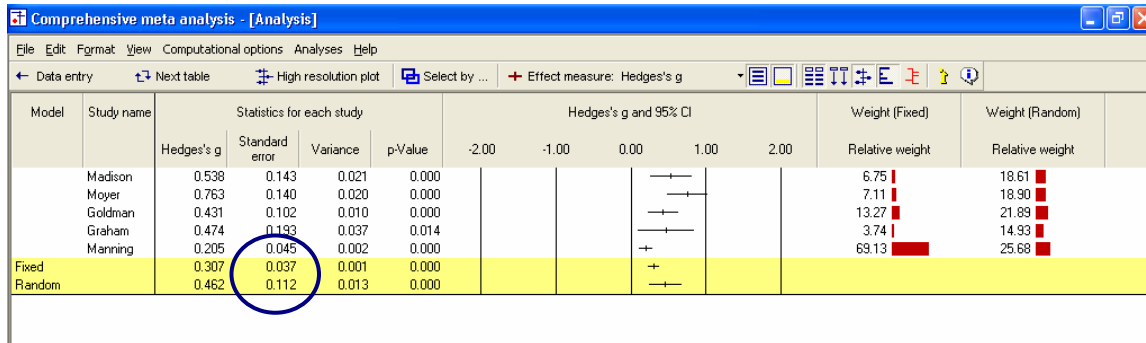
Model	Study name	Statistics for each study				Hedges's g and 95% CI					Weight (Fixed)		Weight (Random)	
		Hedges's g	Standard error	Variance	p-Value	-2.00	-1.00	0.00	1.00	2.00	Relative weight	Relative weight	Relative weight	Relative weight
	Madison	0.538	0.143	0.021	0.000				+		6.75	18.61		
	Moyer	0.763	0.140	0.020	0.000				+		7.11	18.90		
	Goldman	0.431	0.102	0.010	0.000				+		13.27	21.89		
	Graham	0.474	0.193	0.037	0.014				+		3.74	14.93		
	Manning	0.205	0.045	0.002	0.000				+		69.13	25.68		
Fixed		0.307	0.037	0.001	0.000				+					
Random		0.462	0.112	0.013	0.000				+					

## Impact of model on the confidence interval

Under fixed effect, we “set” the between-studies dispersion to zero. Therefore, for the purpose of estimating the mean effect, the only source of uncertainty is within-study error. With a combined total near 1500 subjects per group the within-study error is small, so we have a precise estimate of the combined effect. The standard error is 0.037.

Under random effects, dispersion between studies is considered a real source of uncertainty. And, there is a lot of it. The fact that these five studies vary so much one from the other tells us that the effect will vary depending on details that vary randomly from study to study. If the persons who performed these studies happened to use older subjects, or a shorter duration, for example, the effect size would have changed.

While this dispersion is “real” in the sense that it is caused by real differences among the studies, it nevertheless represents error if our goal is to estimate the mean effect. For computational purposes, the variance due to between-study differences is included in the error term. In our example we have only five studies, and the effect sizes do vary. Therefore, our estimate of the mean effect is not terribly precise, as reflected in a standard error of 0.112, which is about three times that of the fixed effect value.

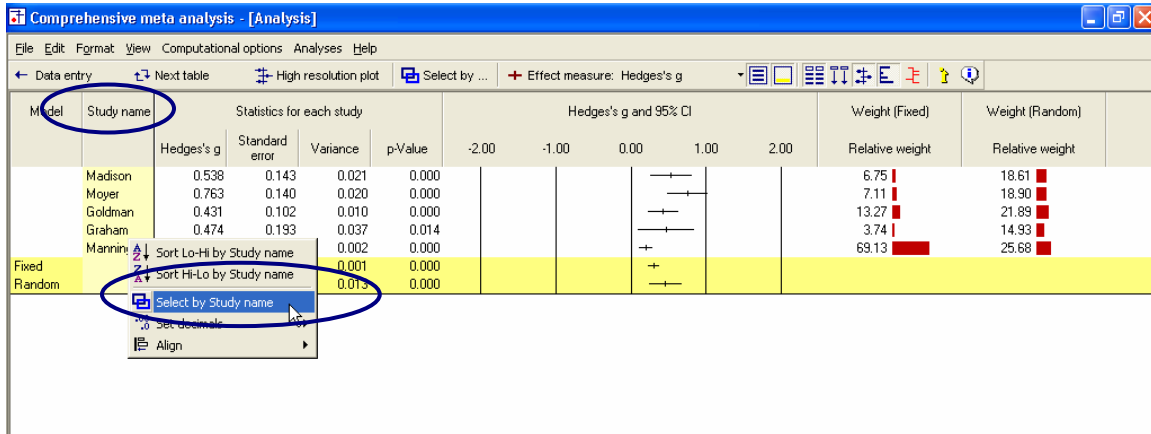




## What would happen if we eliminated Manning?

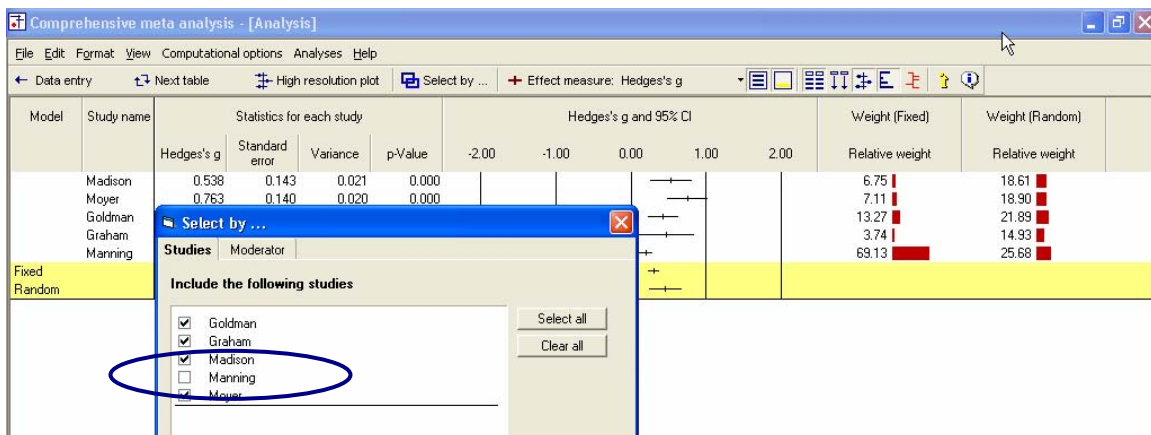
Manning was the largest study, and also the study with the smallest effect size. To better understand the impact of this study under the two models, let's see what would happen if we were to remove this study from the analysis.

- Right-click on STUDY NAME
- Select SELECT BY STUDY NAME

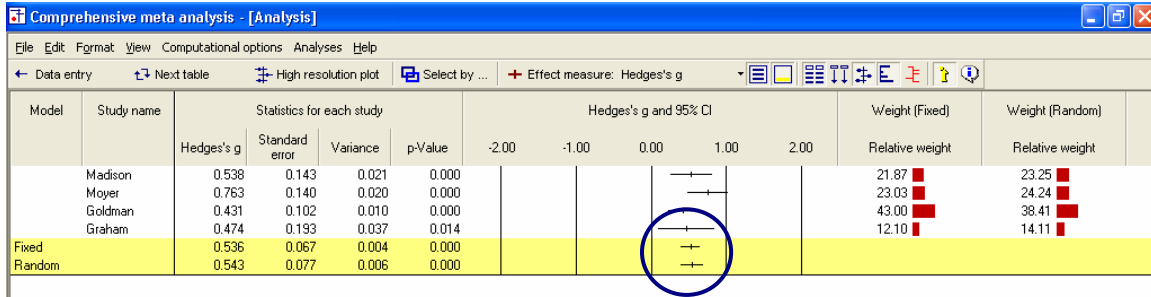


The program opens a dialog box with the names of all studies.

- Remove the check from Manning
- Click OK



The analysis now looks like this.

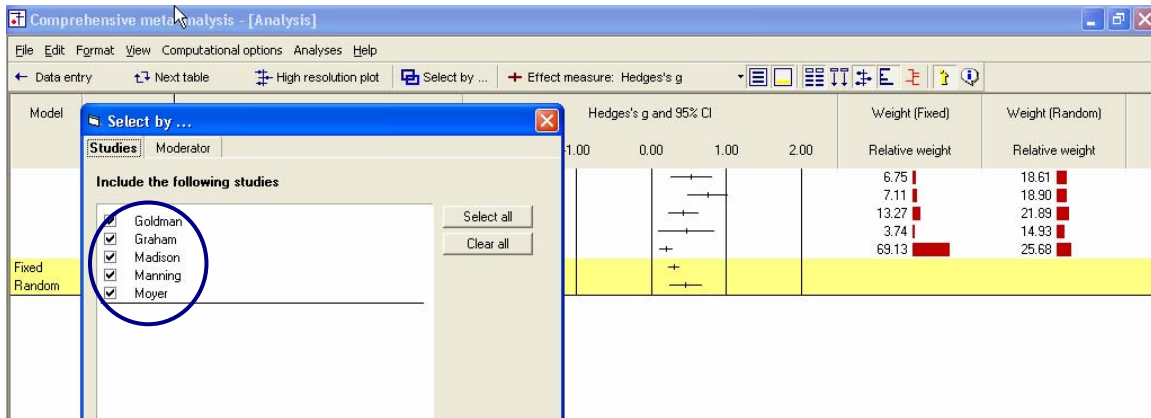


For both the fixed effect and random effects models, the combined effect is now close to 0.54.

- Under fixed effects Manning had pulled the effect down to 0.31
- Under random effects Manning had pulled the effect down to 0.46
- Thus, this study had a substantial impact under either model, but more so under fixed than random effects

→ Right-click on STUDY NAME

→ Add a check for Manning so the analysis again has five studies



## Additional statistics

→ Click NEXT TABLE on the toolbar

Model	Study name	Statistics for each study				Hedges's g and 95% CI					Weight (Fixed)		Weight (Random)	
		Hedges's g	Standard error	Variance	p-Value	-2.00	-1.00	0.00	1.00	2.00	Relative weight	Relative weight	Relative weight	Relative weight
	Madison	0.538	0.143	0.021	0.000			+			6.75	18.61		
	Moyer	0.763	0.140	0.020	0.000			+			7.11	18.90		
	Goldman	0.431	0.102	0.010	0.000			+			13.27	21.89		
	Graham	0.474	0.193	0.037	0.014			+			3.74	14.93		
	Manning	0.205	0.045	0.002	0.000			+			69.13	25.68		
Fixed		0.307	0.037	0.001	0.000			+						
Random		0.462	0.112	0.013	0.000			+						

The program switches to this screen.

The program shows the point estimate, standard error, variance, and confidence interval. These are the same values that had been shown on the forest plot.

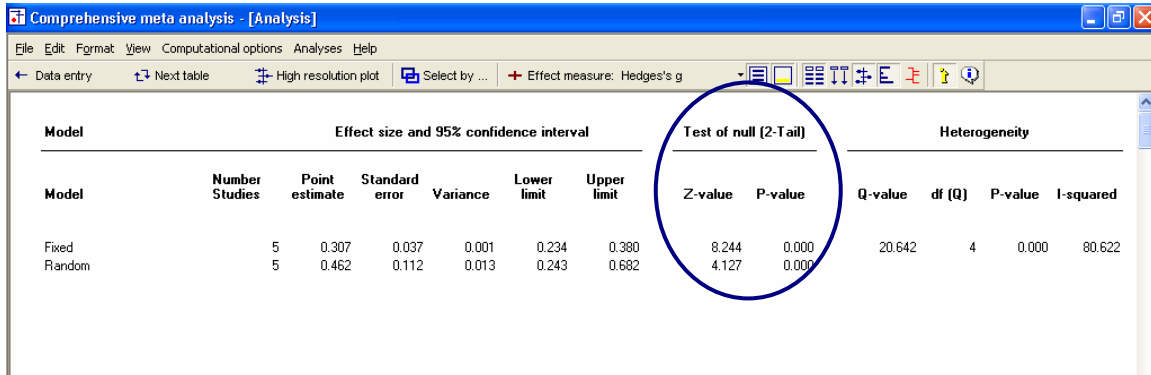
Model	Effect size and 95% confidence interval						Test of null (2-Tail)		Heterogeneity			
	Number Studies	Point estimate	Standard error	Variance	Lower limit	Upper limit	Z-value	P-value	Q-value	df (Q)	P-value	I-squared
Fixed	5	0.307	0.037	0.001	0.234	0.380	8.244	0.000	20.642	4	0.000	80.622
Random	5	0.462	0.112	0.013	0.243	0.682	4.127	0.000				

- Under fixed effect the combined effect is 0.307 with standard error of 0.037 and 95% confidence interval of 0.234 to 0.380
- Under random effects the combined effect is 0.462 with standard error of 0.112 and 95% confidence interval of 0.243 to 0.682

## Test of the null hypothesis

Under the fixed effect model the null hypothesis is that the common effect is zero. Under the random effects model the null hypothesis is that the mean of the true effects is zero.

In either case, the null hypothesis is tested by the z-value, which is computed as  $G/SE$  for the corresponding model.



Model	Effect size and 95% confidence interval						Test of null (2-Tail)		Heterogeneity			
	Number Studies	Point estimate	Standard error	Variance	Lower limit	Upper limit	Z-value	P-value	Q-value	df (Q)	P-value	I-squared
Fixed	5	0.307	0.037	0.001	0.234	0.380	8.244	0.000	20.642	4	0.000	80.622
Random	5	0.462	0.112	0.013	0.243	0.682	4.127	0.000				

For the fixed effect model

$$Z = \frac{.307}{.037} = 8.244$$

For the random effects model

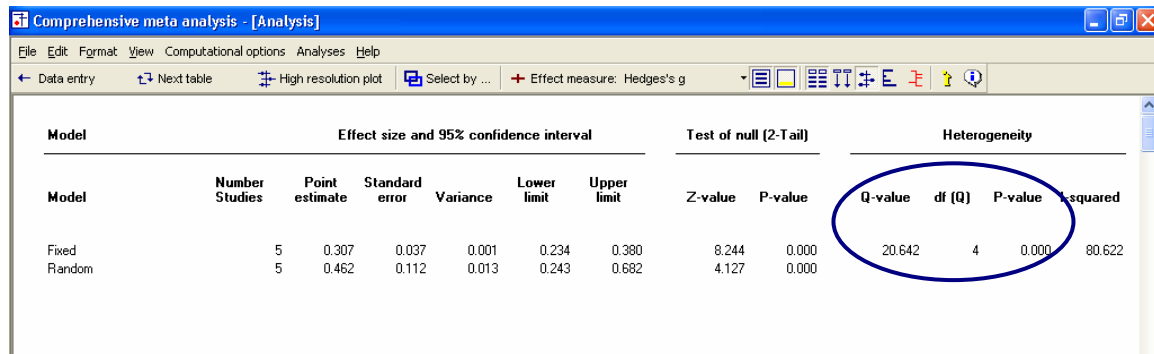
$$Z^* = \frac{0.462}{0.112} = 4.127$$

In either case, the two-tailed p-value is  $< 0.001$ .

## Test of the heterogeneity

The null hypothesis for heterogeneity is that the studies share a common effect size.

The statistics in this section address the question of whether the observed dispersion among effects exceeds the amount that would be expected by chance.



Model	Effect size and 95% confidence interval						Test of null (2-Tail)		Heterogeneity			
	Number Studies	Point estimate	Standard error	Variance	Lower limit	Upper limit	Z-value	P-value	Q-value	df (Q)	P-value	I-squared
Fixed	5	0.307	0.037	0.001	0.234	0.380	8.244	0.000	20.642	4	0.000	80.622
Random	5	0.462	0.112	0.013	0.243	0.682	4.127	0.000				

The Q statistic reflects the observed dispersion. Under the null hypothesis that all studies share a common effect size, the expected value of Q is equal to the degrees of freedom (the number of studies minus 1), and is distributed as Chi-square with  $df = k - 1$  (where  $k$  is the number of studies).

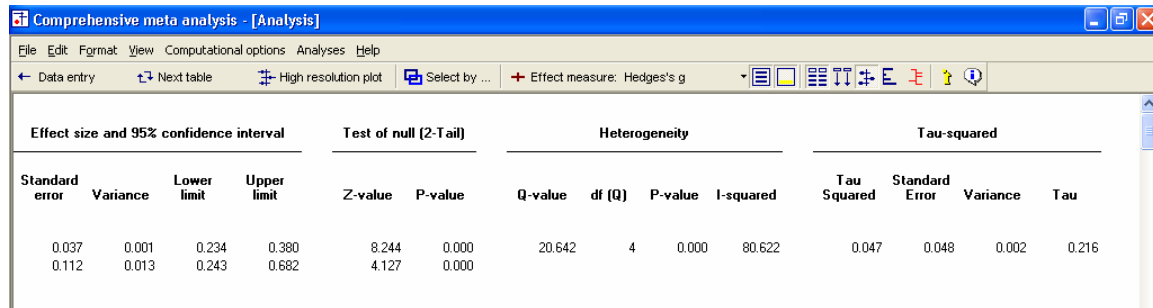
- The Q statistic is 20.64, as compared with an expected value of 4
- The p-value is  $< 0.001$

As discussed in the text, the decision to use a random effects model should be based on our understanding of how the studies were acquired, and should not depend on a statistically significant p-value for heterogeneity. In any event, however, this p-value tells us that the observed dispersion cannot be attributed to random error, and there is real variance among the effects. A fixed effect model would not fit the data.

## Quantifying the heterogeneity

While Q is meant to test the null hypothesis that there is no dispersion across effect sizes, we want also to quantify this dispersion. For this purpose we would turn to I-squared and tau-squared.

→ To see these statistics, scroll the screen toward the right



The screenshot shows the 'Comprehensive meta analysis - [Analysis]' window. The main table displays the following data:

Effect size and 95% confidence interval				Test of null (2-Tail)		Heterogeneity				Tau-squared			
Standard error	Variance	Lower limit	Upper limit	Z-value	P-value	Q-value	df (Q)	P-value	I-squared	Tau Squared	Standard Error	Variance	Tau
0.037	0.001	0.234	0.380	8.244	0.000	20.642	4	0.000	80.622	0.047	0.048	0.002	0.216
0.112	0.013	0.243	0.682	4.127	0.000								

- I-squared is 80.6, which means that 80% of the observed variance between studies is due to real differences in the effect size. Only about 20% of the observed variance would have been expected based on random error.
- Tau-squared is 0.047. This is the “Between studies” variance that was used in computing weights.

The Q statistic and tau-squared are reported on the fixed effect line, and not on the random effects line.

These values are displayed on the fixed effect line because they are computed using fixed effect weights. These values are used in both fixed and random effects analyses, but for different purposes.

For the fixed effect analysis Q addresses the question of whether or not the fixed effect model fits the data (is it reasonable to assume that tau-squared is actually zero). However, tau-squared is actually set to zero for the purpose of assigning weights.

For the random effects analysis, these values are actually used to assign the weights.

Return to the main analysis screen

Click NEXT TABLE again to get back to the other screen.

Model	Effect size and 95% confidence interval						Test of null (2-Tail)		Heterogeneity			
	Number Studies	Point estimate	Standard error	Variance	Lower limit	Upper limit	Z-value	P-value	Q-value	df (Q)	P-value	I-squared
Fixed	5	0.307	0.037	0.001	0.234	0.380	8.244	0.000	20.642	4	0.000	80.622
Random	5	0.462	0.112	0.013	0.243	0.682	4.127	0.000				

Your screen should look like this.

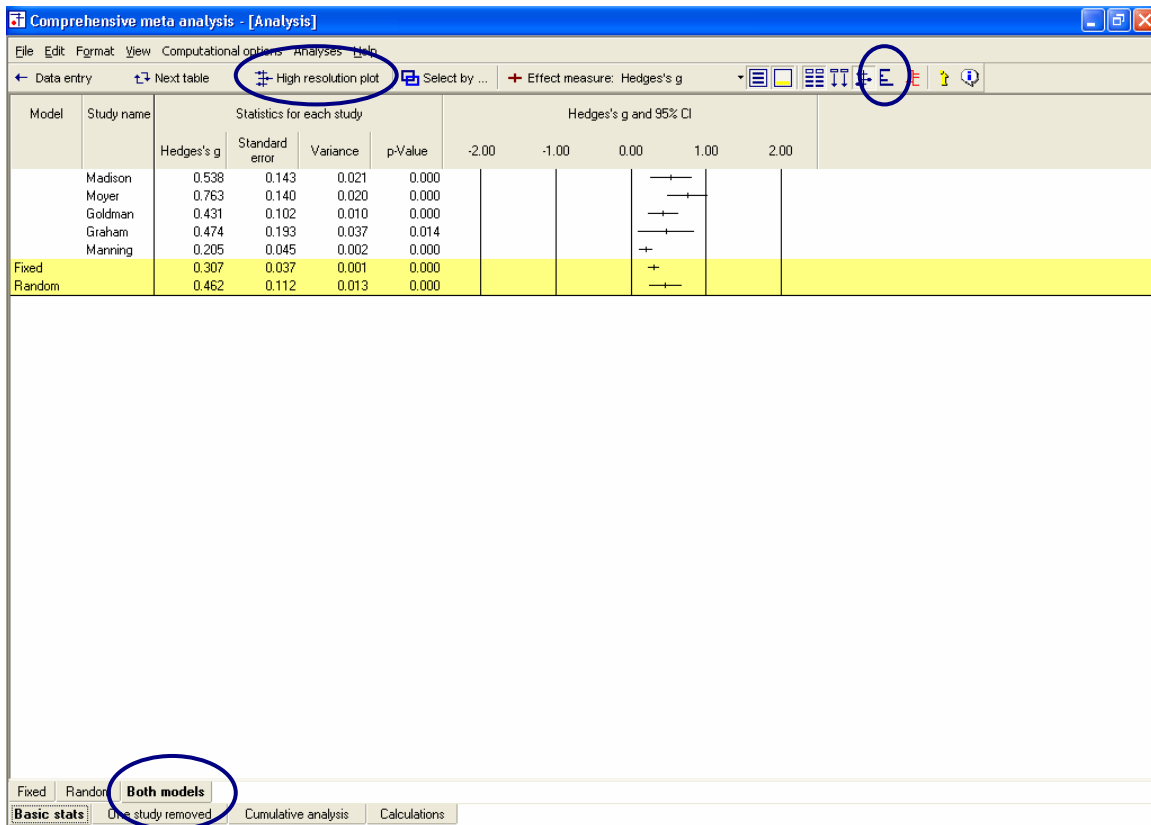
Model	Study name	Statistics for each study				Hedges's g and 95% CI					Weight (Fixed)		Weight (Random)
		Hedges's g	Standard error	Variance	p-Value	-2.00	-1.00	0.00	1.00	2.00	Relative weight	Relative weight	
	Madison	0.538	0.143	0.021	0.000						6.75	18.61	
	Moyer	0.763	0.140	0.020	0.000						7.11	18.90	
	Goldman	0.431	0.102	0.010	0.000						13.27	21.89	
	Graham	0.474	0.193	0.037	0.014						3.74	14.93	
	Manning	0.205	0.045	0.002	0.000						69.13	25.68	
Fixed		0.307	0.037	0.001	0.000								
Random		0.462	0.112	0.013	0.000								

## High-resolution plots

To this point we've used bar graphs to show the weight assigned to each study.

Now, we'll switch to a high-resolution plot, where the weight assigned to each study will determine the size of the symbol representing that study.

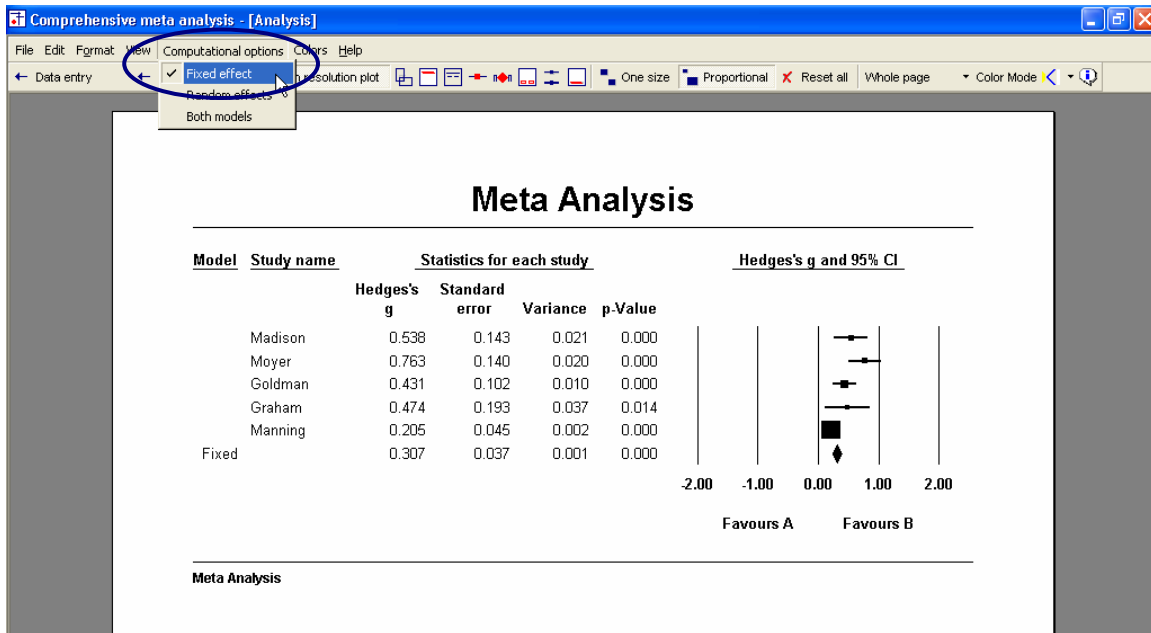
- Select BOTH MODELS at the bottom of the screen
- Unclick the SHOW WEIGHTS button on the toolbar
- Click HIGH-RESOLUTION PLOT on the toolbar



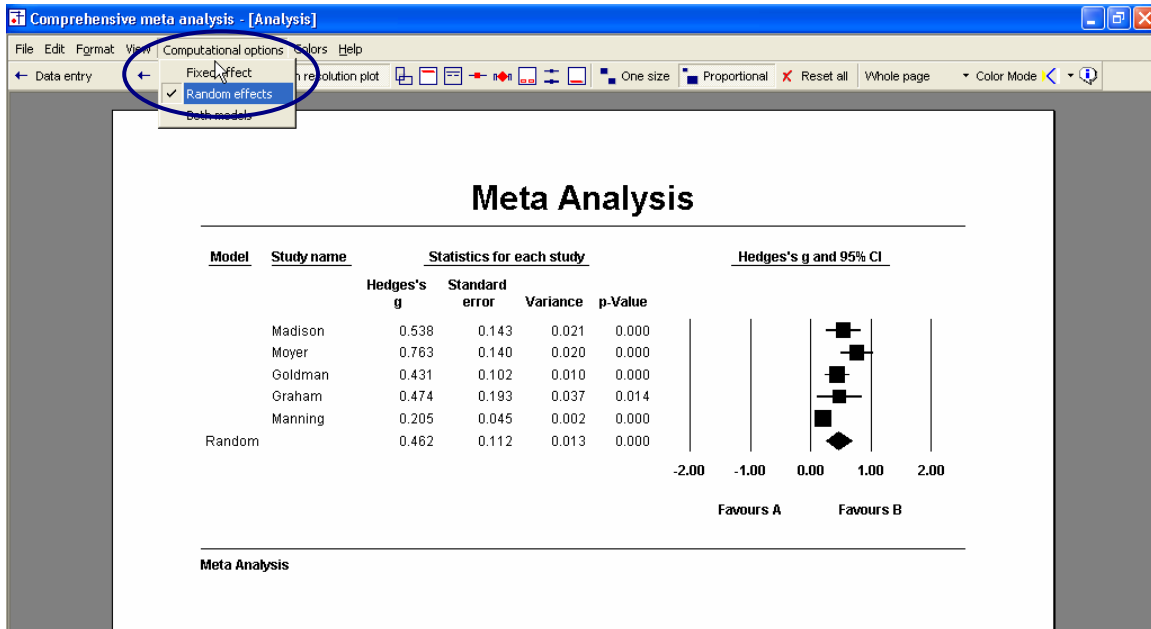


The program shows this screen.

→ Select COMPUTATIONAL OPTIONS > FIXED EFFECT



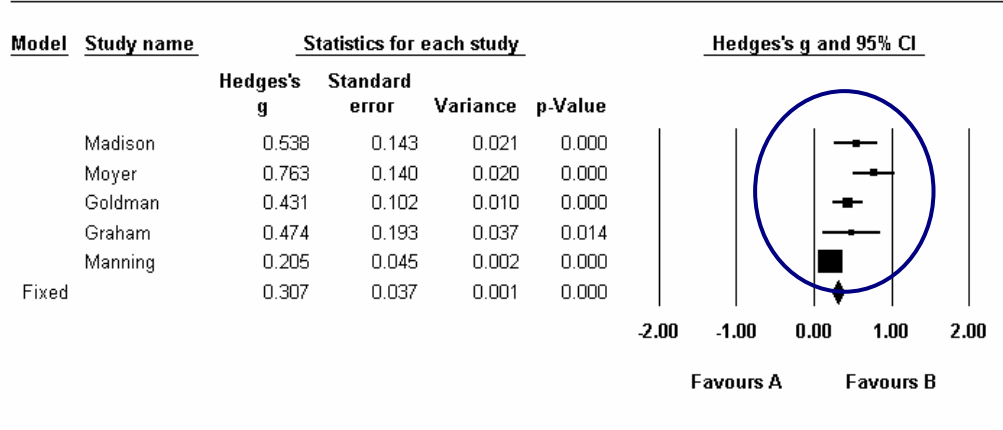
→ Select COMPUTATIONAL OPTIONS > RANDOM EFFECTS



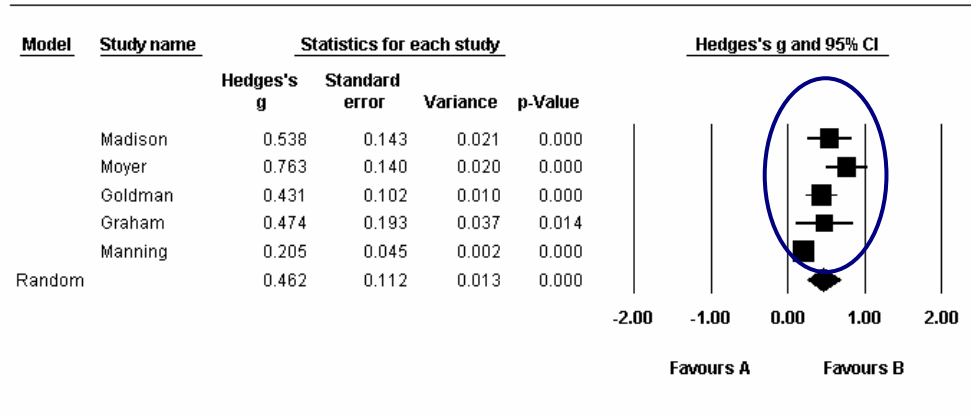
Compare the weights.

In the first plot, using fixed effect weights, the area of the Manning box is about 10 times that of Madison. In the second, using random effects weights, the area of the Manning box is only about 30% larger than Madison.

## Meta Analysis



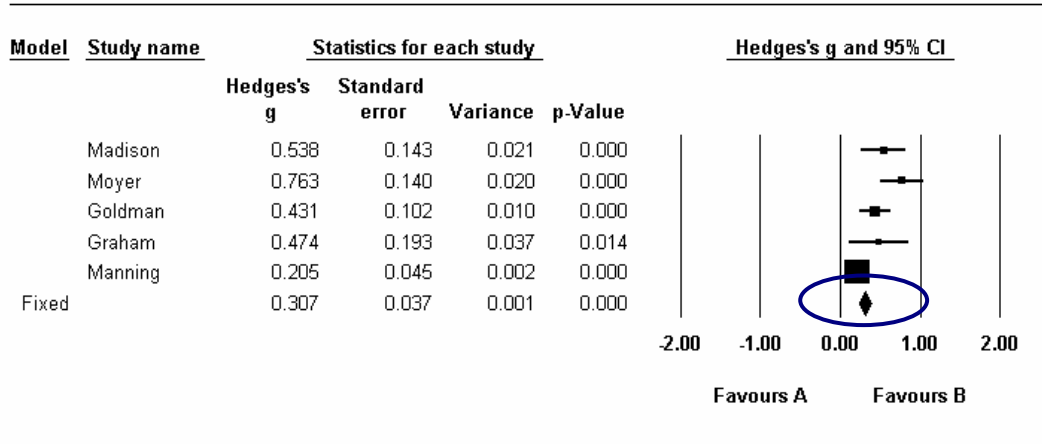
## Meta Analysis



Meta Analysis

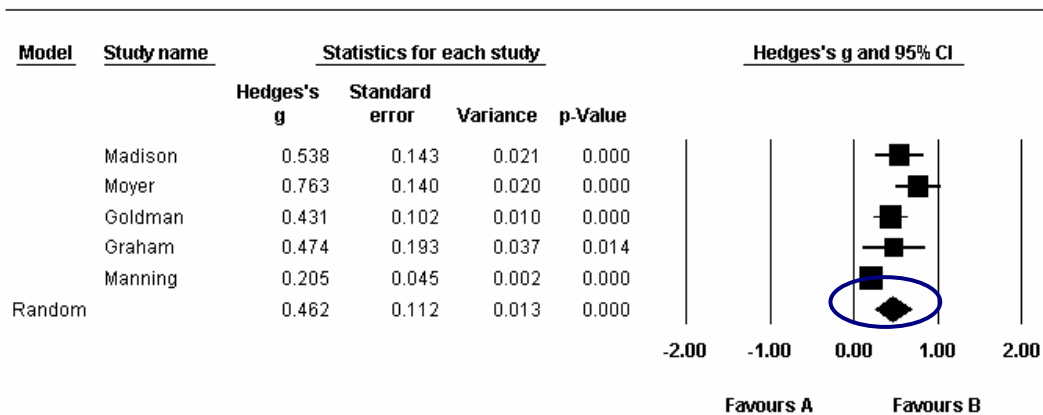
Compare the combined effects and standard error.

## Meta Analysis



Meta Analysis

## Meta Analysis



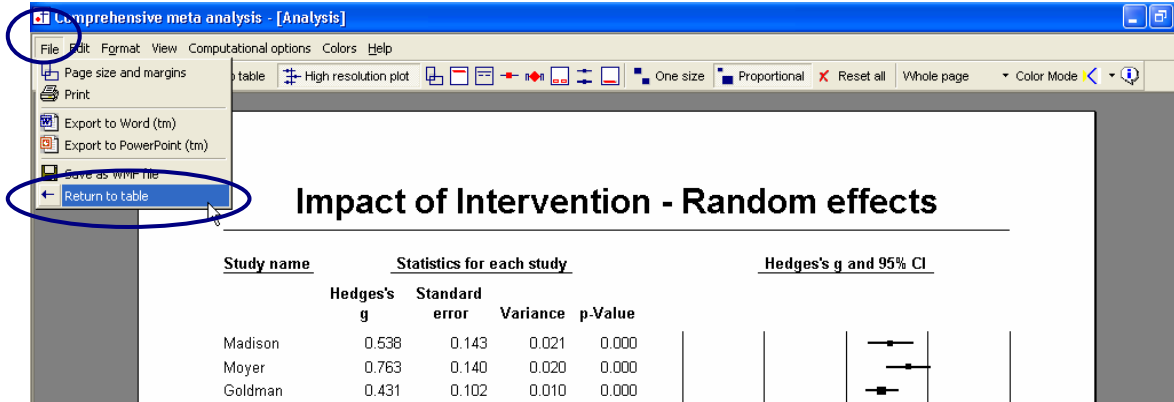
Meta Analysis

In the first case (fixed) Manning is given substantial weight and pulls the combined effect down to 0.307. In the second case (random) Manning is given less weight, and the combined effect is 0.462.

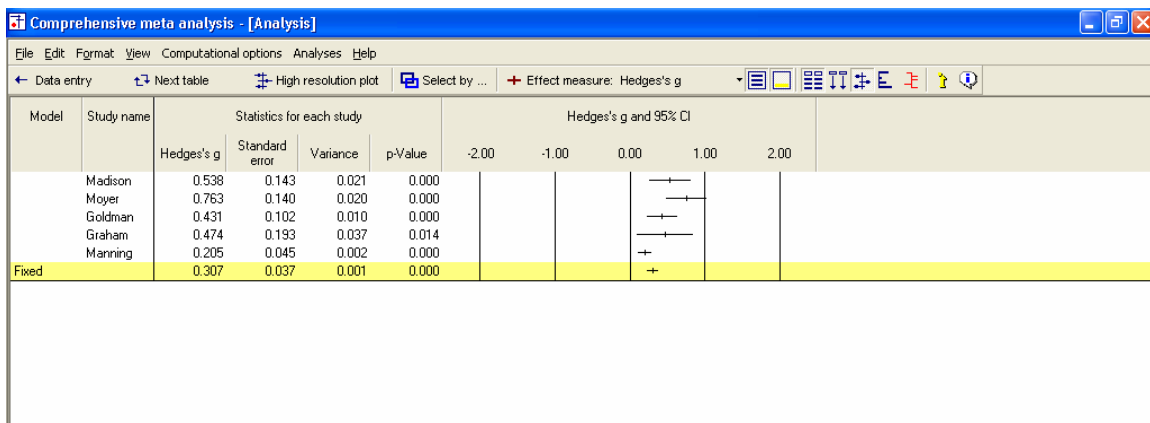
In the first case (fixed) the only source of error is the error within studies. The standard error of the combined effect is 0.037 and the confidence interval about the combined effect is relatively narrow. In the second case (random) the fact that the true effect varies from study to study introduces another level of uncertainty to our estimate. The standard error of the combined effect is 0.112,

and the confidence interval about the combined effect is about three times as wide as that for the fixed effect analysis.

→ Click FILE > RETURN TO TABLE



The program returns to the main analysis screen.



## Computational details

The program allows you to view details of the computations.

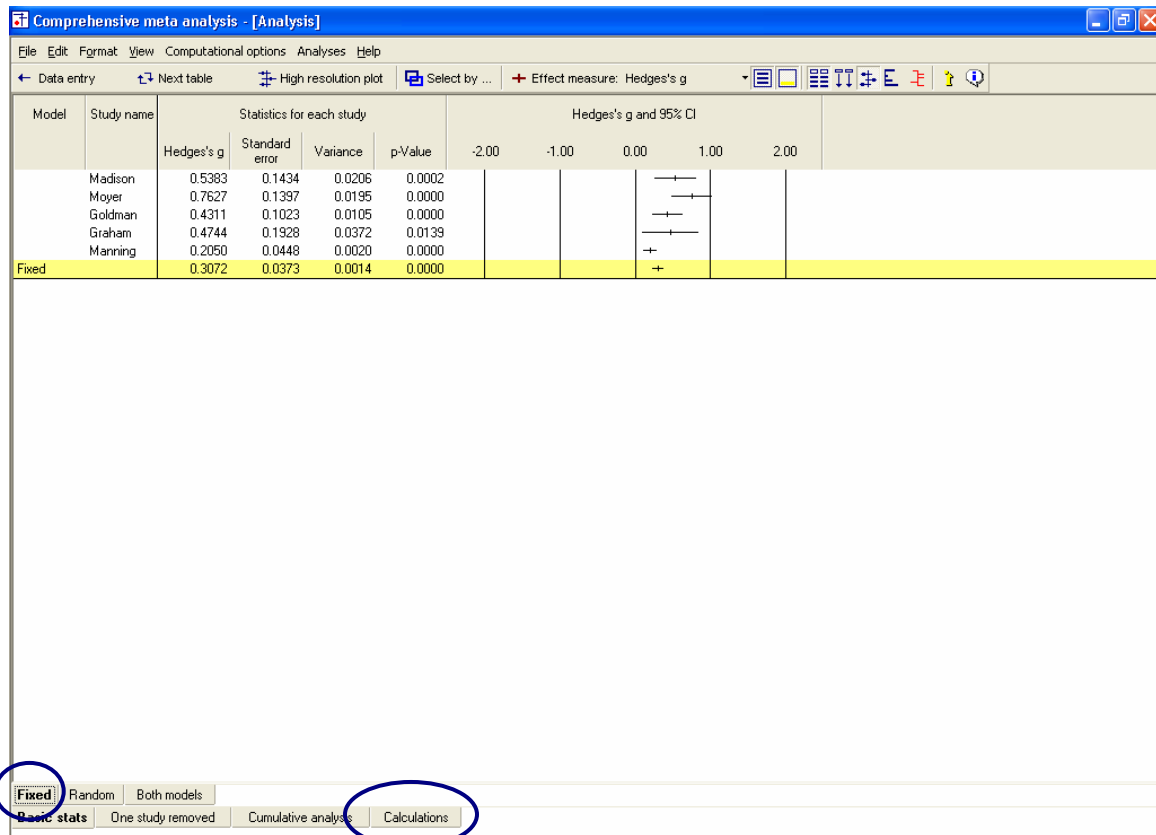
### Computational details for the fixed effect analysis

→ Select **FORMAT > INCREASE DECIMALS** on the menu

This has no effect on the computations, which are always performed using all significant digits, but it makes the example easier to follow.

→ On the bottom of the screen select **FIXED**

→ On the bottom of the screen select **CALCULATIONS**



The program switches to this display.

Model	Study name	Point	Calculations (Fixed)						
			Study Variance	Tau <sup>2</sup> Within	Tau <sup>2</sup> Between	Total Variance	IV-Weight	W	T*W
	Madison	0.5383	0.0206	0.0000	0.0000	0.0206	48.6074	48.6074	26.1651
	Moyer	0.7627	0.0195	0.0000	0.0000	0.0195	51.2049	51.2049	39.0548
	Goldman	0.4311	0.0105	0.0000	0.0000	0.0105	95.5858	95.5858	41.2107
	Graham	0.4744	0.0372	0.0000	0.0000	0.0372	26.8954	26.8954	12.7600
	Manning	0.2050	0.0020	0.0000	0.0000	0.0020	497.7593	497.7593	102.0326
		2.4116	0.0898	0.0000	0.0000	0.0898	720.0528	720.0528	221.2233

For the first study, Madison

$$\text{RawDifference} = 410 - 360 = 50$$

$$SD_{\text{Pooled}} = \sqrt{\frac{((100-1)(90^2)) + ((100-1)(95^2))}{100 + 100 - 2}} = 92.534$$

$$d = \frac{50}{92.534} = 0.5403$$

$$SE_d = \sqrt{\frac{1}{100} + \frac{1}{100} + \frac{0.5403^2}{2 * (100 + 100)}} = 0.1440$$

$$J = 1 - \left( \frac{3}{4 * 198 - 1} \right) = 0.996$$

$$G = 0.996 * 0.5403 = 0.5383$$

$$SE_G = 0.996 * 0.1440 = 0.1434$$

$$SE_G^2 = 0.1434^2 = 0.0206$$

The weight is computed as

$$w_1 = \frac{1}{0.0206 + 0.0000} = 48.6074$$

where the second term in the denominator represents tau-squared, which has been set to zero for the fixed effect analysis.

$$T_1 w_1 = 0.5383 * 48.6074 = 26.1651$$

and so on for the other studies. Then, working with the sums (in the last row)

$$\bar{T}_{\bullet} = \frac{221.2233}{720.0528} = 0.3072$$

$$v_{\bullet} = \frac{1}{720.0528} = 0.0014$$

$$SE(\bar{T}_{\bullet}) = \sqrt{0.0014} = 0.0373$$

$$\text{Lower Limit} = 0.3072 - 1.96 * 0.0373 = 0.2342$$

$$\text{Upper Limit} = 0.3072 + 1.96 * 0.0373 = 0.3803$$

$$Z = \frac{0.3072}{0.0373} = 8.2442$$

$$p_{2T} = 2[1 - (\Phi(|8.2442|))] < 0.0001$$

→ To switch back to the main analysis screen, click BASIC STATS at the bottom

Model	Study name	Statistics for each study				Hedges's g and 95% CI				
		Hedges's g	Standard error	Variance	p-Value	-2.00	-1.00	0.00	1.00	2.00
	Madison	0.5383	0.1434	0.0206	0.0002			+	+	
	Moyer	0.7627	0.1397	0.0195	0.0000			+	+	
	Goldman	0.4311	0.1023	0.0105	0.0000			+	+	
	Graham	0.4744	0.1328	0.0372	0.0139			+	+	
	Manning	0.2050	0.0448	0.0020	0.0000			+		
Fixed		0.3072	0.0373	0.0014	0.0000			+		

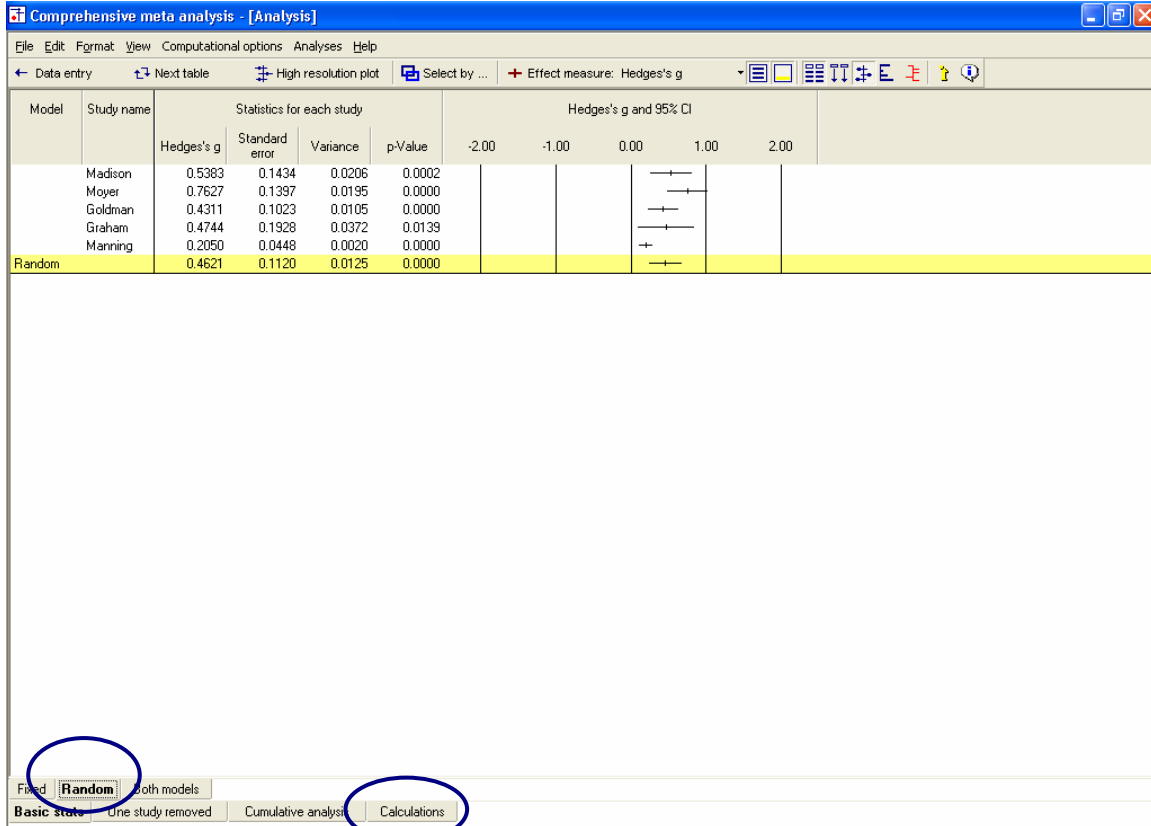
→ These are the same values presented here



## Computational details for the random effects analysis

Now, we can repeat the exercise for random effects.

- On the bottom of the screen select RANDOM
- On the bottom of the screen select CALCULATIONS



The program switches to this display.

Comprehensive meta analysis - [Analysis]									
File Edit Format View Computational options Analyses Help									
← Data entry    ↻ Next table    High resolution plot    Select by ...    + Effect measure: Hedges's g									
Model	Study name	Calculations (Random)							
		Point	Study Variance	Tau <sup>2</sup> Within	Tau <sup>2</sup> Between	Total Variance	I <sup>2</sup> -Weight	W	T*W
	Madison	0.5383	0.0286	0.0468	0.0000	0.0674	14.8356	14.8356	7.9859
	Moyer	0.7627	0.0195	0.0468	0.0000	0.0664	15.0689	15.0689	11.4933
	Goldman	0.4311	0.0105	0.0468	0.0000	0.0573	17.4537	17.4537	7.5250
	Graham	0.4744	0.0372	0.0468	0.0000	0.0840	11.9028	11.9028	5.6471
	Manning	0.2050	0.0020	0.0468	0.0000	0.0488	20.4744	20.4744	4.1969
		2.4116	0.0898	0.2342	0.0000	0.3239	79.7353	79.7353	36.8481

For the first study, Madison

$$\text{RawDifference} = 410 - 360 = 50$$

$$SD_{\text{Pooled}} = \sqrt{\frac{((100-1)(90^2)) + ((100-1)(95^2))}{100 + 100 - 2}} = 92.534$$

$$d = \frac{50}{92.534} = 0.5403$$

$$SE_d = \sqrt{\frac{1}{100} + \frac{1}{100} + \frac{0.5403^2}{2 * (100 + 100)}} = 0.1440$$

$$J = 1 - \left( \frac{3}{4 * 198 - 1} \right) = 0.996$$

$$G = 0.996 * 0.5403 = 0.5383$$

$$SE_G = 0.996 * 0.1440 = 0.1434$$

$$SE_G^2 = 0.1434^2 = .0206$$

The weight is computed as

$$w_1^* = \frac{1}{0.0206 + 0.0468} = \frac{1}{0.0674} = 14.8356$$

where the (\*) indicates that we are using random effects weights, and the second term in the denominator represents tau-squared.

$$T_1^* w_1^* = 0.5383 * 14.8356 = 7.9859$$

and so on for the other studies. Then, working with the sums (in the last row)

$$\bar{T}_\bullet^* = \frac{36.8481}{79.7353} = 0.4621$$

$$v_\bullet^* = \frac{1}{79.7353} = 0.0125$$

$$SE(\bar{T}_\bullet^*) = \sqrt{0.0125} = 0.1120$$

$$\text{Lower Limit}^* = 0.4621 - 1.96 * 0.1120 = 0.2426$$

$$\text{Upper Limit}^* = 0.4621 + 1.96 * 0.1120 = 0.6816$$

$$Z^* = \frac{0.4621}{0.1120} = 4.1266$$

$$p_{2T}^* = 2[1 - (\Phi(|4.1266|))] < 0.0001$$

(Note – The column labeled TAU-SQUARED WITHIN is actually tau-squared between studies, and the column labeled TAU-SQUARED BETWEEN is reserved for a fully random effects analysis, where we are performing an analysis of variance).

→ To switch back to the main analysis screen, click BASIC STATS at the bottom

On this screen, the values presented are the same as those computed above.

Comprehensive meta analysis - [Analysis]										
File Edit Format View Computational options Analyses Help										
← Data entry    ↕ Next table    High resolution plot    Select by ...    + Effect measure: Hedges's g										
Model	Study name	Statistics for each study				Hedges's g and 95% CI				
		Hedges's g	Standard error	Variance	p-Value	-2.00	-1.00	0.00	1.00	2.00
	Madison	0.5383	0.1434	0.0206	0.0002					
	Moyer	0.7627	0.1397	0.0195	0.0000					
	Goldman	0.4311	0.1023	0.0105	0.0000					
	Graham	0.4744	0.1920	0.0372	0.0139					
	Manning	0.2050	0.0448	0.0020	0.0000					
Random		0.4621	0.1120	0.0125	0.0000					

This is the conclusion of the exercise for means.



## Example 3 – Correlational Data

This appendix shows how to use Comprehensive Meta-Analysis (CMA) to perform a meta-analysis for correlations using fixed and random effects models.

To download a free trial copy of CMA go to [www.Meta-Analysis.com](http://www.Meta-Analysis.com)

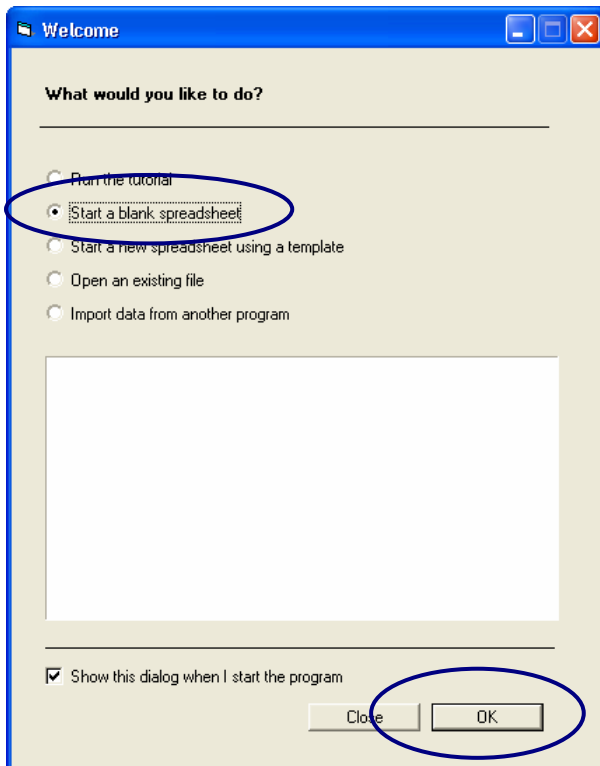
If your trial copy has expired, simply ask for a free extension

- Start the program
- Click “I want to get an unlock code”, to create an e-mail
- In the e-mail, ask for an extension
- We’ll send a code by return e-mail

## Start the program and enter the data

→ Start CMA

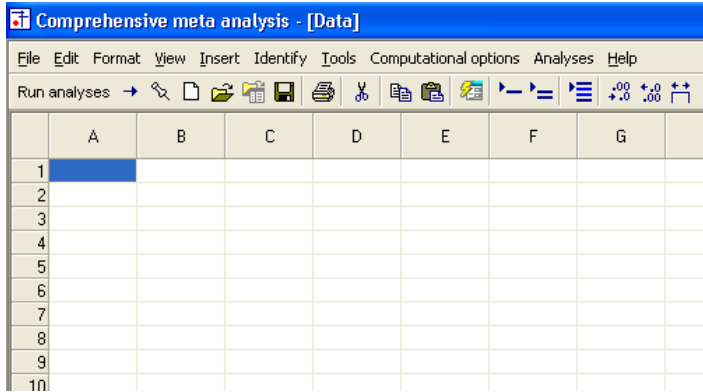
The program shows this dialog box.



→ Select START A BLANK SPREADSHEET

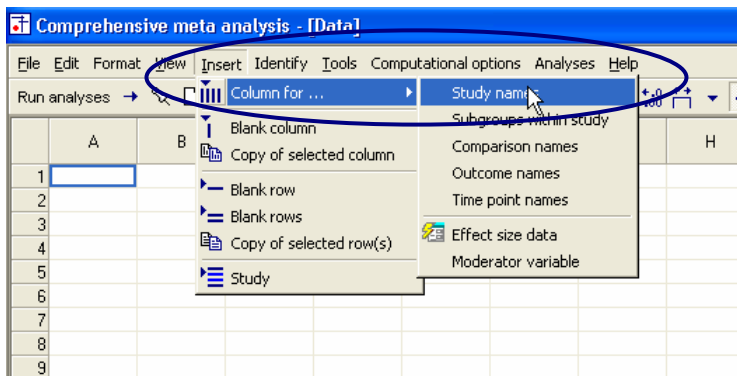
→ Click OK

The program displays this screen.

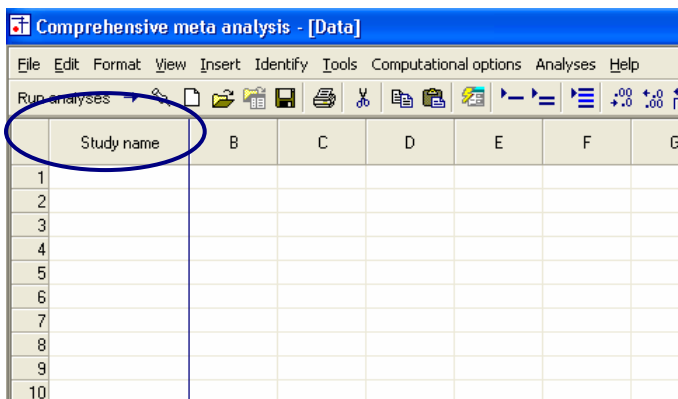


### Insert column for study names

→ Click INSERT > COLUMN FOR > STUDY NAMES



The program has added a column for Study names.

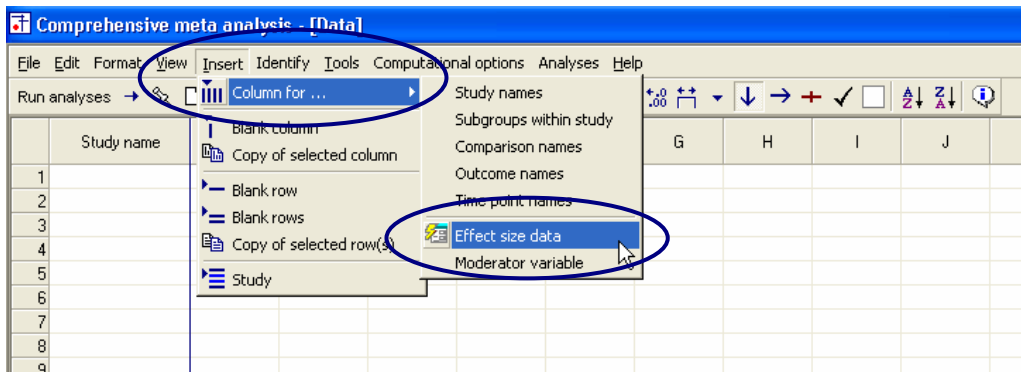


## Insert columns for the effect size data

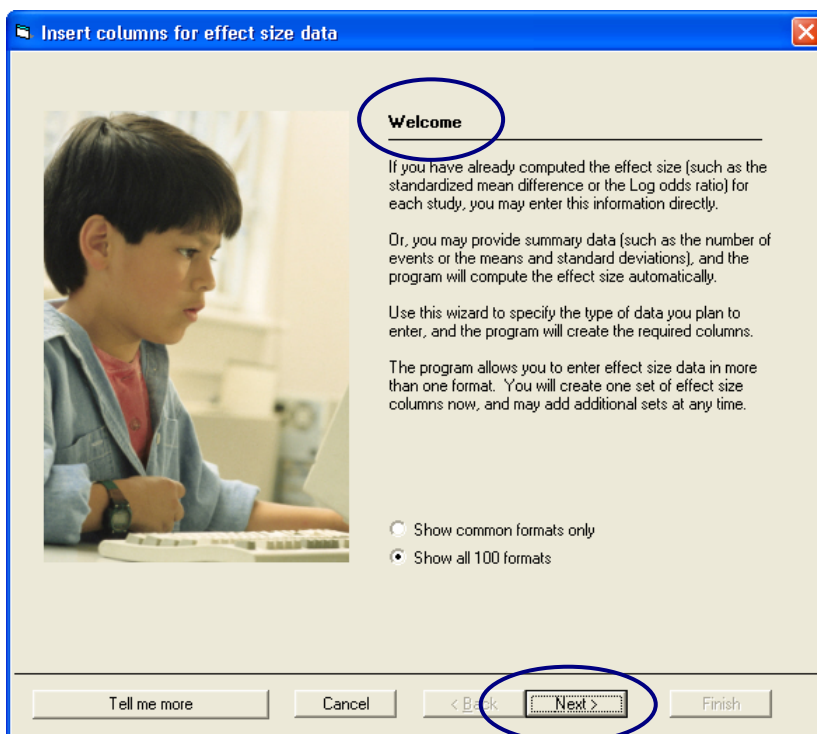
Since CMA will accept data in more than 100 formats, you need to tell the program what format you want to use.

You do have the option to use a different format for each study, but for now we'll start with one format.

→ Click INSERT > COLUMN FOR > EFFECT SIZE DATA



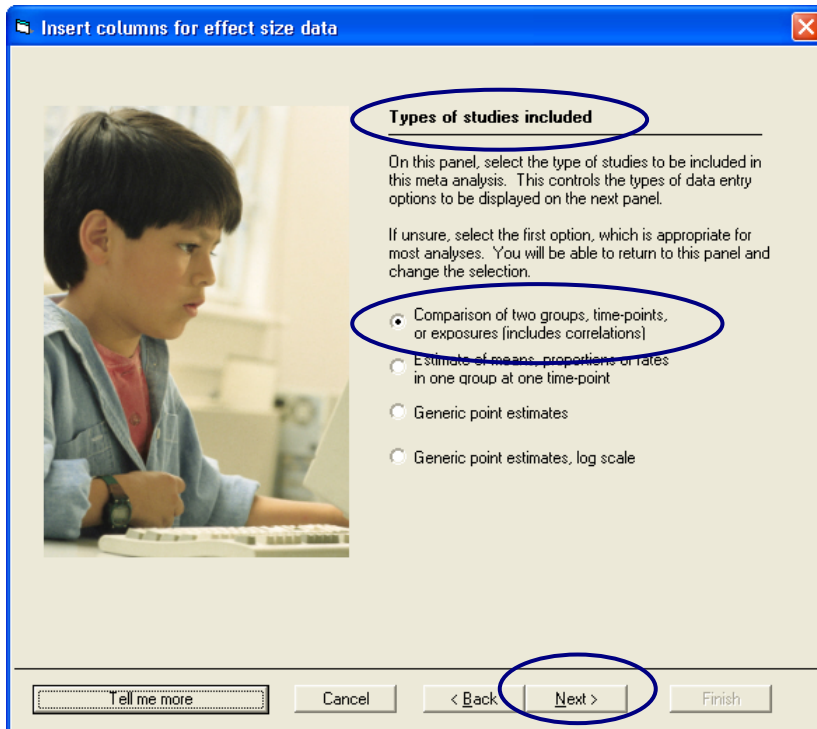
The program shows this dialog box



→ Click NEXT



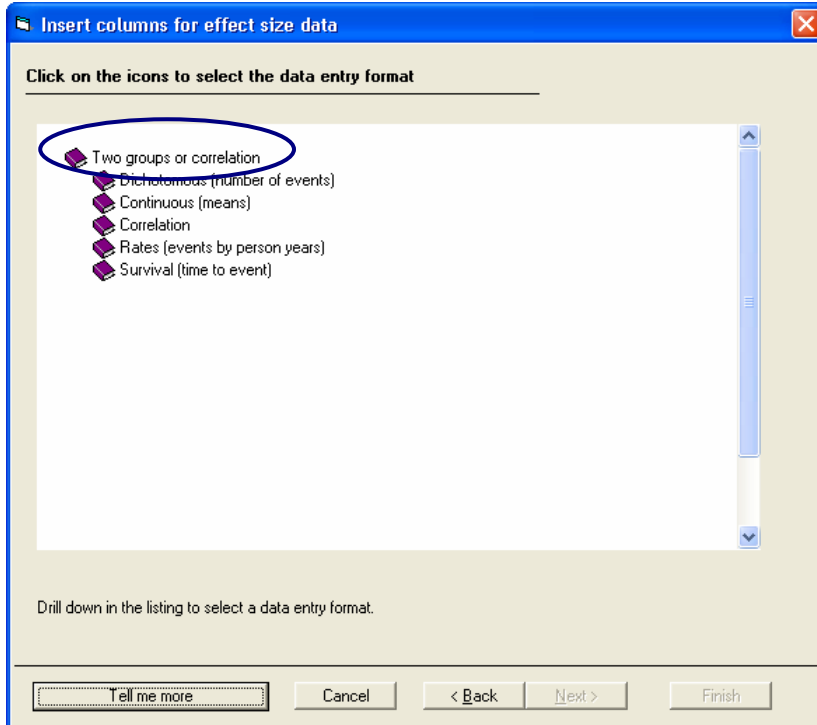
The dialog box lists four sets of effect sizes.



→ Select COMPARISON OF TWO GROUPS, TIME-POINTS, OR EXPOSURES (INCLUDES CORRELATIONS)

→ Click NEXT

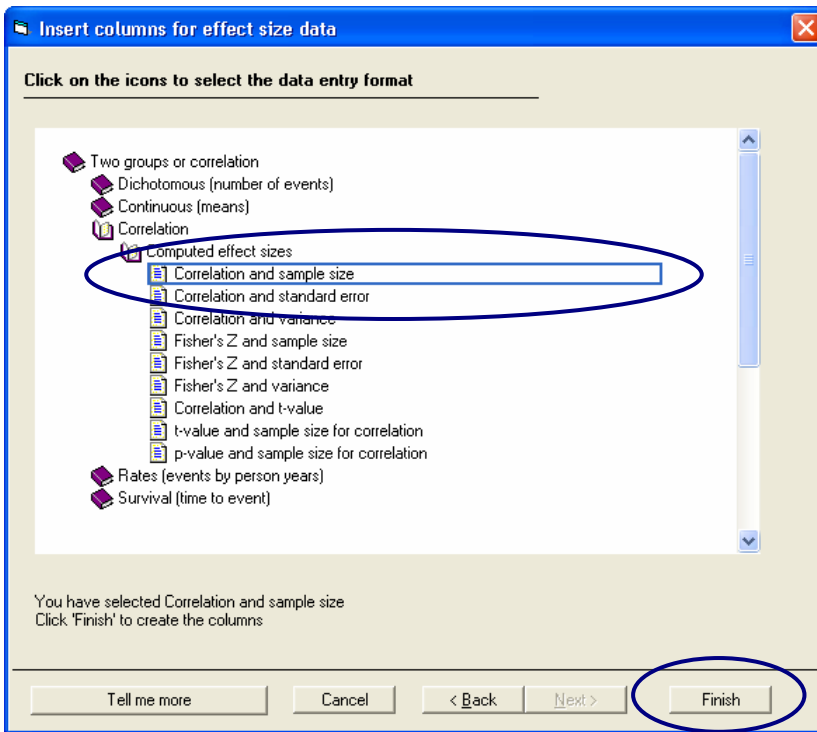
The program displays this dialog box.



→ Drill down to

- CORRELATION
- COMPUTED EFFECT SIZES
- CORRELATION AND SAMPLE SIZE

→ Click FINISH



The program will return to the main data-entry screen.

The program displays the columns needed for the selected format (Correlation, Sample size).

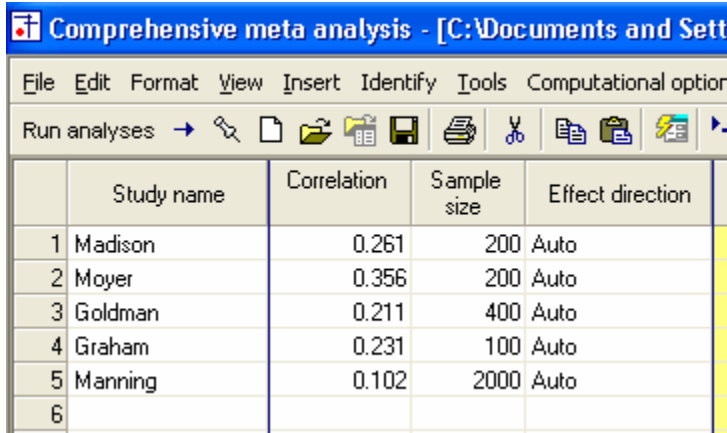
You will enter data into the white columns (at left). The program will compute the effect size for each study and display that effect size in the yellow columns (at right).

Since you elected work with correlational data the program initially displays columns for the correlation and the Fisher's Z transformation of the correlation. You can add other indices as well.

	Study name	Correlation	Sample size	Effect direction	Correlation	Std Err	Variance	Fisher's Z	Std Err	Variance	K	L
1												
2												
3												
4												
5												
6												
7												
8												
9												

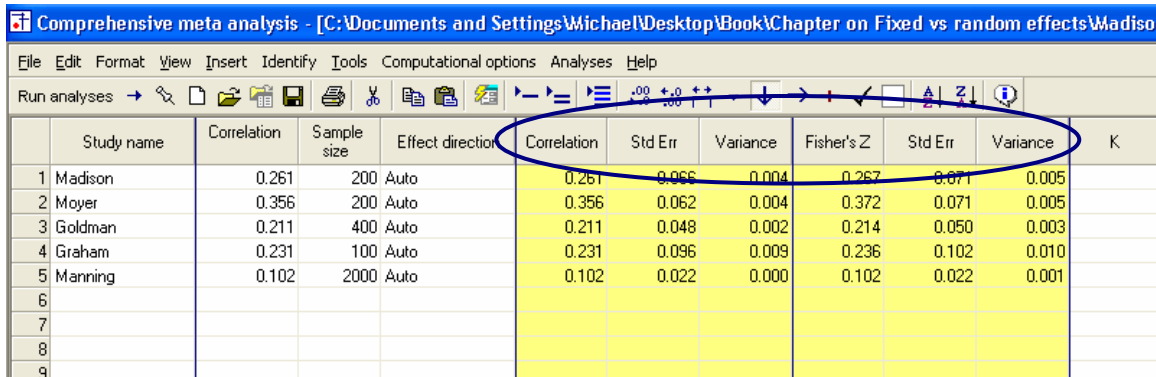
## Enter the data

→ Enter the correlation and sample size for each study as shown here



	Study name	Correlation	Sample size	Effect direction
1	Madison	0.261	200	Auto
2	Moyer	0.356	200	Auto
3	Goldman	0.211	400	Auto
4	Graham	0.231	100	Auto
5	Manning	0.102	2000	Auto
6				

The program will automatically compute the standard error and Fisher's Z values as shown here in the yellow columns.



	Study name	Correlation	Sample size	Effect direction	Correlation	Std Err	Variance	Fisher's Z	Std Err	Variance	K
1	Madison	0.261	200	Auto	0.261	0.066	0.004	0.267	0.071	0.005	
2	Moyer	0.356	200	Auto	0.356	0.062	0.004	0.372	0.071	0.005	
3	Goldman	0.211	400	Auto	0.211	0.048	0.002	0.214	0.050	0.003	
4	Graham	0.231	100	Auto	0.231	0.096	0.009	0.236	0.102	0.010	
5	Manning	0.102	2000	Auto	0.102	0.022	0.000	0.102	0.022	0.001	
6											
7											
8											
9											

## What are Fisher's Z values

The Fisher's Z value is a transformation of the correlation into a different metric.

The standard error of a correlation is a function not only of sample size but also of the correlation itself, with larger correlations (either positive or negative) having a smaller standard error. This can cause problems in a meta-analysis since this would lead the larger correlations to appear more precise and be assigned more weight in the analysis.

To avoid this problem we convert all correlations to the Fisher's Z metric, whose standard error is determined solely by sample size. All computations are performed using Fisher's Z. The results are then converted back to correlations for display.

Fisher's Z should not be confused with the Z statistic used to test hypotheses. The two are not related.

This table shows the correspondence between the correlation value and Fisher's Z for specific correlations. Note that the two are similar for correlations near zero, but diverge substantially as the correlation increases.

	A	B
1	Correlation	Fisher's Z
2	0.0000	0.0000
3	0.1000	0.1003
4	0.2000	0.2027
5	0.3000	0.3095
6	0.4000	0.4236
7	0.5000	0.5493
8	0.6000	0.6931
9	0.7000	0.8673
10	0.8000	1.0986
11	0.9000	1.4722
12	0.9900	2.6467

All conversions are handled automatically by the program. That is, if you enter data using correlations, the program will automatically convert these to Fisher's Z, perform the analysis, and then reconvert the values to correlations for display.

The transformation from correlation to Fisher's Z is given by

$$FisherZ = 0.5 * \text{Log} \left( \frac{1 + \text{Correlation}}{1 - \text{Correlation}} \right)$$

$$SE_{FisherZ} = \frac{1}{\sqrt{N-3}}$$

The transformation from Fisher's Z to correlation is given by

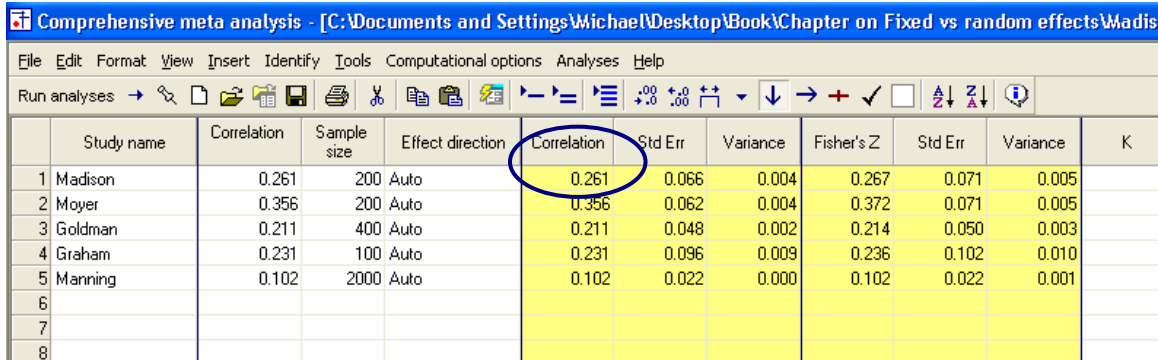
$$C = \text{Exp}(2 * \text{FisherZ})$$

$$\text{Correlation} = \left( \frac{C - 1}{C + 1} \right)$$

$$SE_{\text{Correlation}} = (1 - \text{Correlation}^2) * SE_{FisherZ}$$

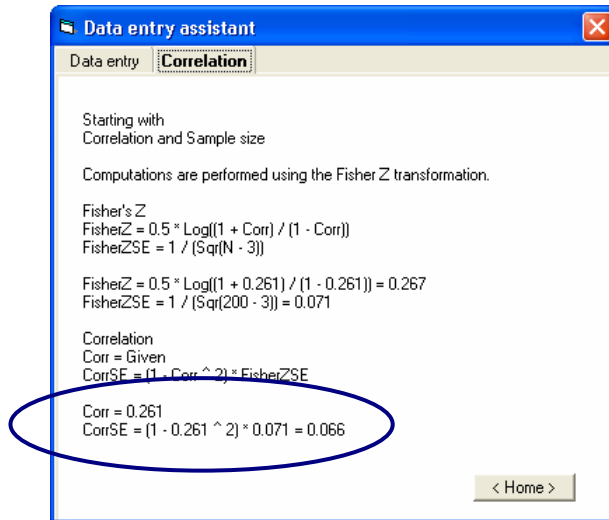
Show details for the computations.

→ Double click on the value 0.261



	Study name	Correlation	Sample size	Effect direction	Correlation	Std Err	Variance	Fisher's Z	Std Err	Variance	K
1	Madison	0.261	200	Auto	0.261	0.066	0.004	0.267	0.071	0.005	
2	Moyer	0.356	200	Auto	0.356	0.062	0.004	0.372	0.071	0.005	
3	Goldman	0.211	400	Auto	0.211	0.048	0.002	0.214	0.050	0.003	
4	Graham	0.231	100	Auto	0.231	0.096	0.009	0.236	0.102	0.010	
5	Manning	0.102	2000	Auto	0.102	0.022	0.000	0.102	0.022	0.001	
6											
7											
8											

The program shows how all related values were computed.



**Data entry assistant**

Data entry **Correlation**

Starting with  
Correlation and Sample size

Computations are performed using the Fisher Z transformation.

Fisher's Z  
 $FisherZ = 0.5 * \text{Log}((1 + \text{Corr}) / (1 - \text{Corr}))$   
 $FisherZSE = 1 / (\text{Sqr}(N - 3))$

$FisherZ = 0.5 * \text{Log}((1 + 0.261) / (1 - 0.261)) = 0.267$   
 $FisherZSE = 1 / (\text{Sqr}(200 - 3)) = 0.071$

Correlation  
Corr = Given  
 $\text{CorrSE} = (1 - \text{Corr}^2) * \text{FisherZSE}$

Corr = 0.261  
 $\text{CorrSE} = (1 - 0.261^2) * 0.071 = 0.066$

< Home >

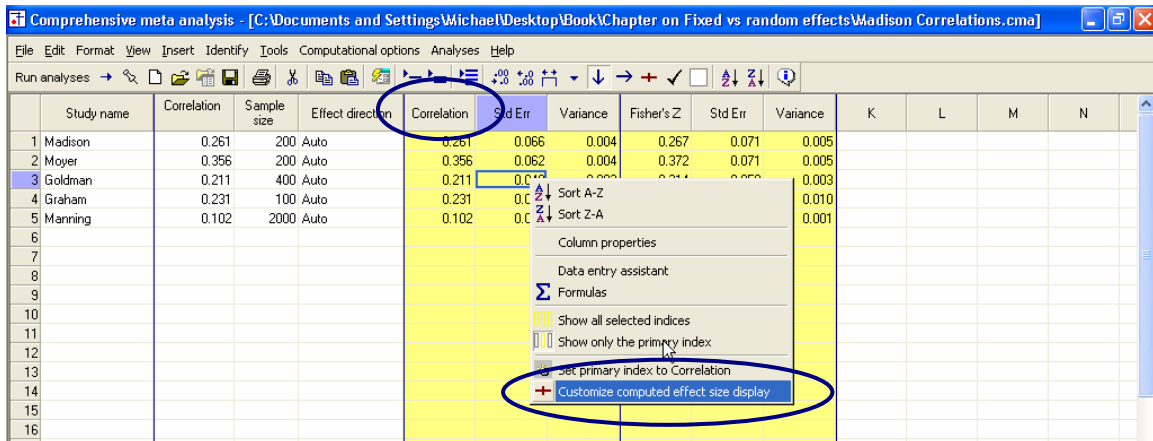


## Set the default index

At this point, the program has displayed the correlation and the Fisher's Z transformation, which we'll be using in this example.

You have the option of adding additional indices, and/or specifying which index should be used as the "Default" index when you run the analysis.

- Right-click on any of the yellow columns
- Select CUSTOMIZE COMPUTED EFFECT SIZE DISPLAY

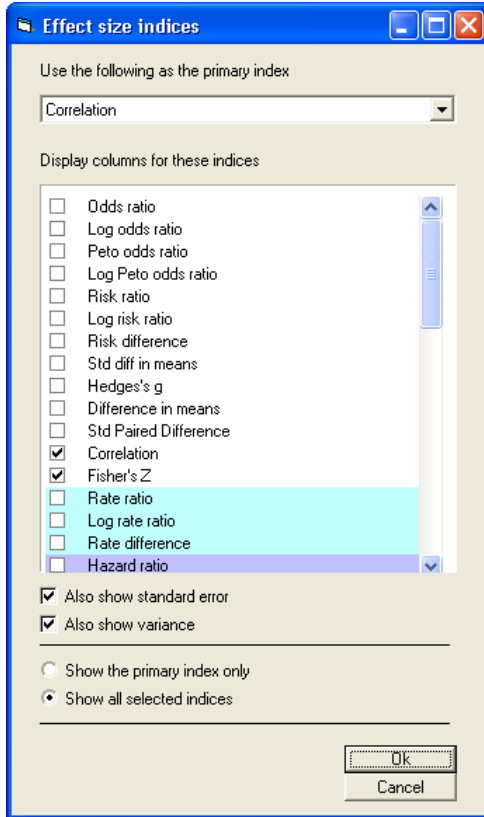


The screenshot shows the 'Comprehensive meta analysis' software interface. The main window displays a table with the following data:

Study name	Correlation	Sample size	Effect direction	Correlation	Std Err	Variance	Fisher's Z	Std Err	Variance	K	L	M	N
1 Madison	0.261	200	Auto	0.261	0.066	0.004	0.267	0.071	0.005				
2 Moyer	0.356	200	Auto	0.356	0.062	0.004	0.372	0.071	0.005				
3 Goldman	0.211	400	Auto	0.211	0.057	0.003	0.217	0.060	0.003				
4 Graham	0.231	100	Auto	0.231	0.077	0.006	0.237	0.087	0.010				
5 Manning	0.102	2000	Auto	0.102	0.025	0.002	0.103	0.025	0.001				

A context menu is open over the 'Std Err' column of the first row. The menu options are: Sort A-Z, Sort Z-A, Column properties, Data entry assistant, Formulas, Show all selected indices, Show only the primary index, Set primary index to Correlation, and **Customize computed effect size display** (highlighted with a blue circle).

The program displays this dialog box.



You could use this dialog box to add or remove effect size indices from the display.

→ Click OK

## Run the analysis

→ Click RUN ANALYSES

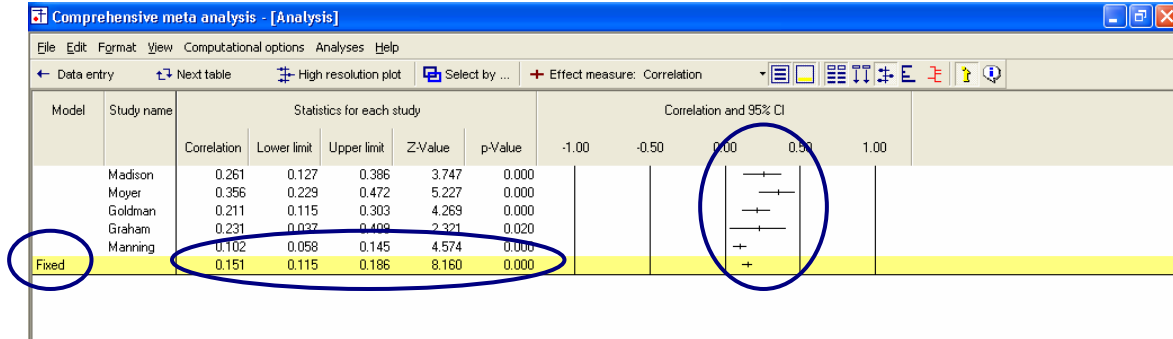
Study name	Correlation	Sample size	Effect direction	Correlation	Std Err	Variance	Fisher's Z	Std Err	Variance	K	L	M	N
1 Madison	0.261	200	Auto	0.261	0.066	0.004	0.267	0.071	0.005				
2 Moyer	0.356	200	Auto	0.356	0.062	0.004	0.372	0.071	0.005				
3 Goldman	0.211	400	Auto	0.211	0.048	0.002	0.214	0.050	0.003				
4 Graham	0.231	100	Auto	0.231	0.096	0.009	0.236	0.102	0.010				
5 Manning	0.102	2000	Auto	0.102	0.022	0.000	0.102	0.022	0.001				
6													
7													
8													
9													
10													

The program displays this screen.

- The default effect size is the correlation
- The default model is fixed effect

Model	Study name	Statistics for each study					Correlation and 95% CI				
		Correlation	Lower limit	Upper limit	Z-Value	p-Value	-1.00	-0.50	0.00	0.50	1.00
	Madison	0.261	0.127	0.386	3.747	0.000					
	Moyer	0.356	0.229	0.472	5.227	0.000					
	Goldman	0.211	0.115	0.303	4.269	0.000					
	Graham	0.231	0.037	0.409	2.321	0.020					
	Manning	0.102	0.058	0.145	4.574	0.000					
Fixed		0.151	0.115	0.186	8.160	0.000					

The screen should look like this.



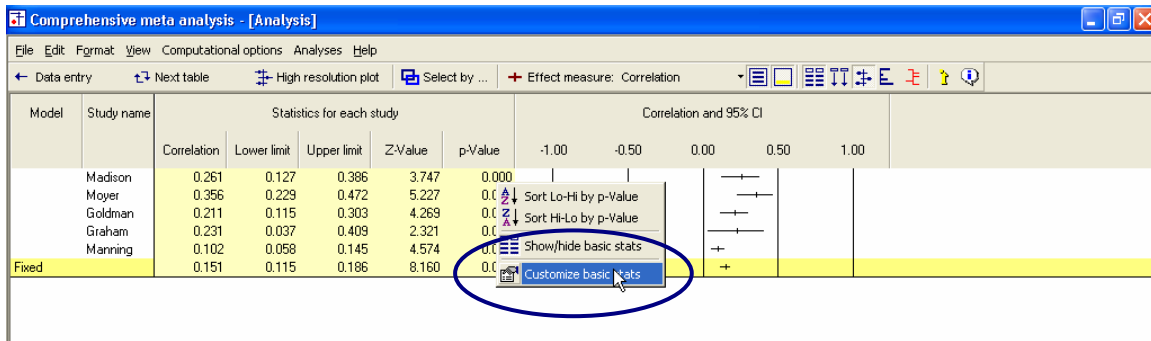
We can immediately get a sense of the studies and the combined effect. For example,

- All the correlations are positive, and they all fall in the range of 0.102 to 0.356
- Some studies are clearly more precise than others. The confidence interval for Graham is substantially wider than the one for Manning, with the other three studies falling somewhere in between
- The combined correlation is 0.151 with a 95% confidence interval of 0.115 to 0.186

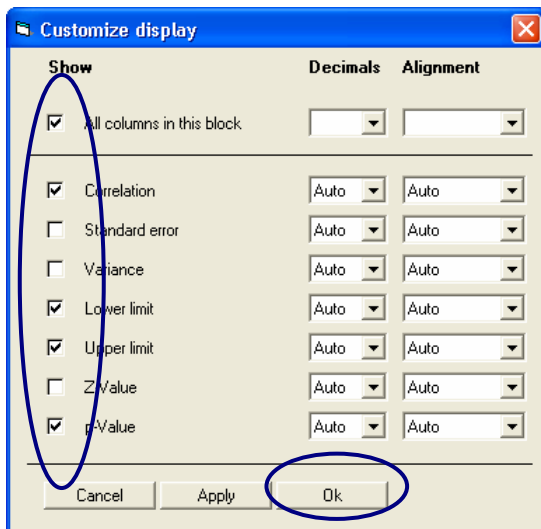
## Customize the screen

We want to hide the column for the z-value.

- Right-click on one of the “Statistics” columns
- Select CUSTOMIZE BASIC STATS

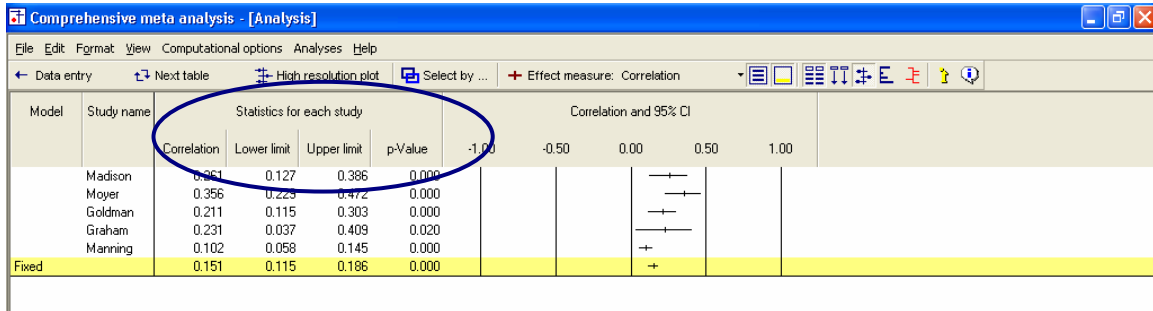


- Assign check-marks as shown here
- Click OK



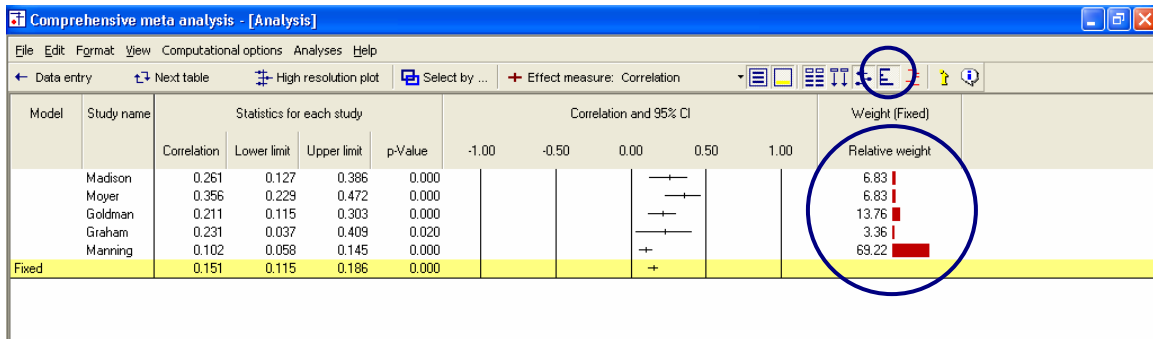
Note – the standard error and variance are never displayed for the correlation. They are displayed when the corresponding boxes are checked and Fisher’s Z is selected as the index.

The program has hidden some of the columns, leaving us more room to work with on the display.



## Display weights

→ Click the tool for SHOW WEIGHTS



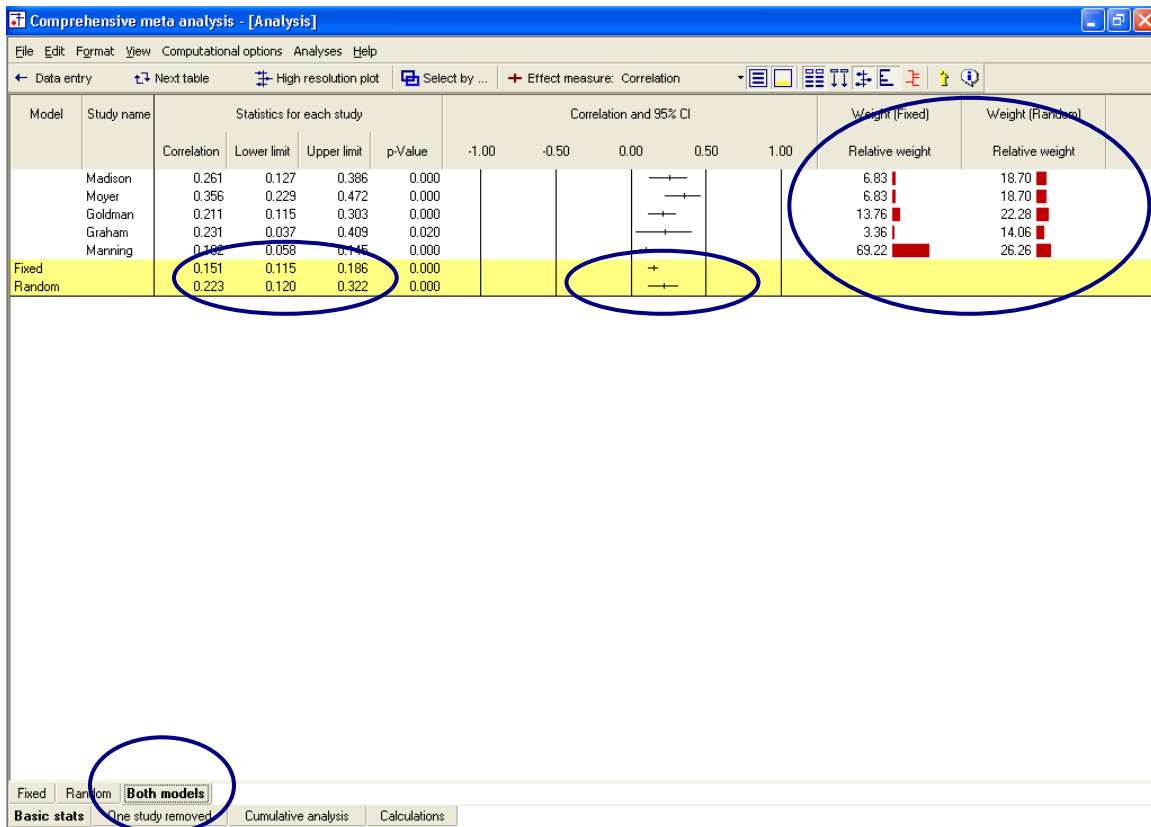
The program now shows the relative weight assigned to each study for the fixed effect analysis. By “relative weight” we mean the weights as a percentage of the total weights, with all relative weights summing to 100%.

For example, Madison was assigned a relative weight of 6.83% while Manning was assigned a relative weight of 69.22%.

## Compare the fixed effect and random effects models

→ At the bottom of the screen, select BOTH MODELS

- The program shows the combined effect and confidence limits for both fixed and random effects models
- The program shows weights for both the fixed effect and the random effects models



## Impact of model on study weights

The Manning study, with a large sample size (N=2000) is assigned 69% of the weight under the fixed effect model, but only 26% of the weight under the random effects model.

This follows from the logic of fixed and random effects models explained earlier.

Under the fixed effect model we assume that all studies are estimating the same value and this study yields a better estimate than the others, so we take advantage of that.

Under the random effects model we assume that each study is estimating a unique effect. The Manning study yields a precise estimate of its population, but that population is only one of many, and we don't want it to dominate the analysis. Therefore, we assign it 26% of the weight. This is more than the other studies, but not the dominant weight that we gave it under fixed effects.

Model	Study name	Statistics for each study				Correlation and 95% CI					Weight (Fixed)		Weight (Random)	
		Correlation	Lower limit	Upper limit	p-Value	-1.00	-0.50	0.00	0.50	1.00	Relative weight	Relative weight	Relative weight	Relative weight
	Madison	0.261	0.127	0.386	0.000						6.83	18.70	6.83	18.70
	Moyer	0.356	0.229	0.472	0.000						6.83	18.70	6.83	18.70
	Goldman	0.211	0.115	0.303	0.000						13.76	37.40	13.76	37.40
	Graham	0.231	0.037	0.409	0.020						3.36	14.06	3.36	14.06
	Manning	0.102	0.058	0.145	0.000						69.22	26.26	69.22	26.26
Fixed		0.151	0.115	0.186	0.000									
Random		0.223	0.120	0.322	0.000									



## Impact of model on the combined effect

As it happens, the Manning study of 0.102 is the smallest effect size in this group of studies. Under the fixed effect model, where this study dominates the weights, it pulls the combined effect down to 0.151. Under the random effects model, it still pulls the effect size down, but only to 0.223.

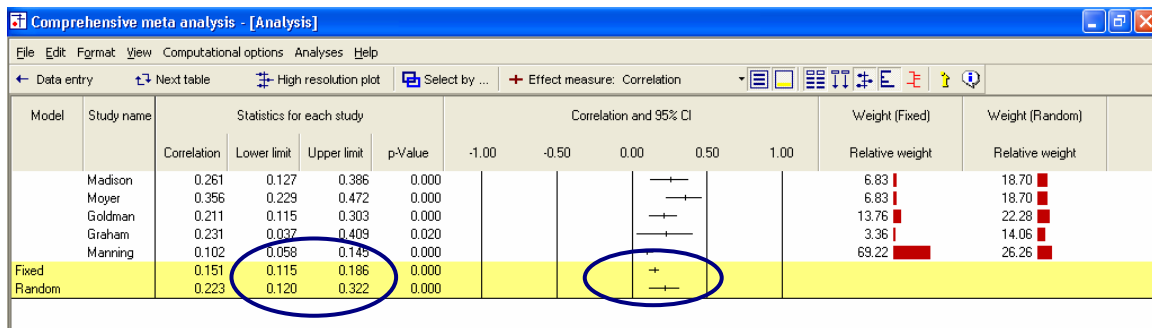
Model	Study name	Statistics for each study				Correlation and 95% CI					Weight (Fixed)		Weight (Random)	
		Correlation	Lower limit	Upper limit	p-Value	-1.00	-0.50	0.00	0.50	1.00	Relative weight	Relative weight	Relative weight	Relative weight
	Madison	0.261	0.127	0.386	0.000			+			6.83		18.70	
	Moyer	0.356	0.229	0.472	0.000			+			6.83		18.70	
	Goldman	0.211	0.115	0.303	0.000			+			13.76		22.28	
	Graham	0.231	0.037	0.409	0.020			+			3.36		14.06	
	Manning	0.102	0.058	0.145	0.000			+			69.22		26.26	
Fixed		0.151	0.115	0.186	0.000			+						
Random		0.223	0.120	0.322	0.000			+						

## Impact of model on the confidence interval

Under fixed effect, we “set” the between-studies dispersion to zero. Therefore, for the purpose of estimating the mean effect, the only source of uncertainty is within-study error. With a combined total near 3000 subjects the within-study error is small, so we have a precise estimate of the combined effect. The confidence interval is relatively narrow, extending from 0.12 to 0.19.

Under random effects, dispersion between studies is considered a real source of uncertainty. And, there is a lot of it. The fact that these five studies vary so much one from the other tells us that the effect will vary depending on details that vary randomly from study to study. If the persons who performed these studies happened to use older subjects, or a shorter duration, for example, the effect size would have changed.

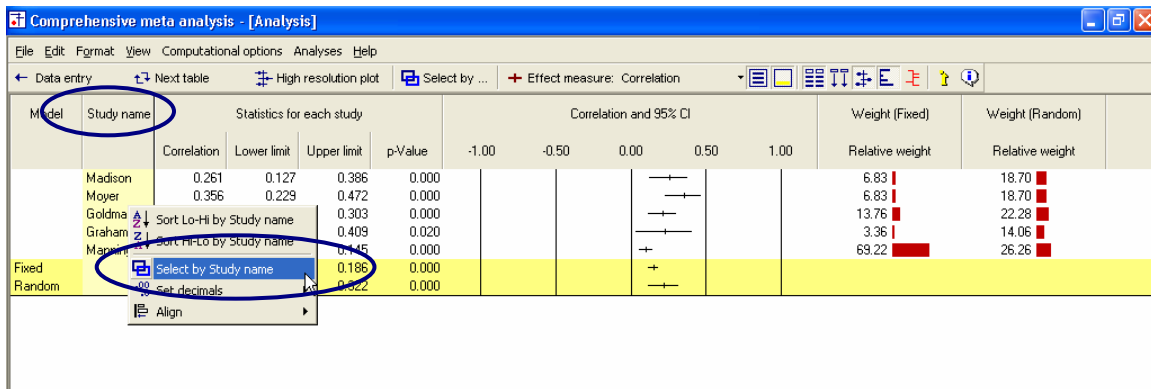
While this dispersion is “real” in the sense that it is caused by real differences among the studies, it nevertheless represents error if our goal is to estimate the mean effect. For computational purposes, the variance due to between-study differences is included in the error term. In our example we have only five studies, and the effect sizes do vary. Therefore, our estimate of the mean effect is not terribly precise, as reflected in the width of the confidence interval, 0.12 to 0.32, substantially wider than that for the fixed effect model.



## What would happen if we eliminated Manning?

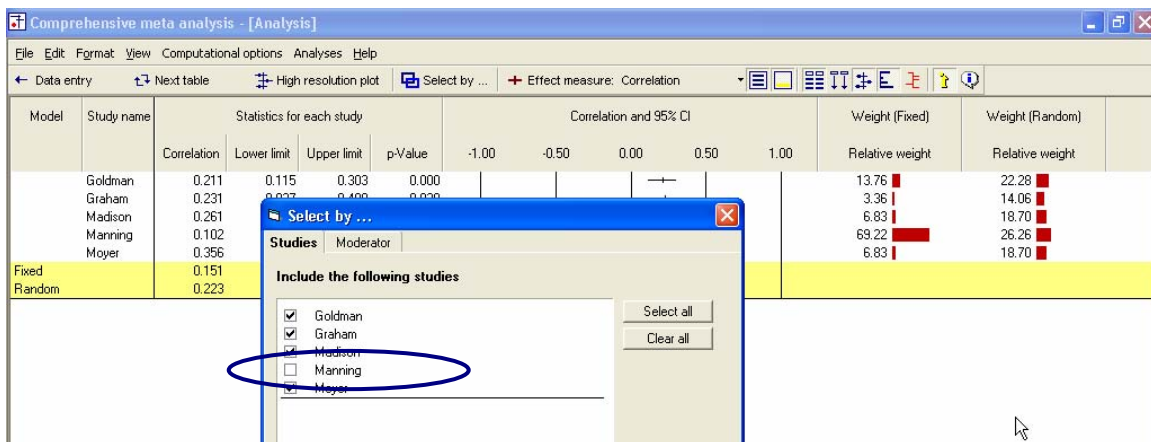
Manning was the largest study, and also the study with the most powerful (left-most) effect size. To better understand the impact of this study under the two models, let's see what would happen if we were to remove this study from the analysis.

- Right-click on STUDY NAME
- Select SELECT BY STUDY NAME

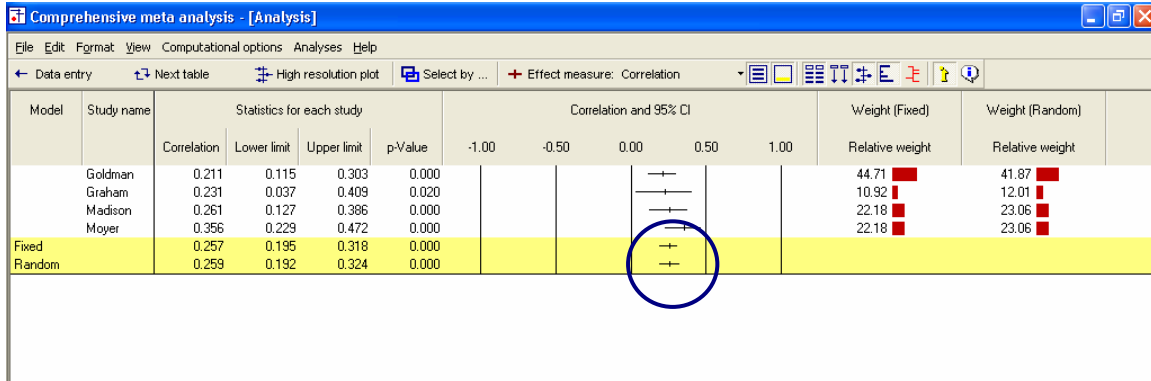


The program opens a dialog box with the names of all studies.

- Remove the check from Manning
- Click OK



The analysis now looks like this.

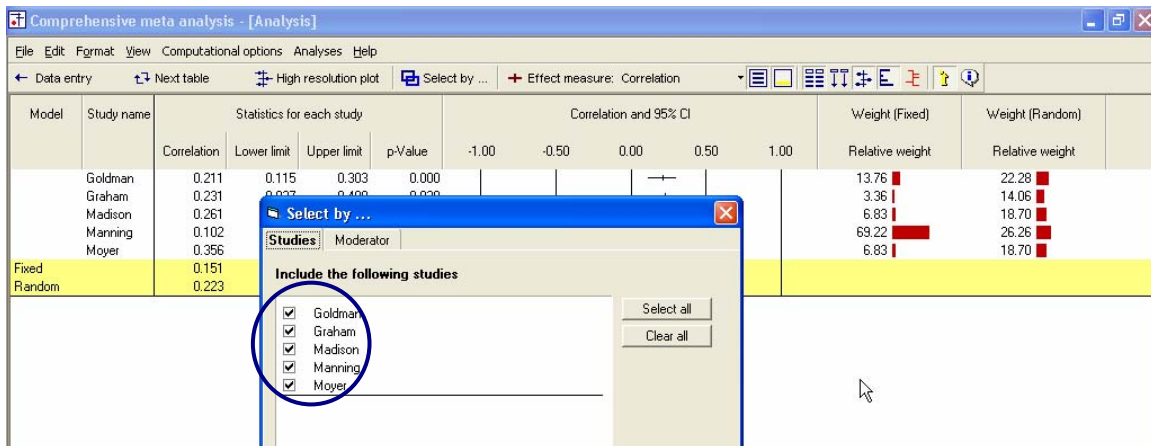


For both the fixed effect and random effects models, the combined effect is now approximately 0.26.

- Under fixed effects Manning had pulled the effect down to 0.15
- Under random effects Manning had pulled the effect down to 0.22
- Thus, this study had a substantial impact under either model, but more so under fixed than random effects

→ Right-click on STUDY NAMES

→ Add a check for Manning so the analysis again has five studies



## Additional statistics

→ Click NEXT TABLE on the toolbar

Model	Study name	Statistics for each study				Correlation and 95% CI					Weight (Fixed)		Weight (Random)	
		Correlation	Lower limit	Upper limit	p-Value	-1.00	-0.50	0.00	0.50	1.00	Relative weight	Relative weight	Relative weight	Relative weight
	Goldman	0.211	0.115	0.303	0.000			+			13.76		22.28	
	Graham	0.231	0.037	0.409	0.020			+			3.36		14.06	
	Madison	0.261	0.127	0.396	0.000			+			6.83		18.70	
	Manning	0.102	0.058	0.145	0.000			+			69.22		26.26	
	Moyer	0.356	0.229	0.472	0.000			+			6.83		18.70	
Fixed		0.151	0.115	0.186	0.000			+						
Random		0.223	0.120	0.322	0.000			+						

The program switches to this screen.

The program shows the point estimate and confidence interval. These are the same values that had been shown on the forest plot.

Model	Effect size and 95% interval				Test of null (2-Tail)		Heterogeneity				Tau-squared	
	Number Studies	Point estimate	Lower limit	Upper limit	Z-value	P-value	Q-value	df (Q)	P-value	I-squared	Tau Squared	Standard Error
Fixed	5	0.151	0.115	0.186	8.160	0.000	19.318	4	0.001	79.294	0.011	0.0
Random	5	0.223	0.120	0.322	4.172	0.000						

- Under fixed effect the combined effect is 0.151 with 95% confidence interval of 0.115 to 0.186
- Under random effects the combined effect is 0.223 with 95% confidence interval of 0.120 to 0.322

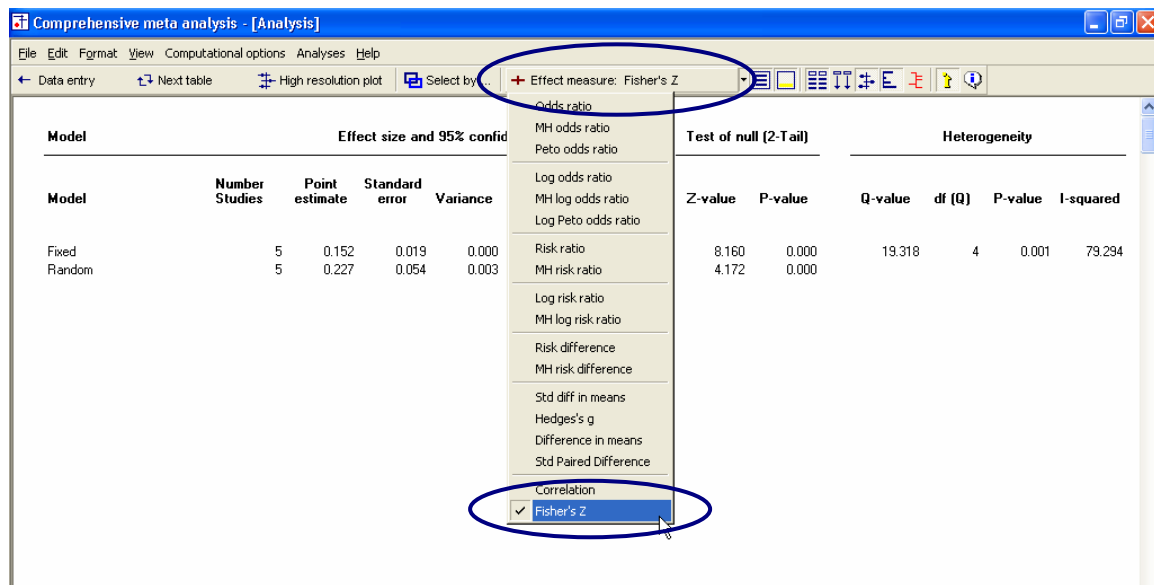
## Test of the null hypothesis

Under the fixed effect model the null hypothesis is that the common effect is zero. Under the random effects model the null hypothesis is that the mean of the true effects is zero.

In either case, the null hypothesis is tested by the z-value, which is computed as Fisher's  $Z/SE$  for the corresponding model.

To this point we've been displaying the correlation rather than the Fisher's  $Z$  value. The test statistic  $Z$  (not to be confused with Fisher's  $Z$ ) is correct as displayed (since it is always based on the Fisher's  $Z$  transform), but to understand the computation we need to switch the display to show Fisher's  $Z$  values.

Select FISHER'S  $Z$  from the drop down box



The screenshot shows the 'Comprehensive meta analysis - [Analysis]' window. The 'Effect measure' dropdown menu is open, showing a list of options. 'Fisher's Z' is selected and highlighted with a blue checkmark. The main window displays a table with columns for 'Model', 'Number Studies', 'Point estimate', 'Standard error', 'Variance', 'Test of null (2-Tail)', and 'Heterogeneity'. The table shows results for 'Fixed' and 'Random' models.

Model	Number Studies	Point estimate	Standard error	Variance	Test of null (2-Tail)	Heterogeneity
Fixed	5	0.152	0.019	0.000	Z-value: 8.160, P-value: 0.000	Q-value: 19.318, df (Q): 4, P-value: 0.001, I-squared: 79.294
Random	5	0.227	0.054	0.003	Z-value: 4.172, P-value: 0.000	

The screen should look like this (use the Format menu to display additional decimal places).

Model	Effect size and 95% confidence interval					Test of null (2-Tail)		Heterogeneity				
	Number Studies	Point estimate	Standard error	Variance	Lower limit	Upper limit	Z-value	P-value	Q-value	df (Q)	P-value	I-squared
Fixed	5	0.1519	0.0186	0.0003	0.1154	0.1884	8.1600	0.0000	19.3179	4.0000	0.0007	79.2938
Random	5	0.2273	0.0545	0.0030	0.1205	0.3341	4.1720	0.0000				

Note that all values are now in Fisher's Z units.

- The point estimate for the fixed effect and random effects models are now 0.152 and 0.227, which are the Fisher's Z transformations of the correlations, 0.151 and 0.223. (The difference between the correlation and the Fisher's Z value becomes more pronounced as the size of the correlation increases)
- The program now displays the standard error and variance, which can be displayed for the Fisher's Z value but not for the correlation

For the fixed effect analysis

$$Z = \frac{0.1519}{0.0186} = 8.1600$$

For the random effect analysis

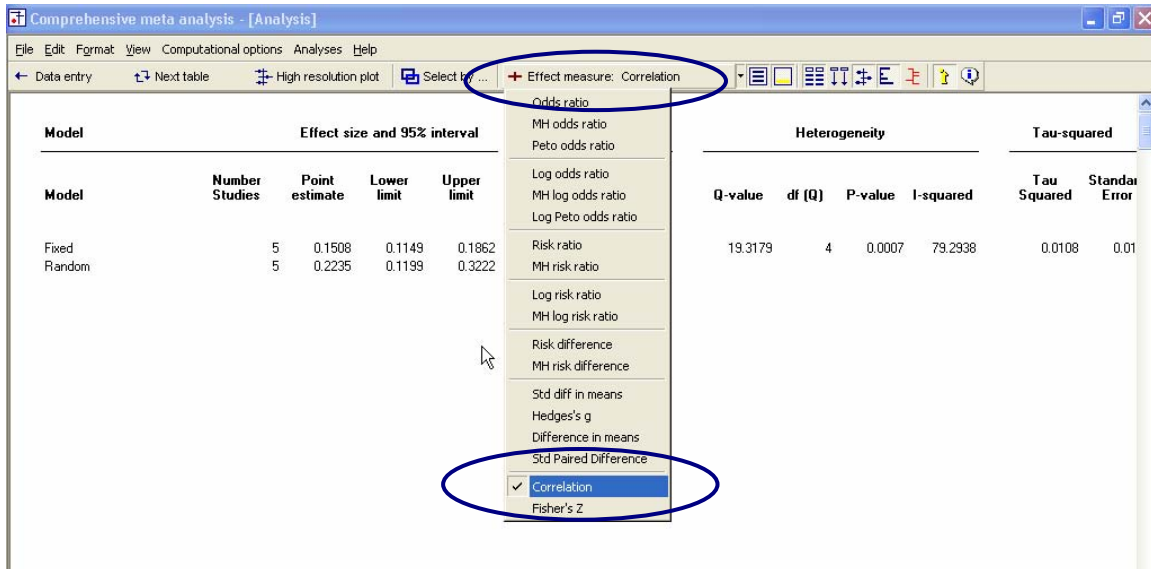
$$Z = \frac{0.2273}{0.0545} = 4.1720$$

In either case, the two-tailed p-value is < 0.0001.

## Test of the heterogeneity

Switch the display back to correlation

→ Select CORRELATION from the drop-down box



Note, however, that the statistics addressed in this section are always computed using the Fisher's Z values, regardless of whether Correlation or Fisher's Z has been selected as the index.

The null hypothesis for heterogeneity is that the studies share a common effect size.

The statistics in this section address the question of whether or the observed dispersion among effects exceeds the amount that would be expected by chance.



Model	Effect size and 95% interval				Test of null (2-Tail)		Heterogeneity				Tau-squared	
	Number Studies	Point estimate	Lower limit	Upper limit	Z-value	P-value	Q-value	df (Q)	P-value	I-squared	Tau Squared	Standard Error
Fixed	5	0.1508	0.1149	0.1862	8.1600	0.0000	19.3179	4	0.0007	79.2938	0.0108	0.01
Random	5	0.2235	0.1199	0.3222	4.1720	0.0000						

The Q statistic reflects the observed dispersion. Under the null hypothesis that all studies share a common effect size, the expected value of Q is equal to the degrees of freedom (the number of studies minus 1), and is distributed as Chi-square with  $df = k-1$  (where k is the number of studies).

- The Q statistic is 19.3179, as compared with an expected value of 4
- The p-value is 0.0007

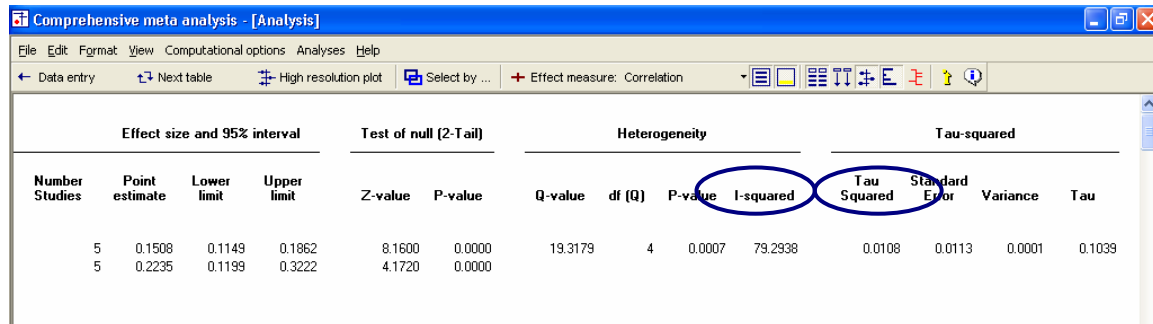
This p-value meets the criterion for statistical significance. It seems clear that there is substantial dispersion, and probably more than we would expect based on random differences. There probably is real variance among the effects.

As discussed in the text, the decision to use a random effects model should be based on our understanding of how the studies were acquired, and should not depend on a statistically significant p-value for heterogeneity. In any event, this p-value does suggest that a fixed effect model does not fit the data.

## Quantifying the heterogeneity

While Q is meant to test the null hypothesis that there is no dispersion across effect sizes, we want also to quantify this dispersion. For this purpose we would turn to I-squared and Tau-squared.

→ To see these statistics, scroll the screen toward the right



Number Studies	Effect size and 95% interval			Test of null (2-Tail)		Heterogeneity			Tau-squared				
	Point estimate	Lower limit	Upper limit	Z-value	P-value	Q-value	df (Q)	P-value	I-squared	Tau Squared	Standard Error	Variance	Tau
5	0.1508	0.1149	0.1862	8.1600	0.0000	19.3179	4	0.0007	79.2938	0.0108	0.0113	0.0001	0.1039
5	0.2235	0.1199	0.3222	4.1720	0.0000								

- I-squared is 79.29, which means that 79% of the observed variance between studies is due to real differences in the effect size. Only about 21% of the observed variance would have been expected based on random error.
- Tau-squared is 0.0108. This is the “Between studies” variance that was used in computing weights.

The Q statistic and tau-squared are reported on the fixed effect line, and not on the random effects line.

These value are displayed on the fixed effect line because they are computed using fixed effect weights. These values are used in both fixed and random effects analyses, but for different purposes.

For the fixed effect analysis Q addresses the question of whether or not the fixed effect model fits the data (is it reasonable to assume that tau-squared is actually zero). However, tau-squared is actually set to zero for the purpose of assigning weights.

For the random effects analysis, these values are actually used to assign the weights.

Return to the main analysis screen

Click NEXT TABLE again to get back to the other screen.

Model	Effect size and 95% interval				Test of null (2-Tail)		Heterogeneity			Tau-squared		
	Number Studies	Point estimate	Lower limit	Upper limit	Z-value	P-value	Q-value	df (Q)	P-value	I-squared	Tau Squared	Standard Error
Fixed	5	0.1508	0.1149	0.1862	8.1600	0.0000	19.3179	4	0.0007	79.2938	0.0108	0.01
Random	5	0.2235	0.1199	0.3222	4.1720	0.0000						

Your screen should look like this

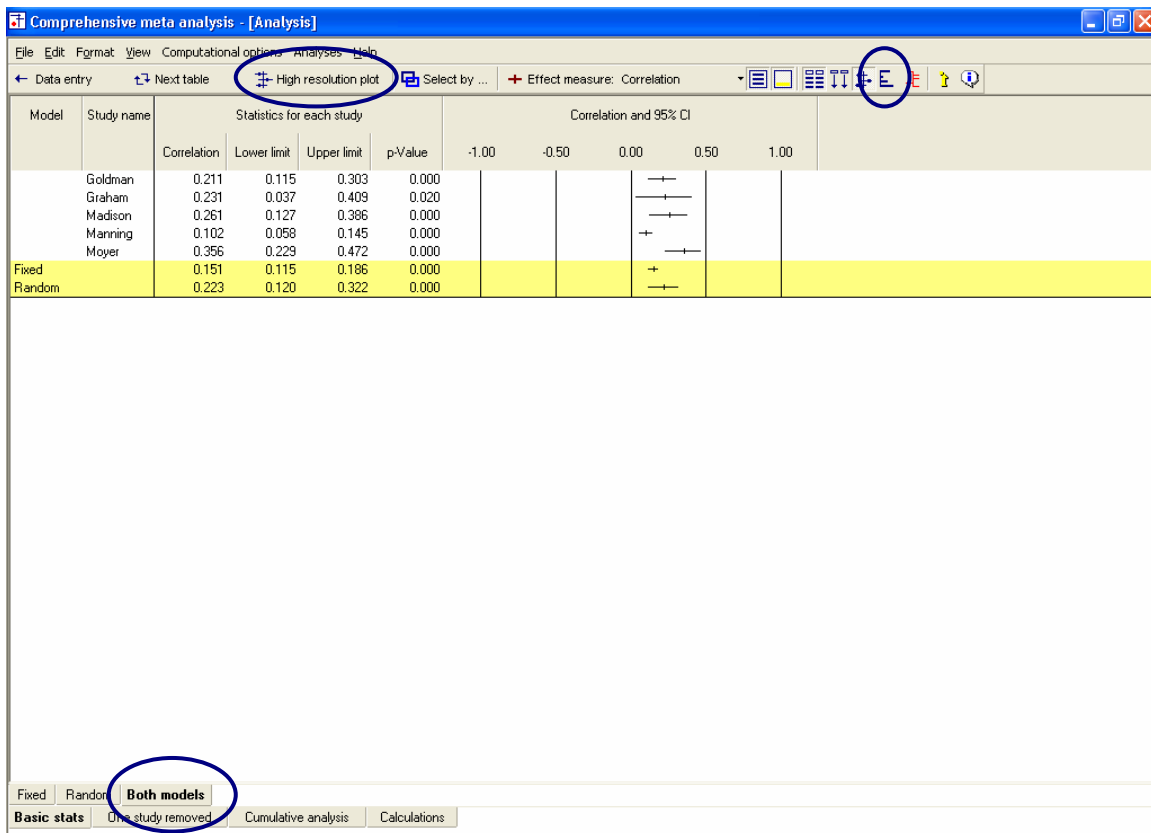
Model	Study name	Statistics for each study				Correlation and 95% CI					Weight (Fixed)		Weight (Random)	
		Correlation	Lower limit	Upper limit	p-Value	-1.00	-0.50	0.00	0.50	1.00	Relative weight	Relative weight	Relative weight	Relative weight
	Goldman	0.211	0.115	0.303	0.000			+			13.76		22.28	
	Graham	0.231	0.037	0.409	0.020			+			3.36		14.06	
	Madison	0.261	0.127	0.386	0.000			+			6.83		18.70	
	Manning	0.102	0.058	0.145	0.000			+			69.22		26.26	
	Moyer	0.356	0.229	0.472	0.000			+			6.83		18.70	
Fixed		0.151	0.115	0.186	0.000			+						
Random		0.223	0.120	0.322	0.000			+						

## High-resolution plots

To this point we've used bar graphs to show the weight assigned to each study.

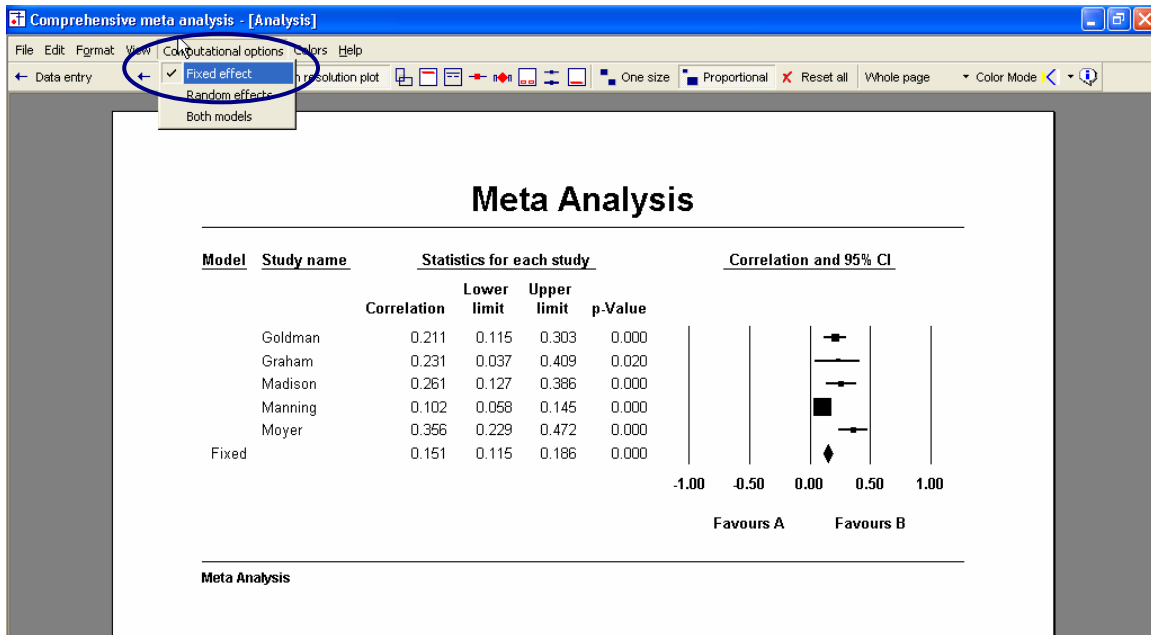
Now, we'll switch to a high-resolution plot, where the weight assigned to each study will be incorporated into the symbol representing that study.

- Select BOTH MODELS at the bottom of the screen
- Unclick the SHOW WEIGHTS button on the toolbar
- Click HIGH-RESOLUTION PLOT on the toolbar

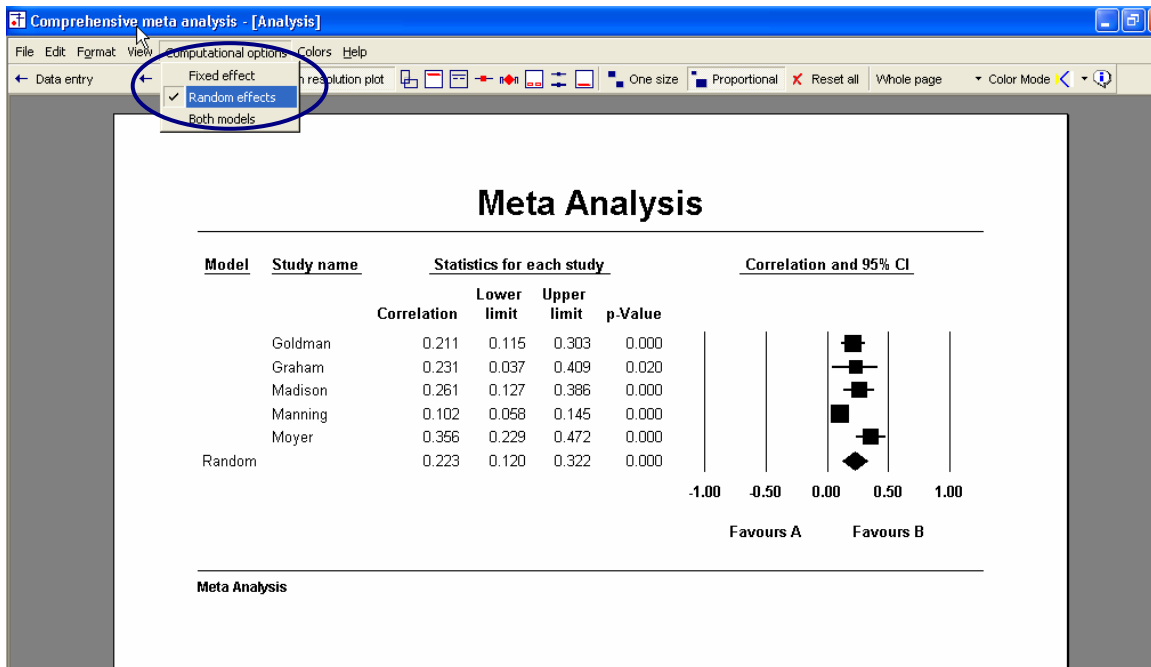


The program shows this screen.

→ Select COMPUTATIONAL OPTIONS > FIXED EFFECT



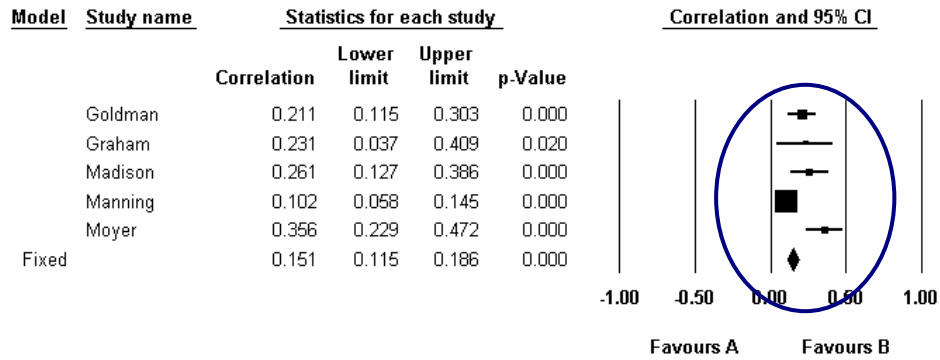
→ Select COMPUTATIONAL OPTIONS > RANDOM EFFECTS



Compare the weights.

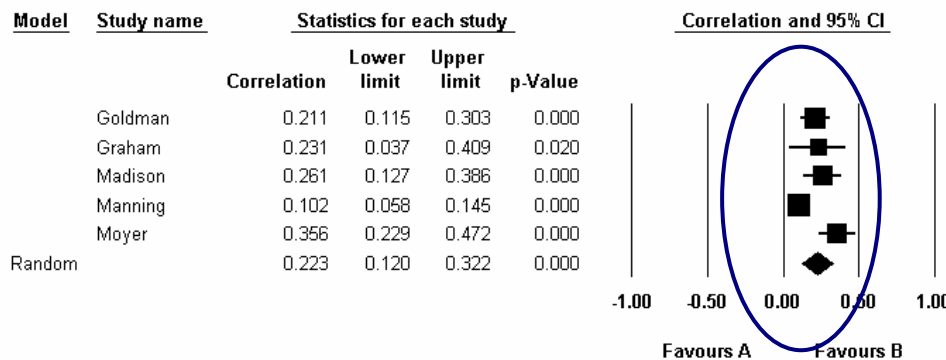
In the first plot, using fixed effect weights, the area of the Manning box was about 20 times that of Graham. In the second, using random effects weights, the area of the Manning box was only about twice as large as Graham.

## Meta Analysis



Meta Analysis

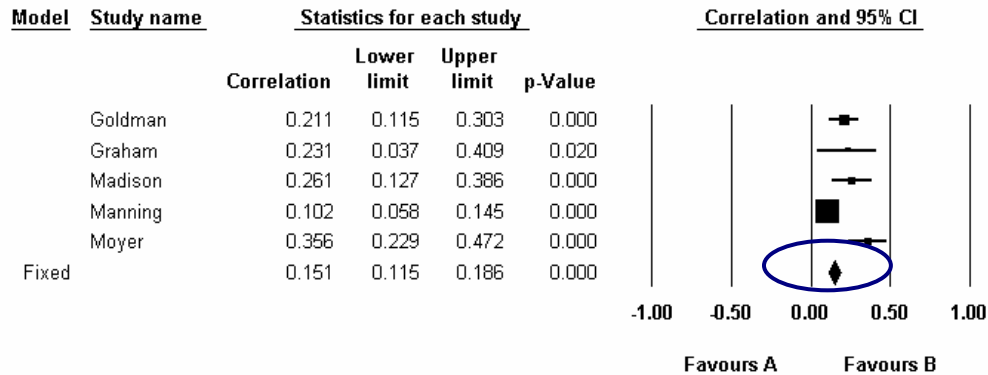
## Meta Analysis



Meta Analysis

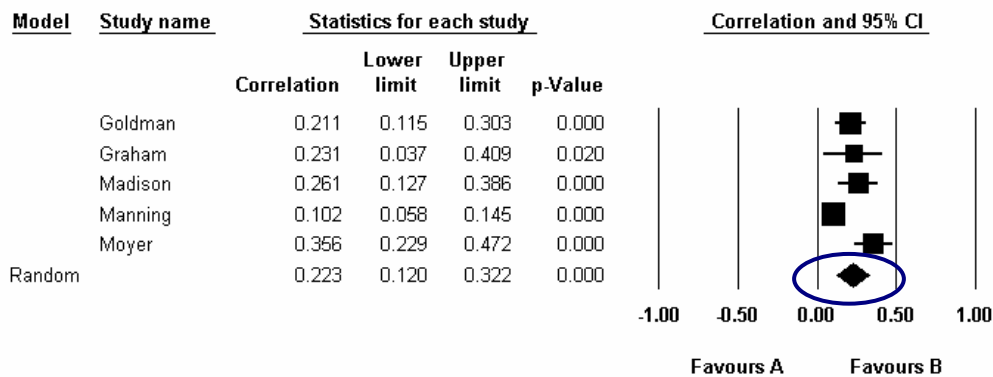
Compare the combined effects and standard error.

## Meta Analysis



Meta Analysis

## Meta Analysis



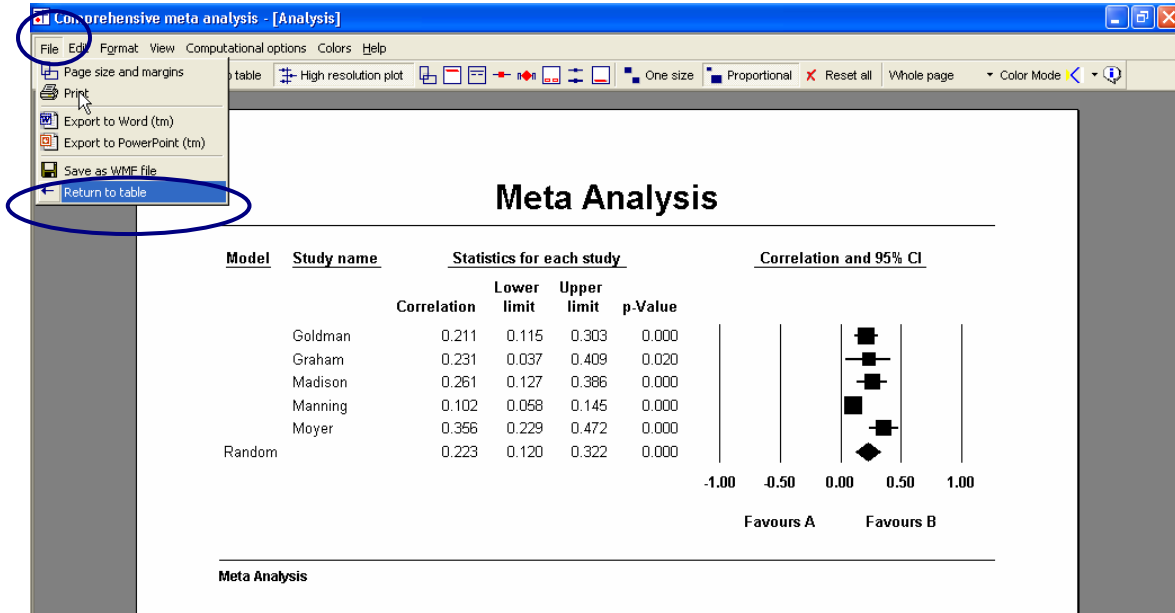
Meta Analysis

In the first case (fixed) Manning is given substantial weight and pulls the combined effect down to .151. In the second case (random) Manning is given less weight, and the combined effect is .223.

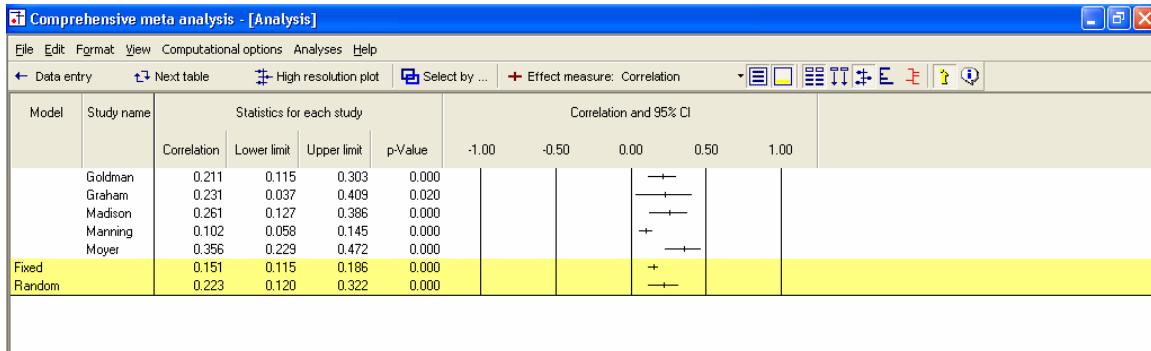
In the first case (fixed) the only source of error is the error within studies and the confidence interval about the combined effect is relatively narrow. In the second case (random) the fact that the true effect varies from study to study introduces

another level of uncertainty to our estimate. The confidence interval about the combined effect is substantially wider than it is for the fixed effect analysis.

→ Click FILE > RETURN TO TABLE



The program returns to the main analysis screen.



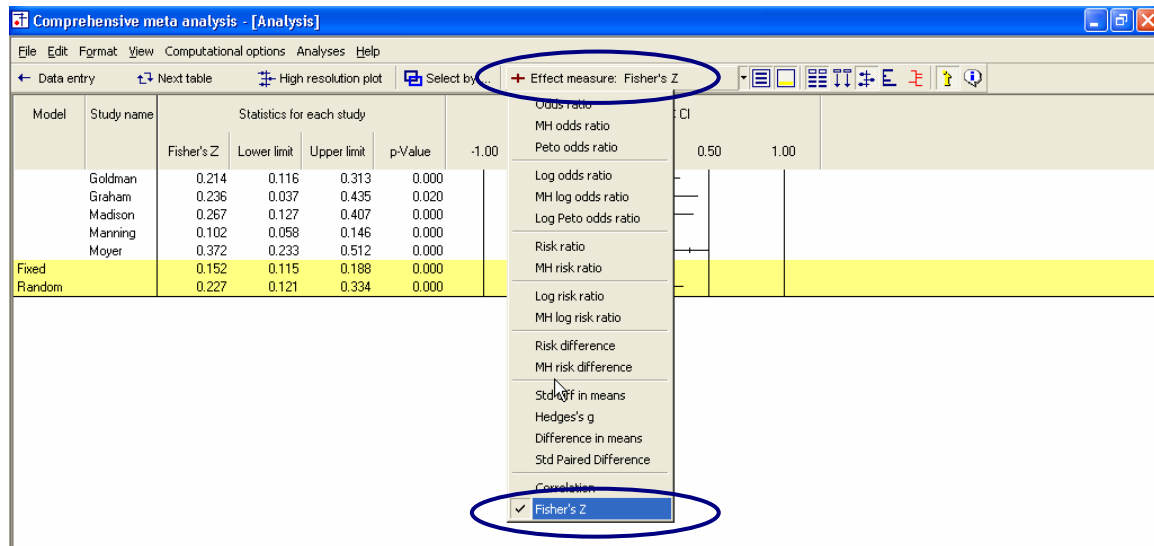


## Computational details

The program allows you to view details of the computations

Since all calculations are performed using Fisher's Z values, they are easier to follow if we switch the screen to use Fisher's Z as the effect size index.

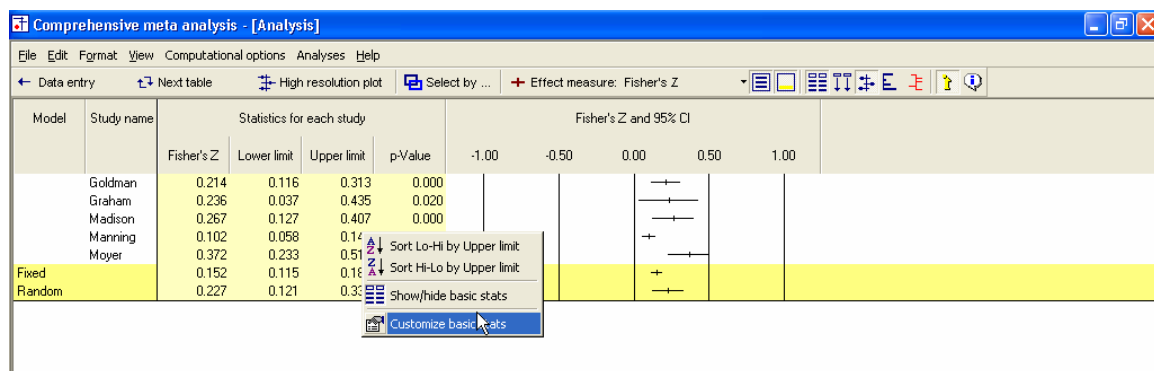
→ Select FISHER'S Z from the drop-down box

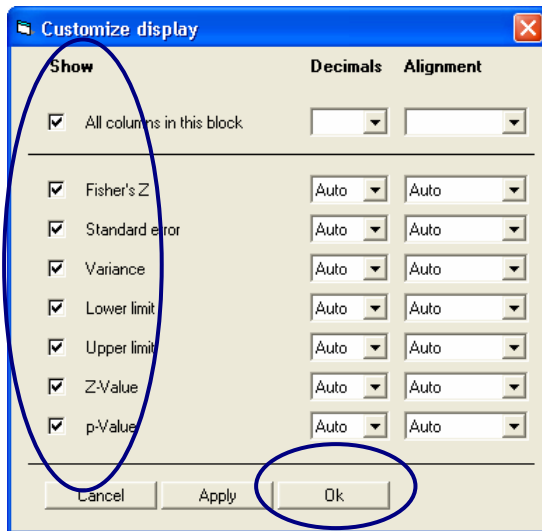


- The program is now showing the effect for each study and the combined effect using values in the Fisher's Z metric.

We want to add columns to display the standard error and variance.

- Right-click on any of the columns in the STATISTICS FOR EACH STUDY section
- Select CUSTOMIZE BASIC STATS





- Check the box next to each statistic
- Click OK

The screen should look like this.

Model	Study name	Statistics for each study							Fisher's Z and 95% CI				
		Fisher's Z	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value	-1.00	-0.50	0.00	0.50	1.00
	Goldman	0.214	0.050	0.003	0.116	0.313	4.269	0.000					
	Graham	0.236	0.102	0.010	0.037	0.435	2.321	0.020					
	Madison	0.267	0.071	0.005	0.127	0.407	3.747	0.000					
	Manning	0.102	0.022	0.001	0.058	0.146	4.574	0.000					
	Moyer	0.372	0.071	0.005	0.233	0.512	5.227	0.000					
Fixed		0.152	0.019	0.000	0.115	0.188	8.160	0.000					
Random		0.227	0.054	0.003	0.121	0.334	4.172	0.000					

- Note that we now have columns for the standard error and variance, and all values are in Fisher's Z units

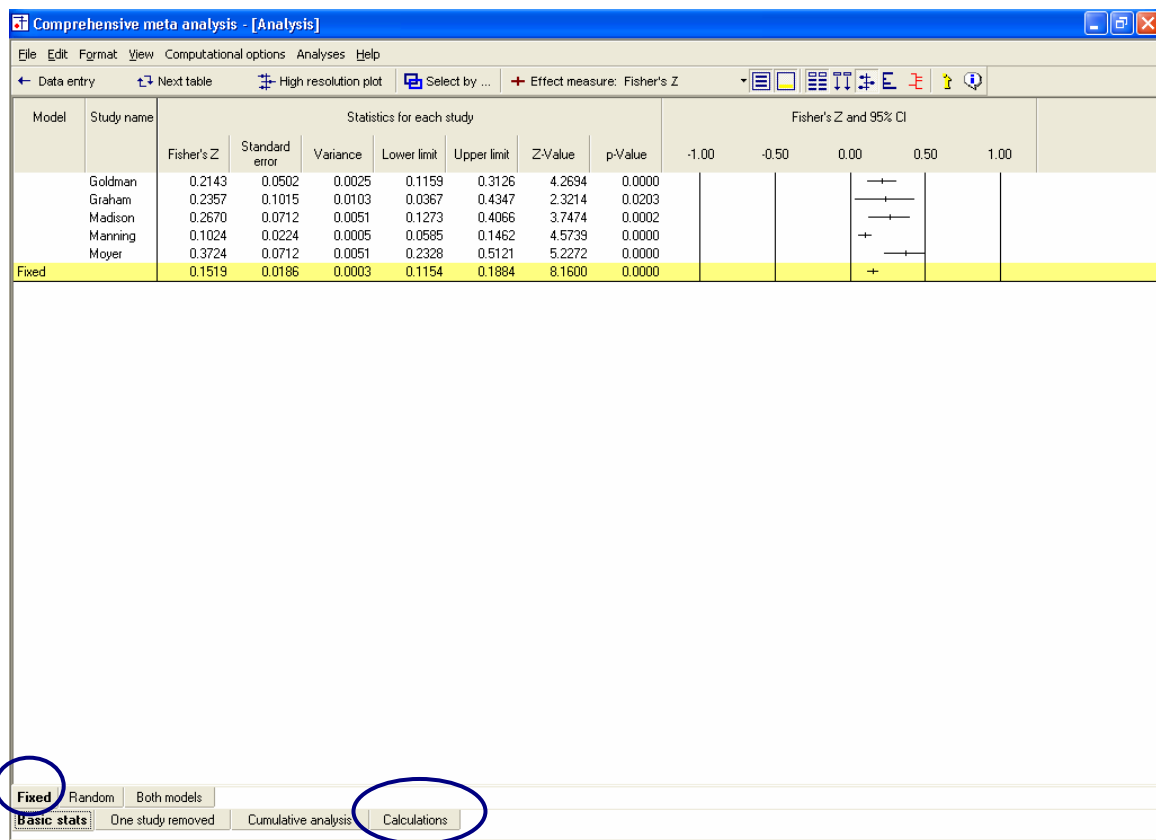
## Computational details for the fixed effect analysis

→ Select **FORMAT > INCREASE DECIMALS** on the menu

This has no effect on the computations, which are always performed using all significant digits, but it makes the example easier to follow.

→ On the bottom of the screen select **FIXED**

→ On the bottom of the screen select **CALCULATIONS**



The program switches to this display.

Model	Study name	Point	Calculations (Fixed)						
			Study Variance	Tau <sup>2</sup> Within	Tau <sup>2</sup> Between	Total Variance	IV-Weight	W	T*W
	Goldman	0.2143	0.0025	0.0000	0.0000	0.0025	337.0000	337.0000	85.0662
	Graham	0.2357	0.0103	0.0000	0.0000	0.0103	97.0000	97.0000	22.8630
	Madison	0.2670	0.0051	0.0000	0.0000	0.0051	197.0000	197.0000	52.5967
	Manning	0.1024	0.0005	0.0000	0.0000	0.0005	1997.0000	1997.0000	204.3962
	Moyer	0.3724	0.0051	0.0000	0.0000	0.0051	197.0000	197.0000	73.3673
		1.1917	0.0235	0.0000	0.0000	0.0235	2885.0000	2885.0000	438.2894

For the first study, Madison, the correlation was entered as .261 with a sample size of 200. The program computed the Fisher's Z value and its variance as follows (to see these computations return to the data entry screen and double-click on the computed value).

$$FisherZ = 0.5 * \text{Log} \left( \frac{1 + 0.261}{1 - 0.261} \right) = 0.267$$

$$SE = \frac{1}{\sqrt{200-3}} = 0.0712$$

$$Variance = 0.0712^2 = 0.0051$$

The weight is computed as

$$w_1 = \frac{1}{0.0051 + 0.0000} = 197.0000$$

Where the second term in the denominator represents tau-squared, which has been set to zero for the fixed effect analysis.

$$T_1 w_1 = (0.2670)(197.0000) = 52.5967$$

and so on for the other studies. Then, working with the sums (in the last row)

$$\bar{T} = \frac{438.2894}{2885.0000} = 0.1519$$

$$v_i = \frac{1}{2885.0000} = 0.0003$$

$$SE(\bar{T}_i) = \sqrt{0.0003} = 0.0186$$

$$\text{Lower Limit} = 0.1519 - 1.96 * 0.0186 = 0.1154$$

$$\text{Upper Limit} = 0.1519 + 1.96 * 0.0186 = 0.1884$$

$$Z = \frac{0.1519}{0.0186} = 8.1600$$

→ To switch back to the main analysis screen, click BASIC STATS at the bottom

On this screen, the values presented are the same as those computed above.

Model	Study name	Statistics for each study						Fisher's Z and 95% CI					
		Fisher's Z	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value	-1.00	-0.50	0.00	0.50	1.00
	Madison	0.2670	0.0712	0.0051	0.1273	0.4066	3.7474	0.0002					
	Moyer	0.3724	0.0712	0.0051	0.2328	0.5121	5.2272	0.0000					
	Goldman	0.2143	0.0502	0.0025	0.1159	0.3126	4.2694	0.0000					
	Graham	0.2367	0.1015	0.0103	0.0367	0.4347	2.3214	0.0203					
Fixed	Manning	0.1024	0.0224	0.0005	0.0585	0.1462	4.5739	0.0000					
		0.1519	0.0186	0.0003	0.1154	0.1884	8.1600	0.0000					

Finally, if we select Correlation as the index, the program takes the effect size and confidence interval, and displays them as correlations.

Model	Study name	Statistics for each study					Correlation and 95% CI				
		Correlation	Lower limit	Upper limit	Z-Value	p-Value	-1.00	-0.50	0.00	0.50	1.00
	Madison	0.2608	0.1267	0.3856	3.7474	0.0002					
	Moyer	0.3561	0.2287	0.4716	5.2272	0.0000					
	Goldman	0.2111	0.1154	0.3028	4.2694	0.0000					
	Graham	0.2314	0.0367	0.4092	2.3214	0.0203					
	Manning	0.1020	0.0584	0.1452	4.5739	0.0000					
Fixed		0.1508	0.1149	0.1862	8.1600	0.0000					

To transform the combined effect (Fisher's Z = 0.1519) to a correlation

$$C = \exp(2 * 0.1519) = 1.355$$

$$\bar{r} = \frac{1.355 - 1}{1.355 + 1} = 0.151$$

To transform the lower limit (Fisher's Z = 0.1154) to a correlation

$$C = \exp(2 * 0.1154) = 1.260$$

$$LowerLimit = \frac{1.260 - 1}{1.260 + 1} = 0.1149$$

To transform the upper limit (Fisher's Z = 0.1884) to a correlation

$$C = \exp(2 * 0.1884) = 1.458$$

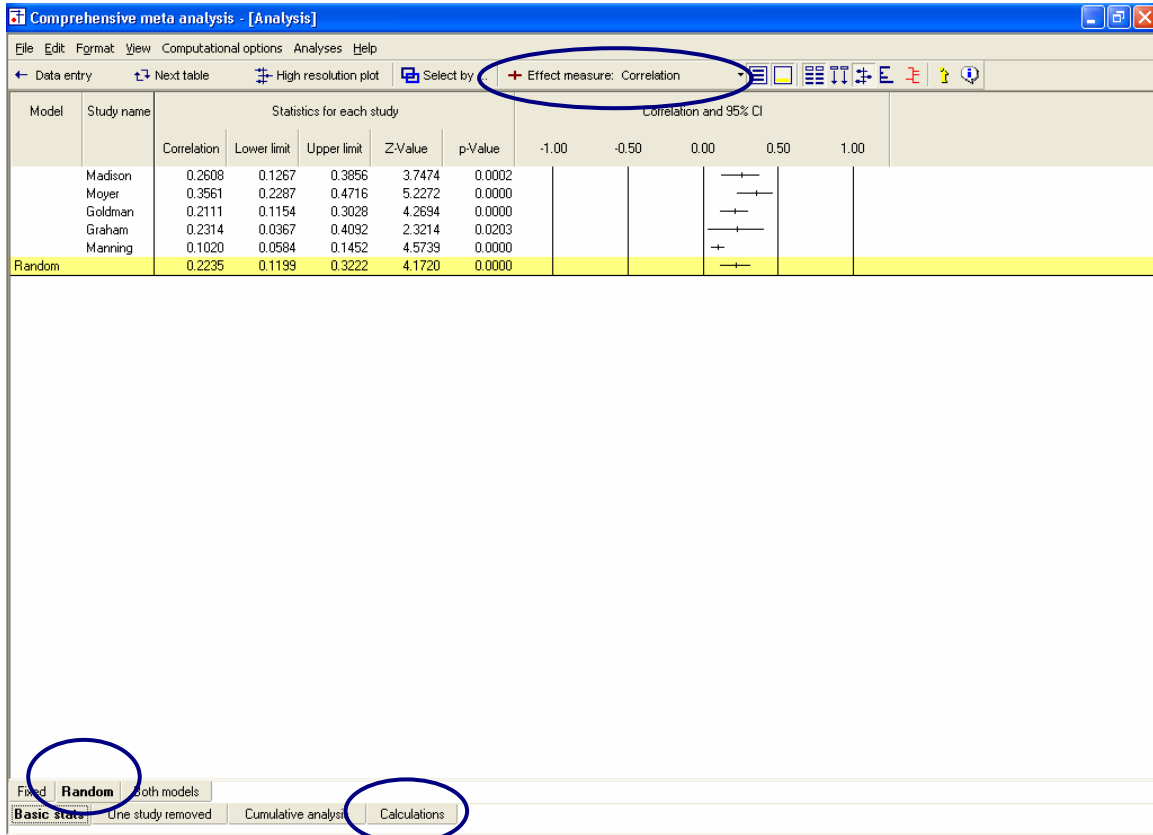
$$UpperLimit = \frac{1.458 - 1}{1.458 + 1} = 0.1862$$

The columns for variance and standard error are hidden. The z-value and p-value that had been computed using Fisher's Z values apply here as well, and are displayed without modification.

## Computational details for the random effects analysis

Now, we can repeat the exercise for random effects.

- Select FISHER'S Z as the index
- On the bottom of the screen select RANDOM
- On the bottom of the screen select CALCULATIONS



The program switches to this display

Model	Study name	Point	Study Variance	Tau <sup>2</sup> Within	Tau <sup>2</sup> Between	Total Variance	IV-Weight	W	T*W
	Madison	0.2670	0.0051	0.0108	0.0000	0.0159	62.9742	62.9742	16.8134
	Moyer	0.3724	0.0051	0.0108	0.0000	0.0159	62.9742	62.9742	23.4530
	Goldman	0.2143	0.0025	0.0108	0.0000	0.0133	75.0623	75.0623	16.0838
	Graham	0.2357	0.0103	0.0108	0.0000	0.0211	47.3650	47.3650	11.1640
	Manning	0.1024	0.0005	0.0108	0.0000	0.0113	88.4633	88.4633	9.0544
		1.1917	0.0235	0.0540	0.0000	0.0775	336.8390	336.8390	76.5686

For the first study, Madison, the correlation was entered as 0.261 with a sample size of 200. The program computed the Fisher's Z value and its variance as follows (to see these computations return to the data entry screen and double-click on the computed value).

$$FisherZ = 0.5 * \text{Log}\left(\frac{1 + 0.261}{1 - 0.261}\right) = 0.267$$

$$SE = \frac{1}{\sqrt{200-3}} = 0.0712$$

$$Variance = 0.0712^2 = 0.0051$$

The weight is computed as

$$w_1^* = \frac{1}{0.0051 + 0.0108} = \frac{1}{0.0159} = 62.9742$$

Where the (\*) indicates that we are using random effects weights, and the second term in the denominator represents tau-squared.

$$T_1^* w_1^* = (.2670)(62.9742) = 16.8134$$

and so on for the other studies. Then, working with the sums (in the last row)



$$\bar{T}_* = \frac{76.5686}{336.8390} = 0.2273$$

$$v_* = \frac{1}{336.8390} = 0.0030$$

$$SE(\bar{T}_*) = \sqrt{0.0030} = .0545$$

$$\text{Lower Limit}^* = 0.2273 - 1.96 * 0.0545 = 0.1205$$

$$\text{Upper Limit}^* = 0.2273 + 1.96 * 0.0545 = 0.3341$$

$$Z^* = \frac{0.2273}{0.0545} = 4.1720$$

$$p_{2T}^* = 2[1 - (\Phi(|4.1720|))] < 0.0001$$

(Note – The column labeled TAU-SQUARED WITHIN is actually tau-squared between studies, and the column labeled TAU-SQUARED BETWEEN is reserved for a fully random effects analysis, where we are performing an analysis of variance).

→ To switch back to the main analysis screen, click BASIC STATS at the bottom

On this screen, the values presented are the same as those computed above.

Model	Study name	Statistics for each study						Fisher's Z and 95% CI					
		Fisher's Z	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value	-1.00	-0.50	0.00	0.50	1.00
	Madison	0.2670	0.0712	0.0051	0.1273	0.4066	3.7474	0.0002					
	Moyer	0.3724	0.0712	0.0051	0.2328	0.5121	5.2272	0.0000					
	Goldman	0.2143	0.0502	0.0025	0.1159	0.3126	4.2694	0.0000					
	Graham	0.2357	0.1015	0.0103	0.0367	0.4347	2.3214	0.0203					
	Manning	0.1684	0.0224	0.0005	0.0585	0.1462	4.5739	0.0000					
Random		0.2273	0.0545	0.0030	0.1205	0.3341	4.1720	0.0000					

Finally, if we select Correlation as the index, the program takes the effect size and confidence interval, and displays them as correlations. The columns for variance and standard error are then hidden.

Model	Study name	Statistics for each study					Correlation and 95% CI				
		Correlation	Lower limit	Upper limit	Z-Value	p-Value	-1.00	-0.50	0.00	0.50	1.00
	Madison	0.2608	0.1267	0.3856	3.7474	0.0002				+	
	Moyer	0.3561	0.2287	0.4716	5.2272	0.0000				+	
	Goldman	0.2111	0.1154	0.3028	4.2694	0.0000				+	
	Graham	0.2314	0.0367	0.4092	2.3214	0.0203				+	
	Manning	0.1020	0.0584	0.1452	4.5739	0.0000				+	
Random		0.2235	0.1199	0.3222	4.1720	0.0000				+	

To transform the combined effect (Fisher's Z = 0.2273) to a correlation

$$C = \exp(2 * 0.2273) = 1.576$$

$$\bar{r} = \frac{1.576 - 1}{1.576 + 1} = 0.2235$$

To transform the lower limit (Fisher's Z = 0.1205) to a correlation

$$C = \exp(2 * 0.1205) = 1.273$$

$$LowerLimit = \frac{1.273 - 1}{1.273 + 1} = 0.1199$$

To transform the upper limit (Fisher's Z = 0.3341) to a correlation

$$C = \exp(2 * 0.3341) = 1.951$$

$$UpperLimit = \frac{1.951 - 1}{1.951 + 1} = 0.3222$$

The columns for variance and standard error are hidden. The z-value and p-value that had been computed using log values apply here as well, and are displayed without modification.

This is the conclusion of the exercise for correlational data.