Example name	BCG
Effect size	Risk ratio
Analysis type	Subgroups analysis, Regression
Level	Advanced

Synopsis

This analysis includes studies where patients were randomized to receive either a vaccine or a placebo Outcome was the proportion of patients each group who developed TB. The analysis focused on the risk ratio.

There was substantial dispersion in the risk ratio. The researchers looked to see if this was related to the study's latitude (a surrogate for the vaccine's potency and for the natural immunity in the population) and several other variables.

We use this example to show

- How to interpret a basic analysis using Risk Ratio
- How to understand the heterogeneity statistics
- How to perform a subgroups analysis
- How a regression analysis corresponds to a subgroups analysis
- How to perform a regression analysis using a categorical predictor
- How to perform a regression analysis using a continuous predictor

To open a CMA file > Download and Save file | Start CMA | Open file from within CMA

Download CMA file for computers that use a period to indicate decimals
Download CMA file for computers that use a comma to indicate decimals

Download this PDF
Download data in Excel
Download trial of CMA

QUICK START

1) On the data-entry screen
   a) Create a column for study name
   b) Create a set of columns for the effect size
   c) Identify one or more columns as "Moderators" and set the subtype to either "Integer", "Decimal", or "Categorical
   d) Enter the data
2) On the main analysis screen
   a) Optionally, select the effect size index
   b) Optionally, select the studies to be included in the regression
   c) Optionally, specify how to work with studies that included multiple subgroups, outcomes, time-point, or comparisons.
   d) Click Analyses > Meta-regression 2
3) On the regression screen – define the regression
   a) Select the covariates to be included in the model
   b) Optionally, define "Sets" of covariates
   c) Optionally, define multiple models
   d) Optionally, select statistical options
   e) Run the analysis
4) On the regression screen – navigate the results
   a) Click "Fixed" or "Random" to select the model
   b) Click the model name (when several models have been created)
   c) Use the toolbar to move between the main analysis screen, the scatterplot, diagnostics, increments, model comparisons, and other screens
5) On the regression screen – save the analysis
6) On the regression screen – export the results

STEP 1: ENTER THE DATA

Insert column for study names

Click Insert > Column for > Study names

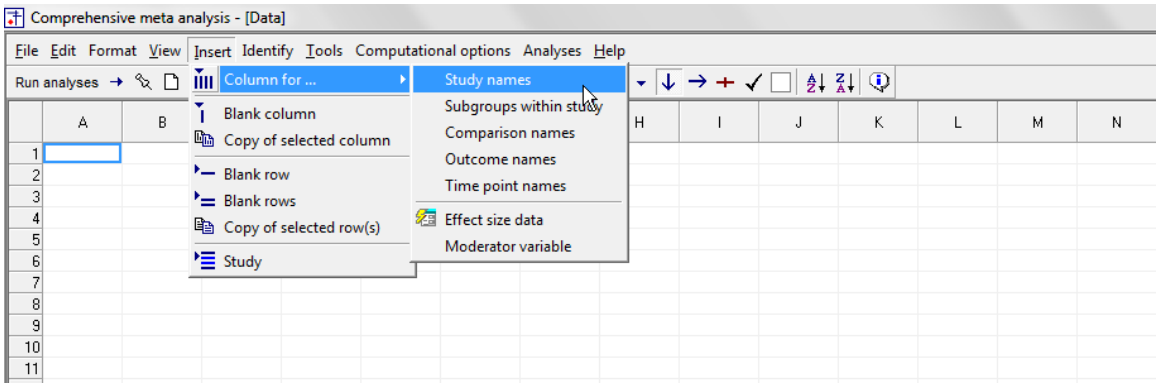


Figure 1

The program creates a column labeled “Study name”

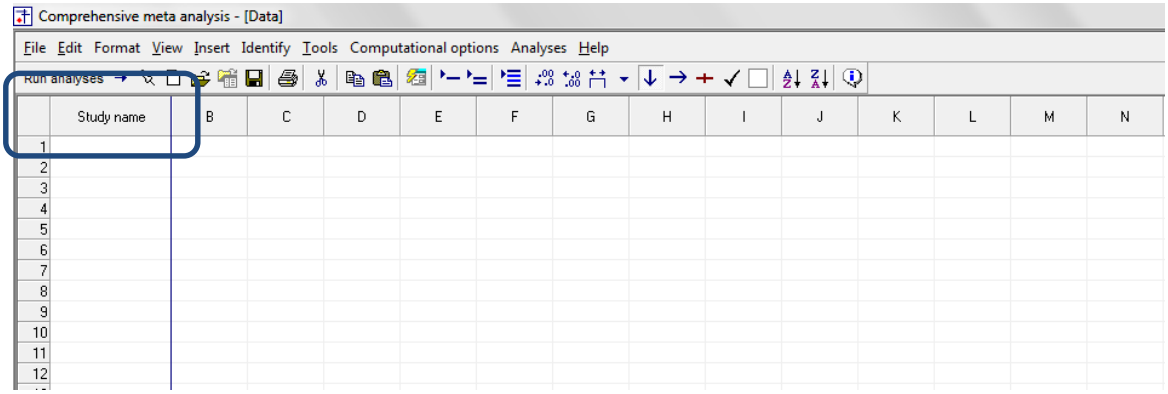


Figure 2

Insert columns for effect size data

Click Insert > Column for > Effect size data

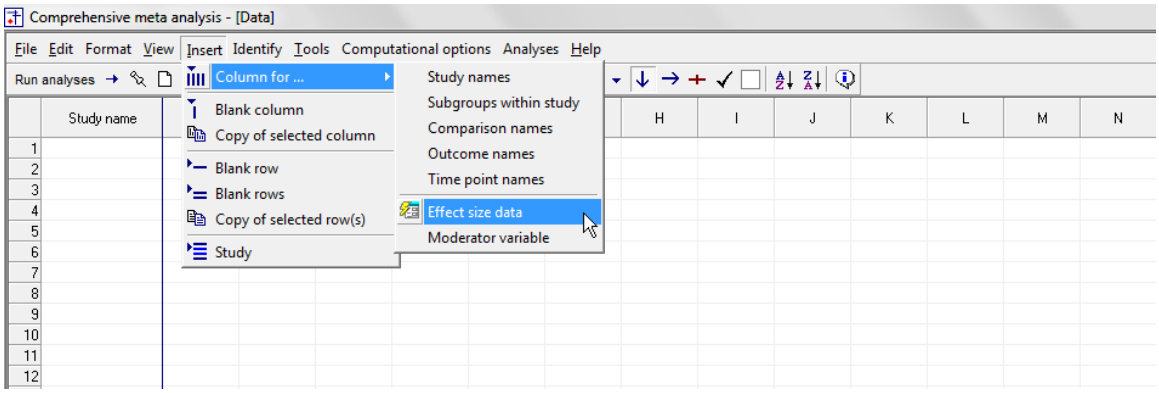


Figure 3

The program opens a wizard that allows you to specify the kind of summary data you will enter

- Select <Show all 100 formats>
- Click <Next>

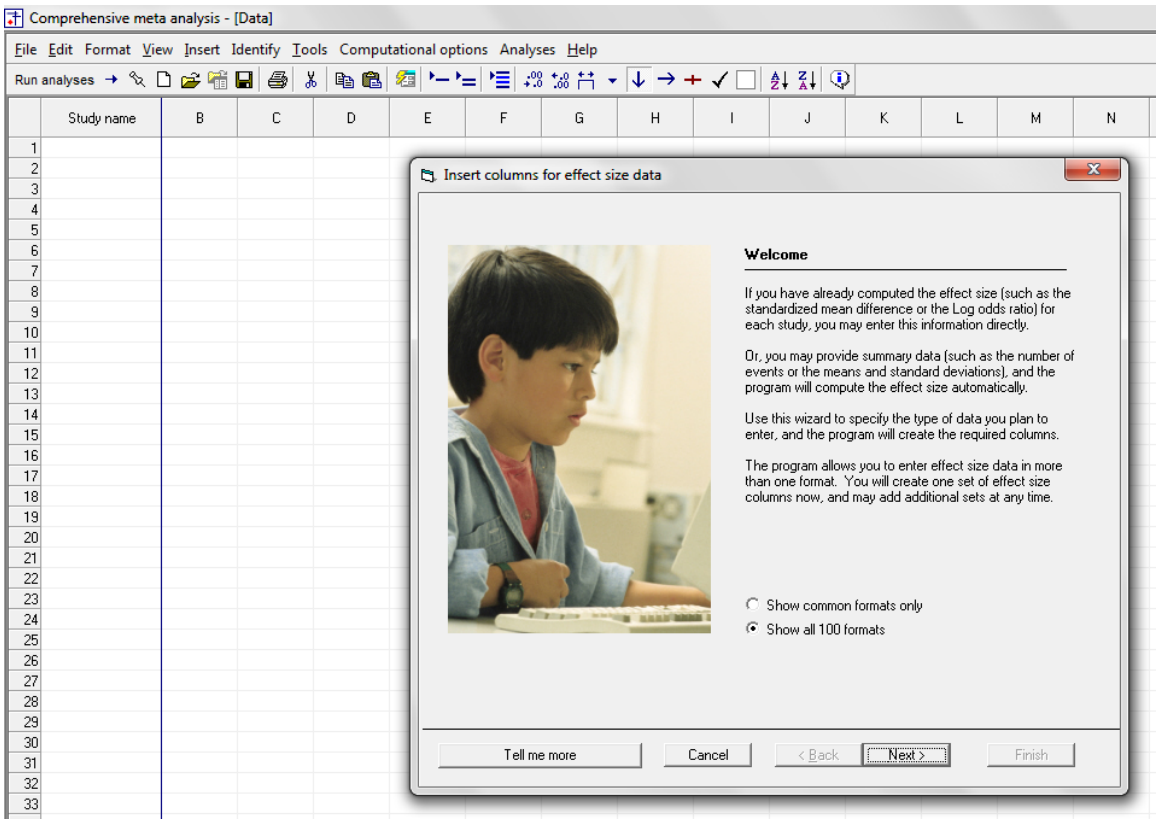


Figure 4

- Select the top option button
- Click <Next>

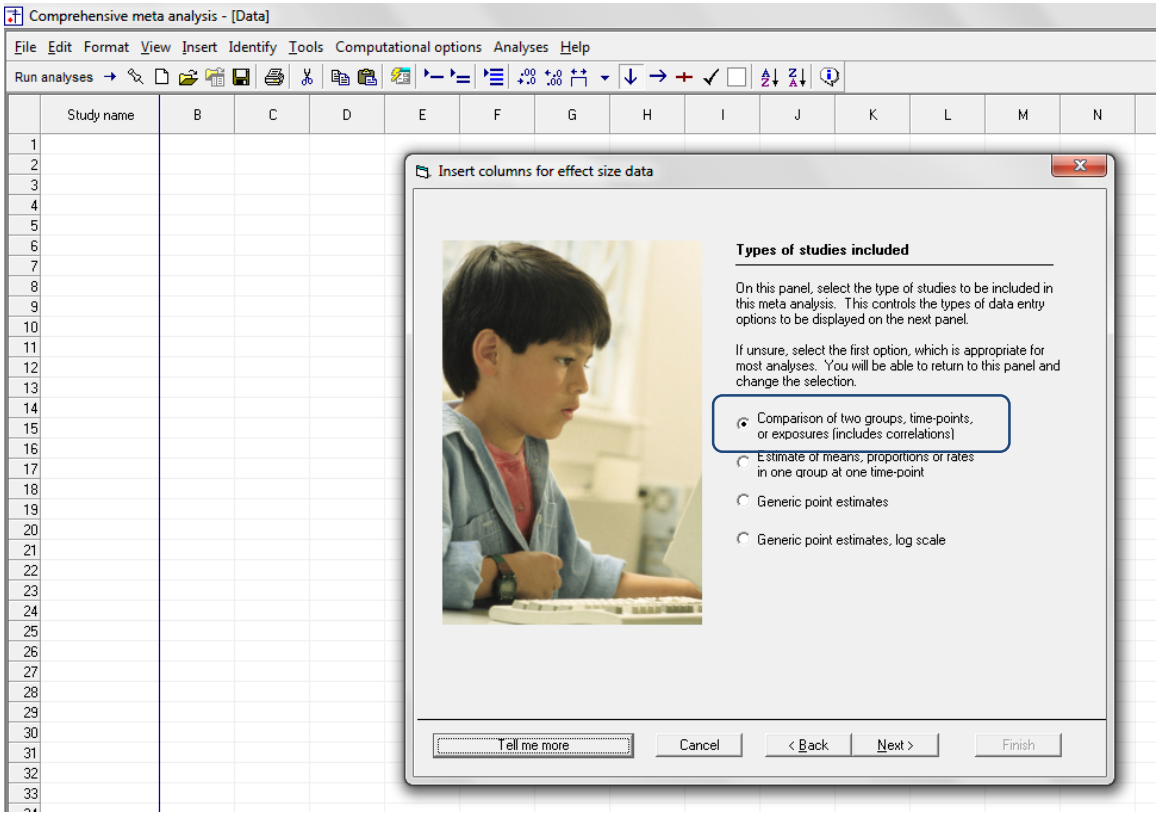


Figure 5

On this screen, drill down to

- Dichotomous (number of events)
- Unmatched groups, prospective (e.g., controlled trials, cohort studies)
- Events and sample size in each group [H]

Then, click <Finish>

Note that we will be entering events and sample size (*N*) for each group. Some of the texts that use the BCG example report events and non-events rather than events and *N*.

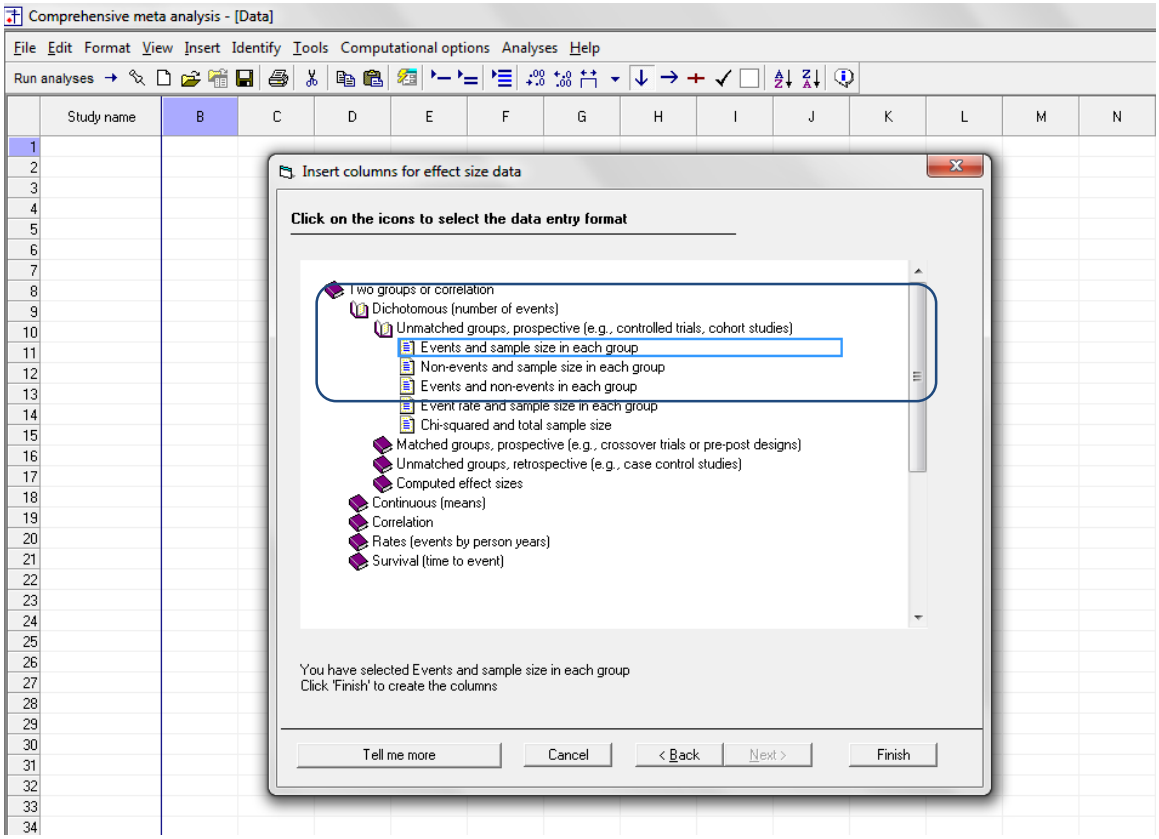


Figure 6

The program creates columns as shown here. It also opens a wizard that allows you to label the columns.

- Enter Vaccine/Control as names for the two groups
- Enter TB/Ok as names for the two outcomes

Then, click Ok

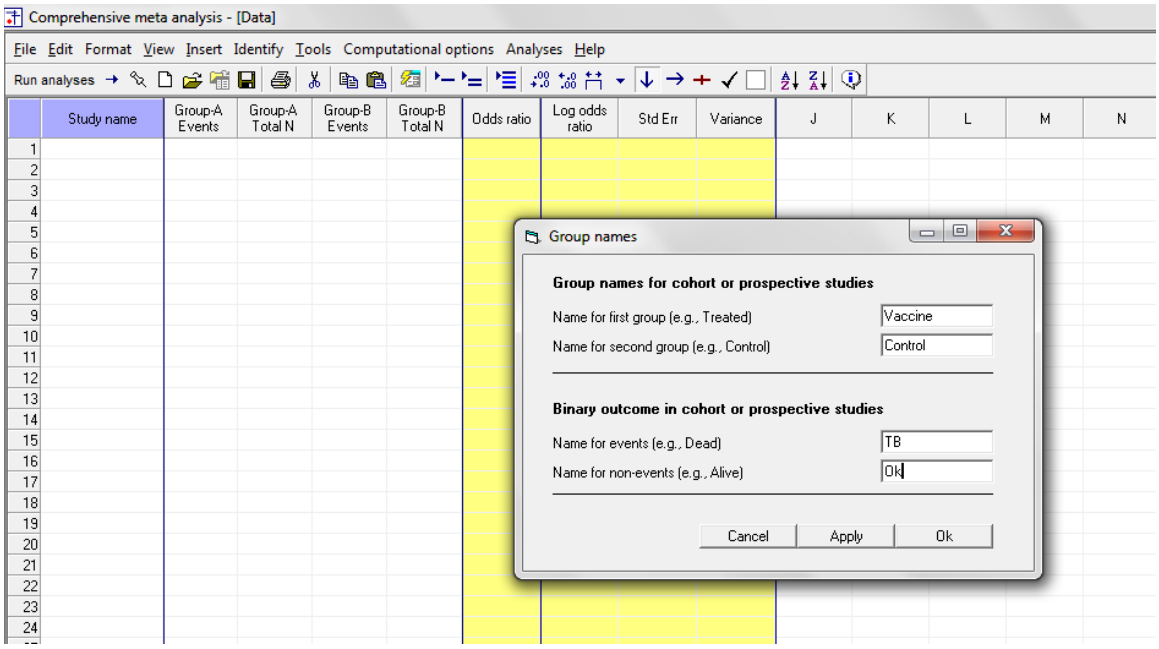


Figure 7

The program applies the labels as shown here [K].

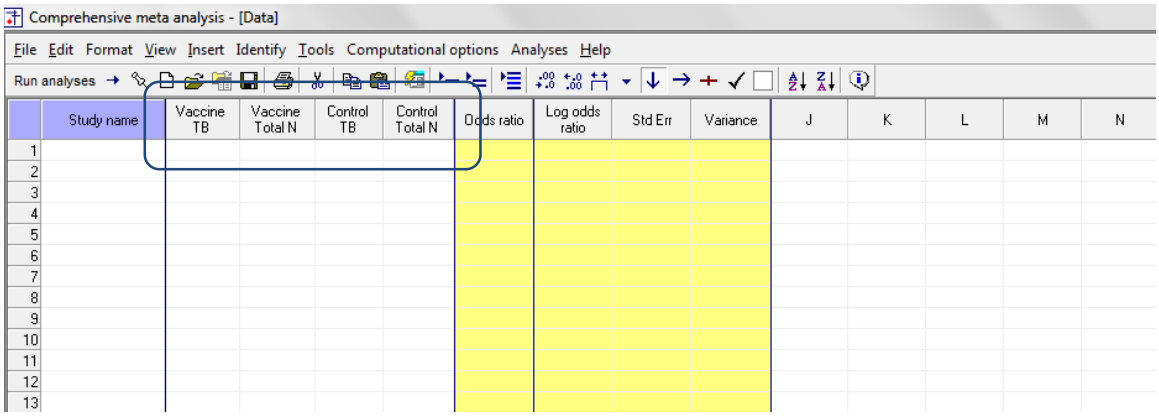


Figure 8

Insert columns for moderators (covariates)

Next, we need to create columns for the moderator variables.

- Click Insert > Column for > Moderator variable

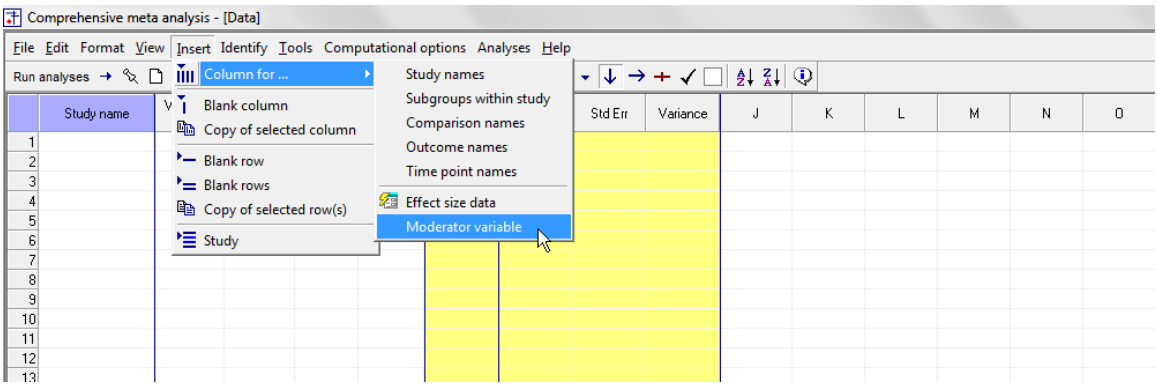


Figure 9

The program opens a wizard

- Set the variable name to “Latitude”
- Set the column function to Moderator
- Set the data type to Integer

Then, click OK

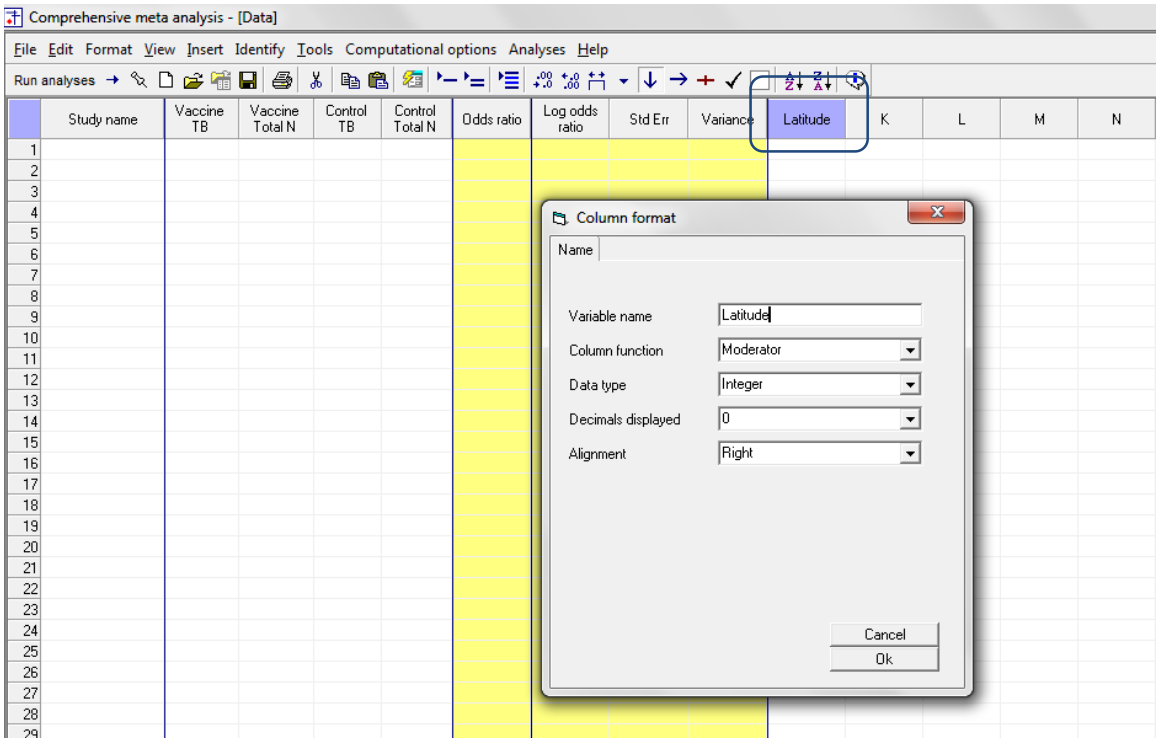


Figure 10

Click Insert > Column for > Moderator variable

- Set the variable name to “Year”
- Set the column function to Moderator
- Set the data type to Integer

(This is the year the study was performed, not the year of publication)

Then, click OK

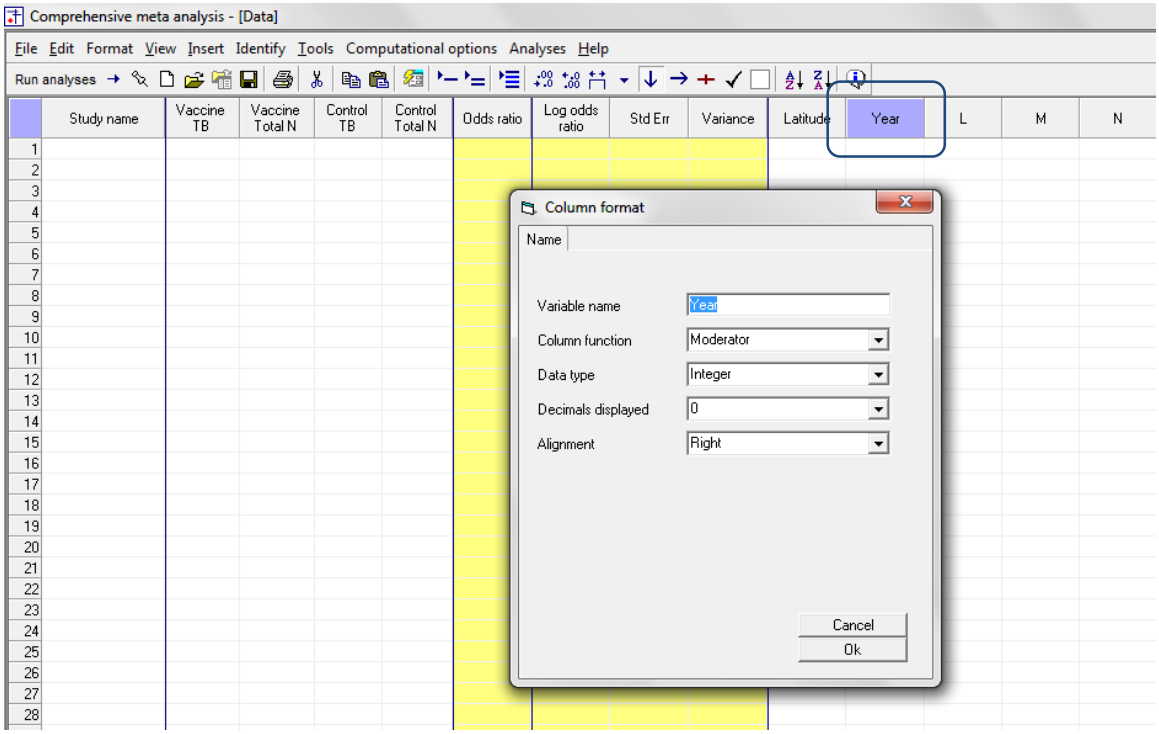


Figure 11

Click Insert > Column for > Moderator variable

- Set the variable name to “Allocation”
- Set the column function to Moderator
- Set the data type to Categorical
-

This moderator tracks the mechanism utilized to assign people to be vaccinated (or not). The possibilities are random, alternate, and systematic.

Then, click [Ok]

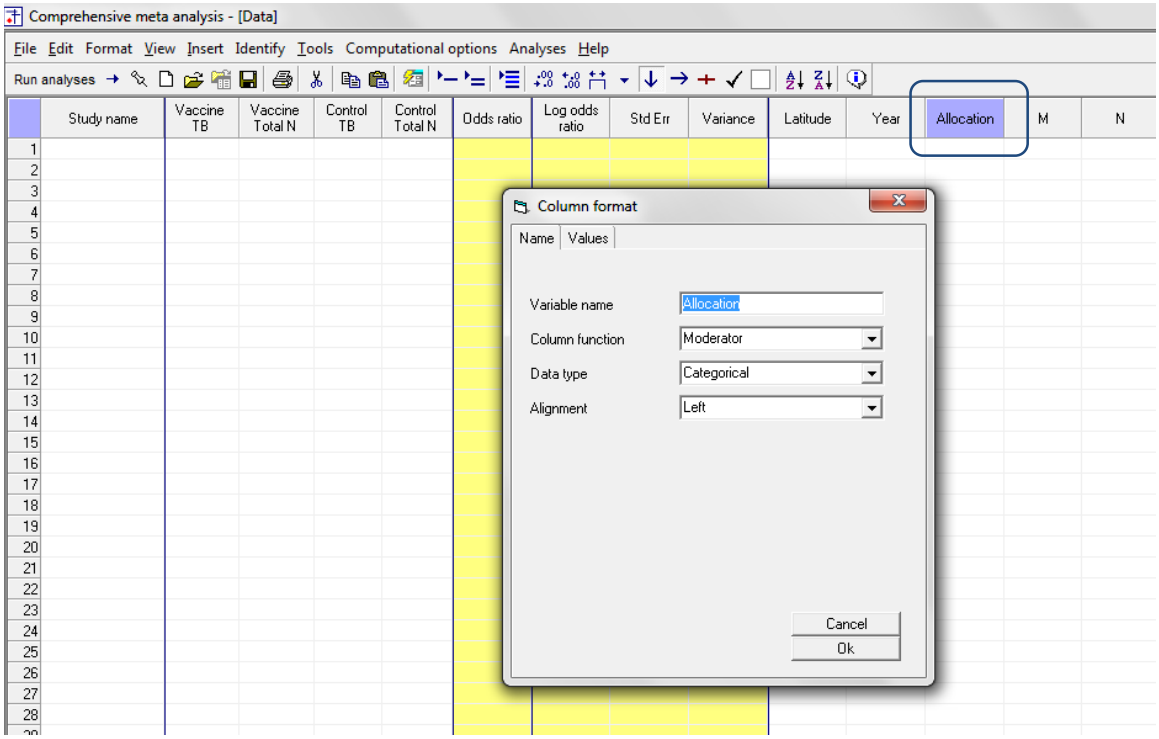


Figure 12

Click Insert > Column for > Moderator variable

- Set the variable name to “Climate”
- Set the column function to Moderator
- Set the data type to Categorical
-

This moderator tracks the climate. The possibilities are Cold and Hot.

Then, click [OK]

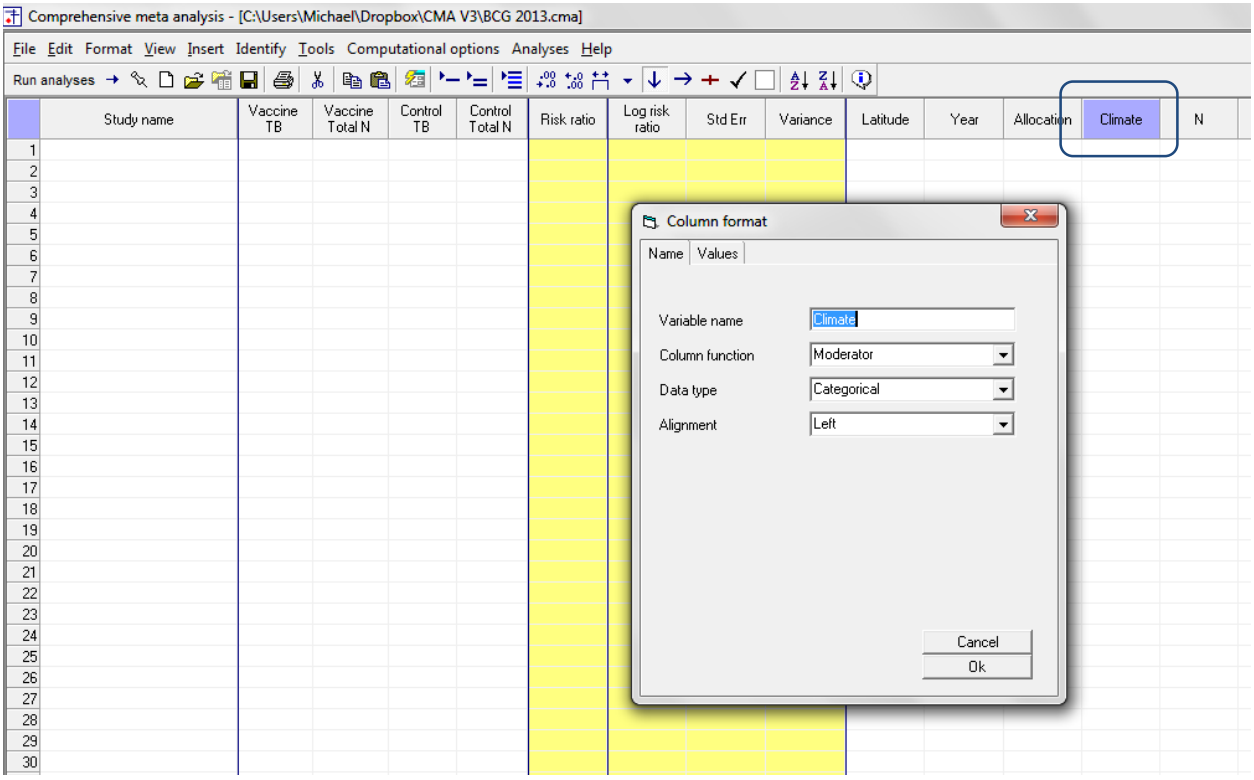


Figure 13

Customize the screen

The program initially displays the odds ratio. Since we want to work with the log risk ratio we need to customize the display. (We will use the log risk ratio rather than the risk ratio since the computations are easier to understand in log units).

- Right-click in any yellow column
- Click <Customize computed effect size display> [A]

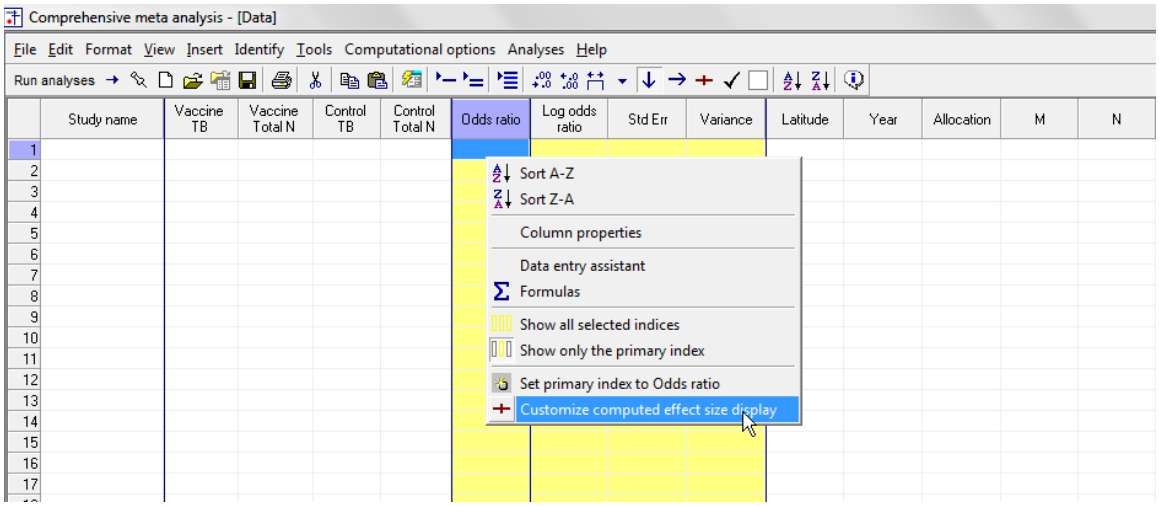


Figure 14

The program displays this wizard

- Tick the box for Risk ratio
- Tick the box for Log risk ratio

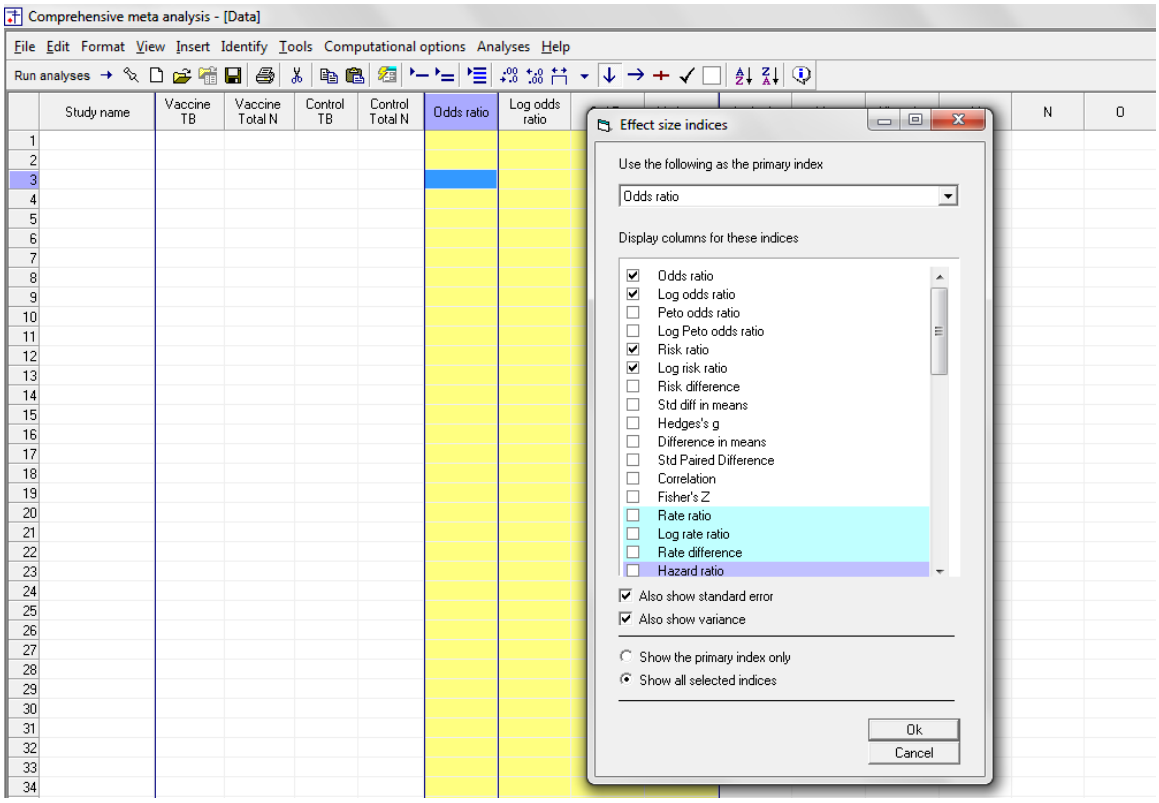


Figure 15

To keep things simple we can set Log risk ratio as the default effect size, and also hide the odds ratio

- In the drop-down box, select Log risk ratio as the primary index
- Un-check the box for odds ratio
- Un-check the box for log odds ratio

Then click [Ok]

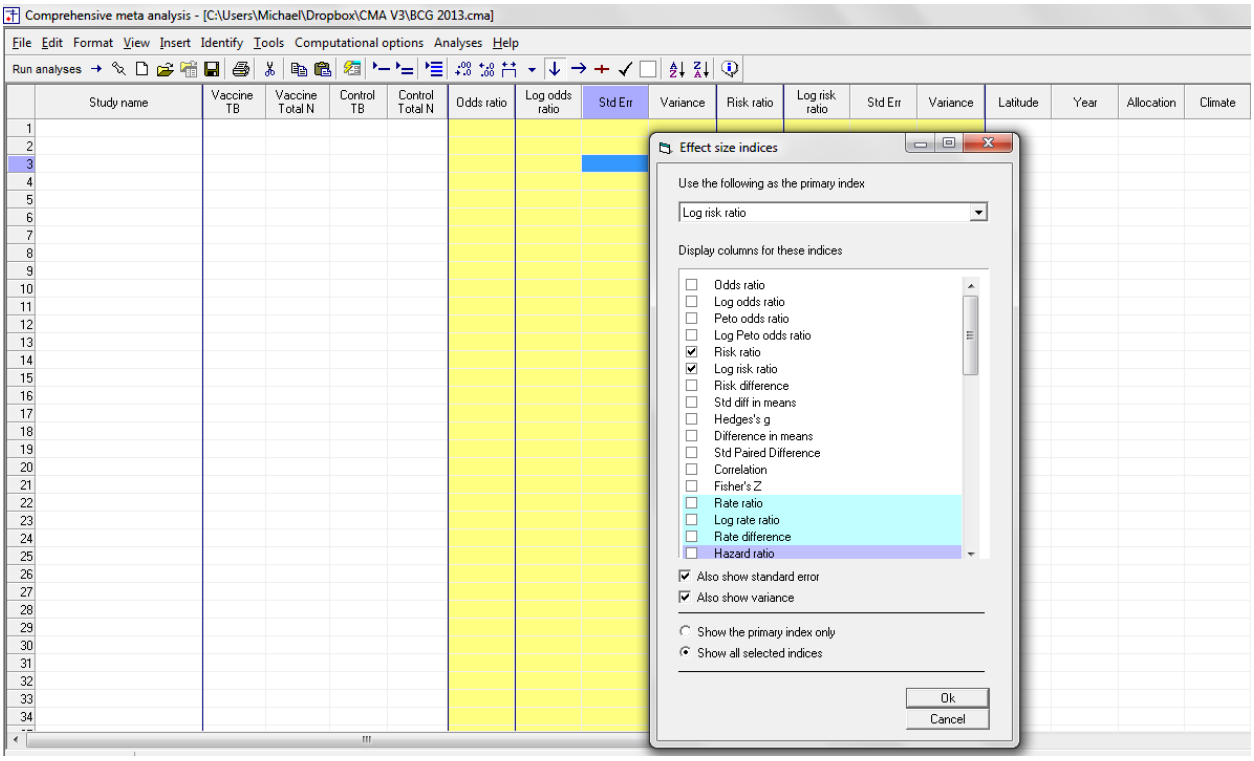


Figure 16

The screen now looks like this

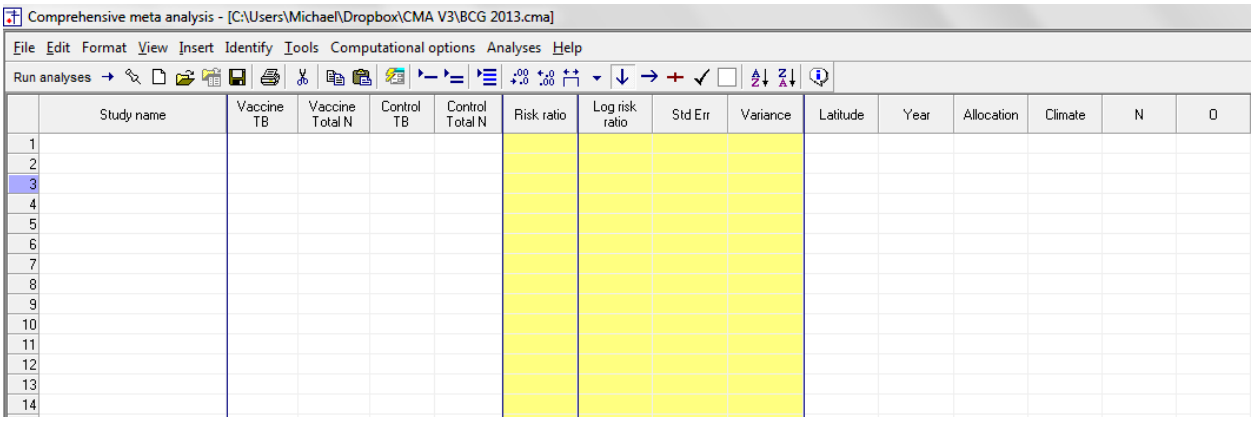
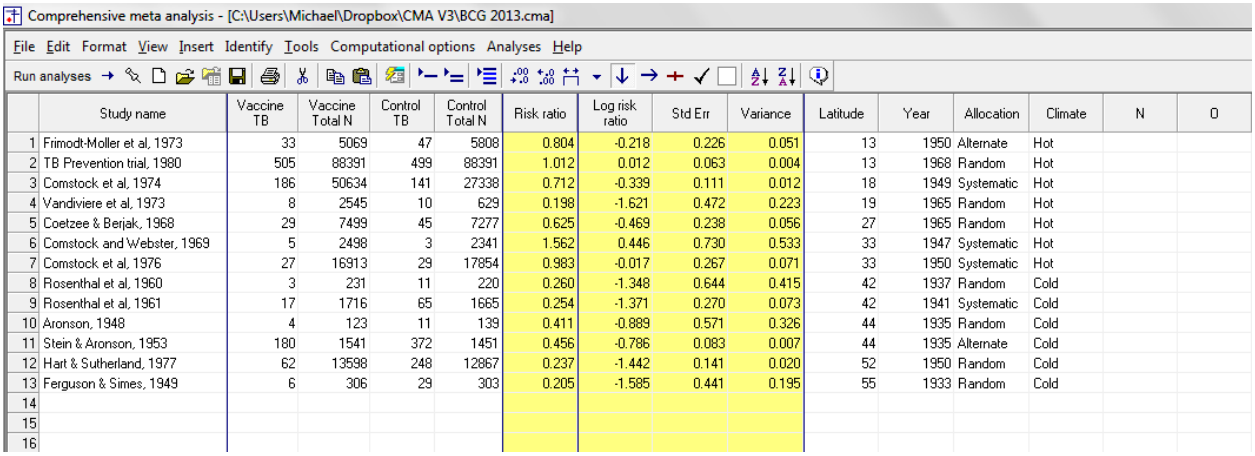


Figure 17

Enter the data

You can enter the data manually, or copy and paste from Excel ™ or another source (see appendix)

Note that you enter data into the white columns. The program automatically computes the values in the yellow columns.

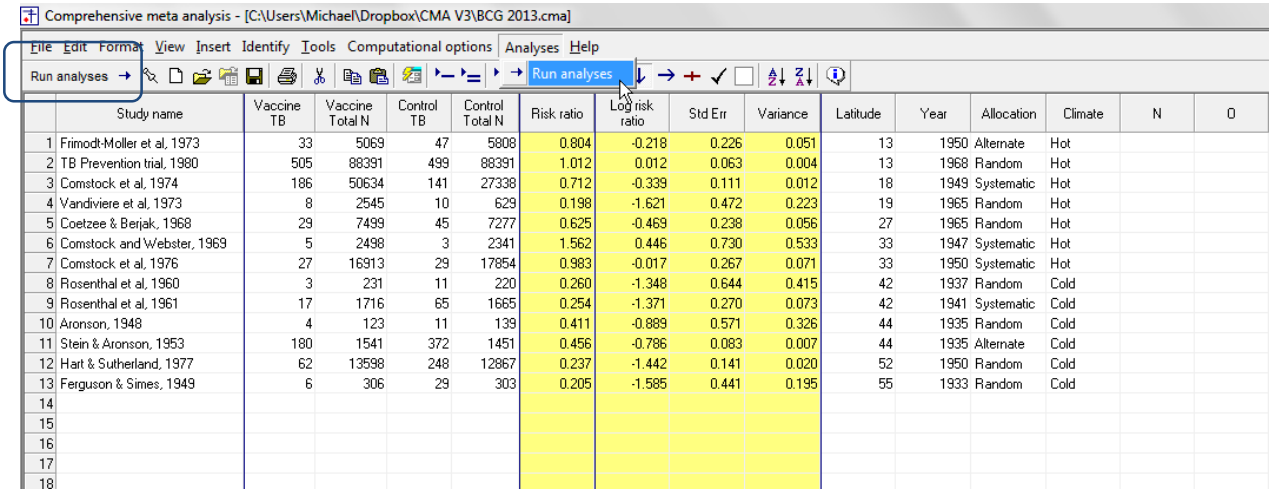


	Study name	Vaccine TB	Vaccine Total N	Control TB	Control Total N	Risk ratio	Log risk ratio	Std Err	Variance	Latitude	Year	Allocation	Climate	N	O
1	Frimodt-Moller et al, 1973	33	5069	47	5808	0.804	-0.218	0.226	0.051	13	1950	Alternate	Hot		
2	TB Prevention trial, 1980	505	88391	499	88391	1.012	0.012	0.063	0.004	13	1968	Random	Hot		
3	Comstock et al, 1974	186	50634	141	27338	0.712	-0.339	0.111	0.012	18	1949	Systematic	Hot		
4	Vandiviere et al, 1973	8	2545	10	629	0.198	-1.621	0.472	0.223	19	1965	Random	Hot		
5	Coetzee & Berjak, 1968	29	7499	45	7277	0.625	-0.469	0.238	0.056	27	1965	Random	Hot		
6	Comstock and Webster, 1969	5	2498	3	2341	1.562	0.446	0.730	0.533	33	1947	Systematic	Hot		
7	Comstock et al, 1976	27	16913	29	17854	0.983	-0.017	0.267	0.071	33	1950	Systematic	Hot		
8	Rosenthal et al, 1960	3	231	11	220	0.260	-1.348	0.644	0.415	42	1937	Random	Cold		
9	Rosenthal et al, 1961	17	1716	65	1665	0.254	-1.371	0.270	0.073	42	1941	Systematic	Cold		
10	Aronson, 1948	4	123	11	139	0.411	-0.889	0.571	0.326	44	1935	Random	Cold		
11	Stein & Aronson, 1953	180	1541	372	1451	0.456	-0.786	0.083	0.007	44	1935	Alternate	Cold		
12	Hart & Sutherland, 1977	62	13598	248	12867	0.237	-1.442	0.141	0.020	52	1950	Random	Cold		
13	Ferguson & Simes, 1949	6	306	29	303	0.205	-1.585	0.441	0.195	55	1933	Random	Cold		
14															
15															
16															

Figure 18

STEP 2: RUN THE BASIC META-ANALYSIS

To run the analysis, click <Run Analysis>



	Study name	Vaccine TB	Vaccine Total N	Control TB	Control Total N	Risk ratio	Log risk ratio	Std Err	Variance	Latitude	Year	Allocation	Climate	N	O
1	Frimodt-Moller et al, 1973	33	5069	47	5808	0.804	-0.218	0.226	0.051	13	1950	Alternate	Hot		
2	TB Prevention trial, 1980	505	88391	499	88391	1.012	0.012	0.063	0.004	13	1968	Random	Hot		
3	Comstock et al, 1974	186	50634	141	27338	0.712	-0.339	0.111	0.012	18	1949	Systematic	Hot		
4	Vandiviere et al, 1973	8	2545	10	629	0.198	-1.621	0.472	0.223	19	1965	Random	Hot		
5	Coetzee & Berjak, 1968	29	7499	45	7277	0.625	-0.469	0.238	0.056	27	1965	Random	Hot		
6	Comstock and Webster, 1969	5	2498	3	2341	1.562	0.446	0.730	0.533	33	1947	Systematic	Hot		
7	Comstock et al, 1976	27	16913	29	17854	0.983	-0.017	0.267	0.071	33	1950	Systematic	Hot		
8	Rosenthal et al, 1960	3	231	11	220	0.260	-1.348	0.644	0.415	42	1937	Random	Cold		
9	Rosenthal et al, 1961	17	1716	65	1665	0.254	-1.371	0.270	0.073	42	1941	Systematic	Cold		
10	Aronson, 1948	4	123	11	139	0.411	-0.889	0.571	0.326	44	1935	Random	Cold		
11	Stein & Aronson, 1953	180	1541	372	1451	0.456	-0.786	0.083	0.007	44	1935	Alternate	Cold		
12	Hart & Sutherland, 1977	62	13598	248	12867	0.237	-1.442	0.141	0.020	52	1950	Random	Cold		
13	Ferguson & Simes, 1949	6	306	29	303	0.205	-1.585	0.441	0.195	55	1933	Random	Cold		
14															
15															
16															
17															
18															

Figure 19

The main analysis screen

The program displays the main analysis screen.

The program shows the current effect size, which is “Log risk ratio”. If you want to switch to another effect size, click on “Log risk ratio”.

The next few pages outline the main analysis in *CMA*. However, this is optional. You can use the menu Analysis > Meta-regression 2 to proceed immediately to the regression module.

The initial meta-analysis

At the top, we have selected [Log risk ratio] as the effect size.

At the bottom the <Fixed> tab is selected, so the program is displaying the results for a fixed-effect analysis.

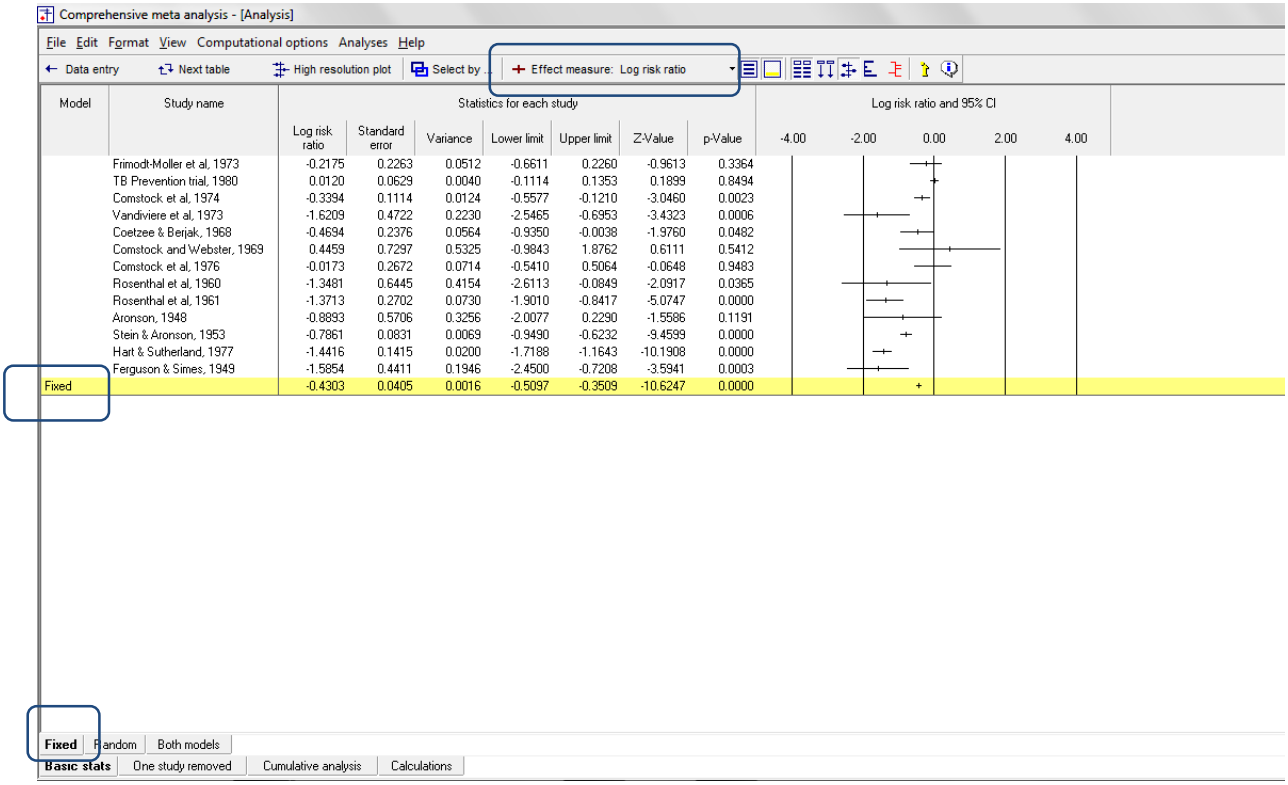


Figure 20

Click the tab for <Random>. The program displays results for a random-effects analysis.

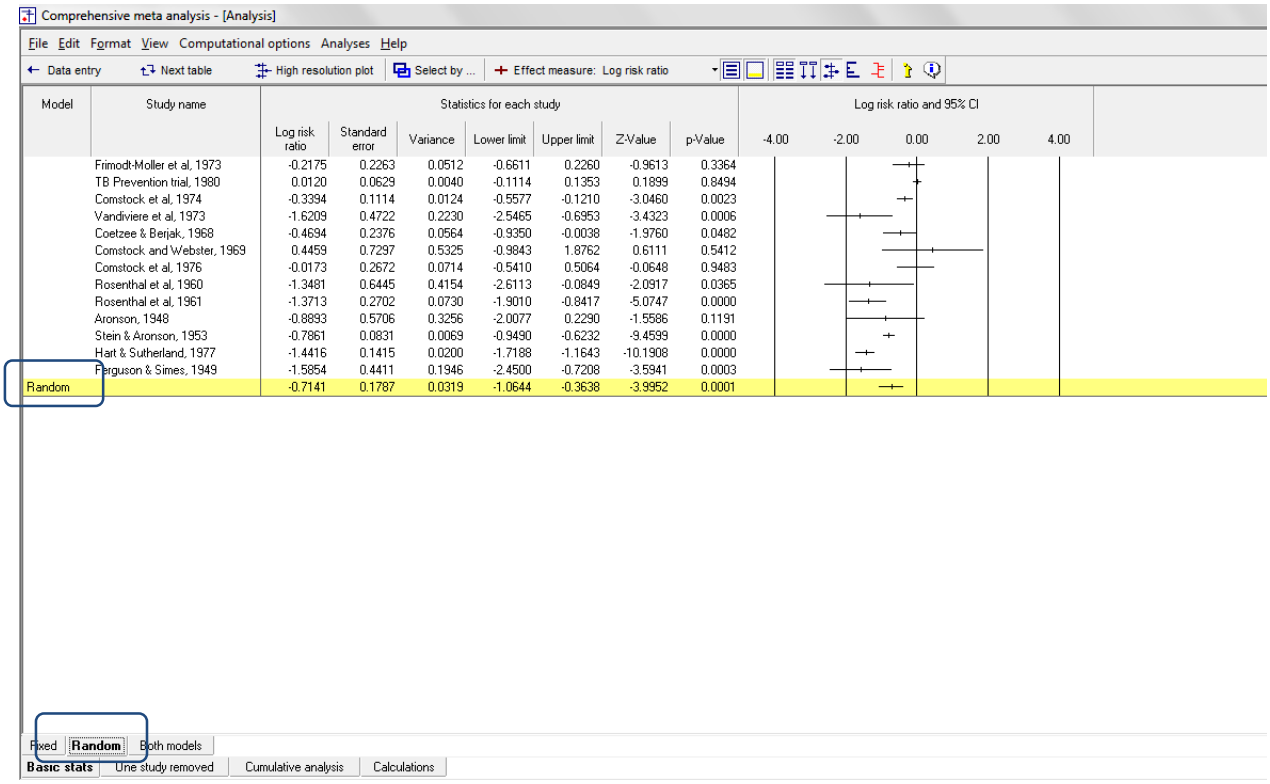


Figure 21

Display moderator variables

Next, we want to display the moderator variables on the plot.
Note that this is optional, and has no effect on the regression.

Click View > Columns > Moderators

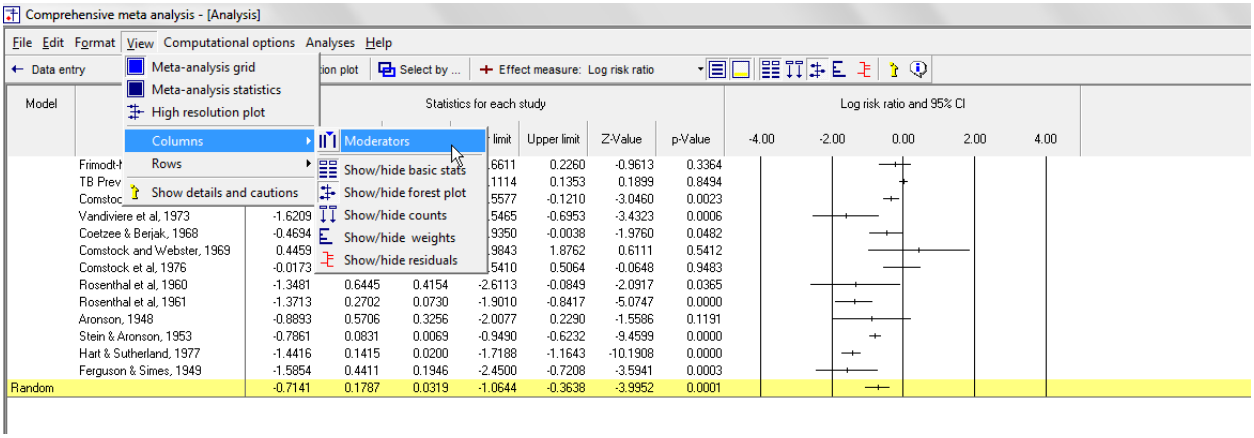


Figure 22

The program displays a list of all variables that had been defined as moderators on the data-entry screen.

Drag and drop each of these onto the main screen, to the right of the "*p*-value" column [G].

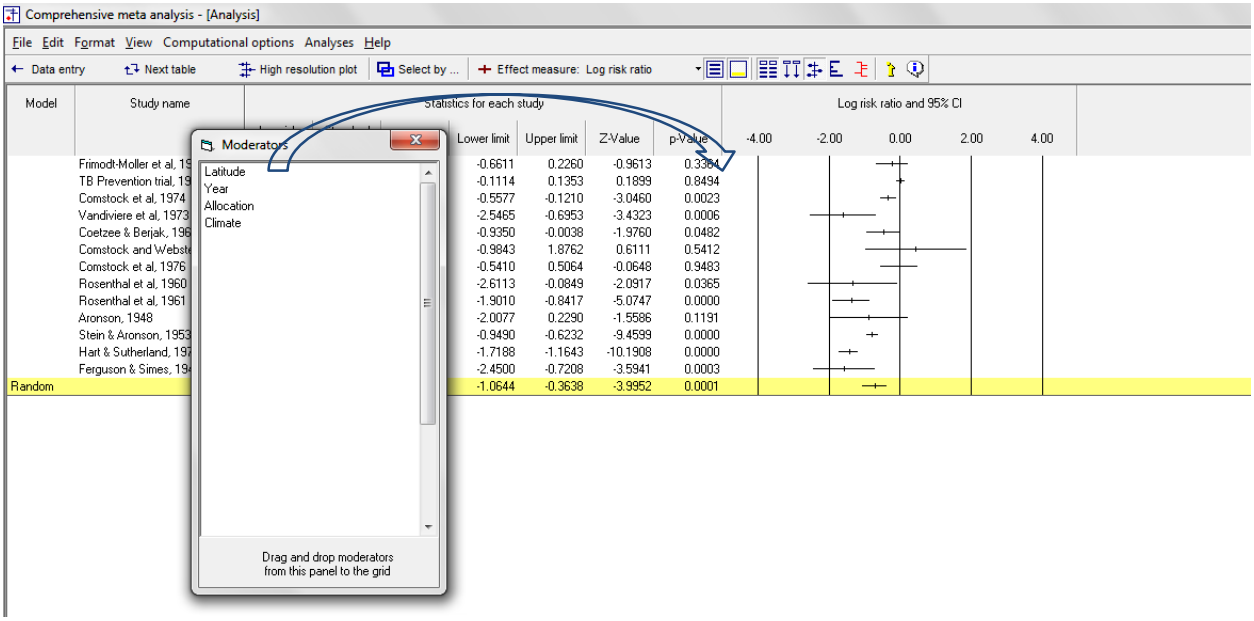


Figure 23

The screen should now look like this.

Since the data had been sorted by latitude on the data-entry screen, the program initially displays the studies in that sequence. It appears that the effect size is minor (near 1.0) toward the top (for studies near the equator)and larger (as extreme as 0.20) toward the bottom (for studies in colder climates).

You can right-click on any column and sort by that column.

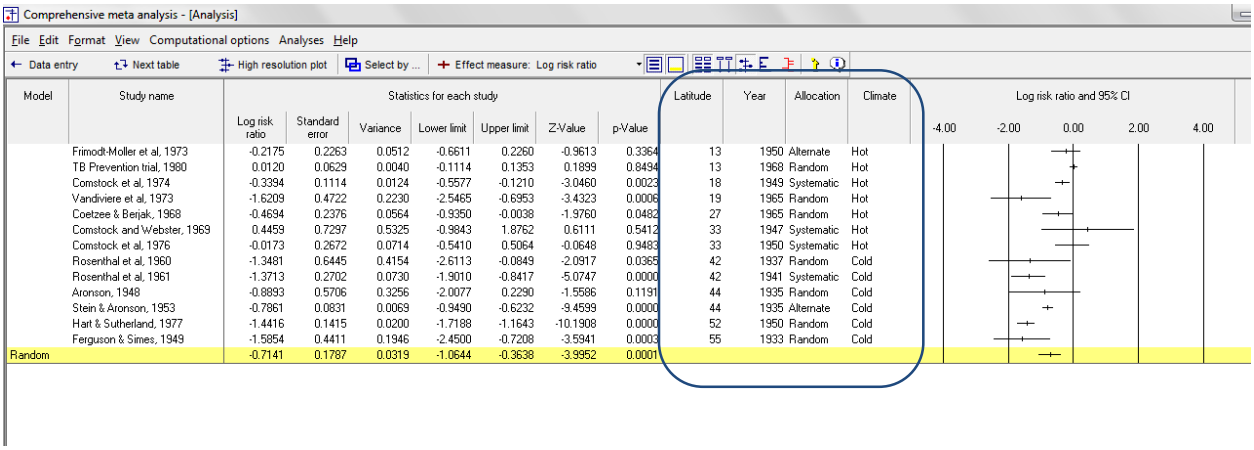


Figure 24

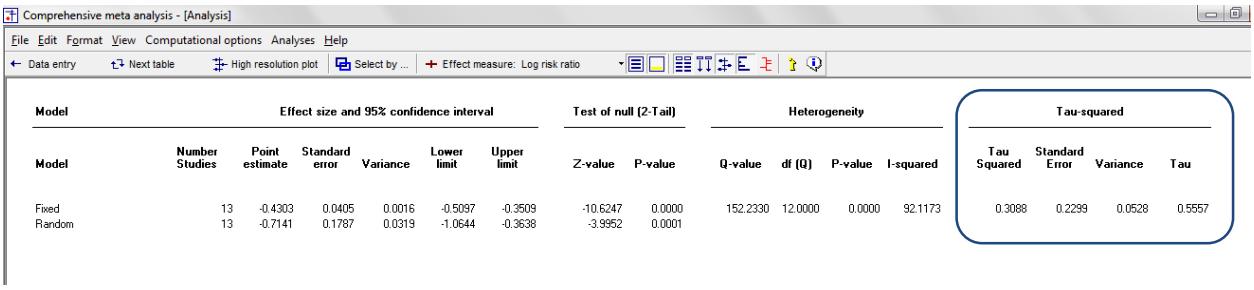
Display statistics

Click <Next table> to display the statistics shown here.

Using random-effects weights [J], the summary log risk ratio is -0.7141. The Z-value is -3.995 with a corresponding *p*-value of 0.0001. Thus, we can reject the null hypothesis that log risk ratio is 0.0 (or that the risk ratio is 1.0). If we assume that the studies are valid and that random assignment was carried out properly, we can conclude that, on the average, the vaccine does prevent TB.

At the same time, there is also a substantial amount of dispersion in the effect size. Tau-squared [K] is 0.3088 and Tau is 0.5557. To get a general sense of the true dispersion we can assume that the true effects are balanced about the random-effects estimate of the mean effect, and that some 95% of all true effects fall within 1.96 *T* of this mean. Then (in log units) most true effects fall in the range of −1.8032 to +0.3750. This corresponds to risk ratios of approximately 0.16 (a strongly protective effect) to 1.46 (a harmful effect).

It would be very important to understand the reason for this dispersion, and for this purpose we turn to meta-regression



Model		Effect size and 95% confidence interval					Test of null (2-Tail)		Heterogeneity				Tau-squared			
Model	Number Studies	Point estimate	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
Fixed	13	-0.4303	0.0405	0.0016	-0.5097	-0.3509	-10.6247	0.0000	152.2330	12.0000	0.0000	92.1173	0.3088	0.2299	0.0528	0.5557
Random	13	-0.7141	0.1787	0.0319	-1.0644	-0.3638	-3.9952	0.0001								

Figure 25

STEP 3: RUN THE META-REGRESSION

At this point we proceed to the meta-regression.

On the analysis screen, select Analysis > Meta-regression 2

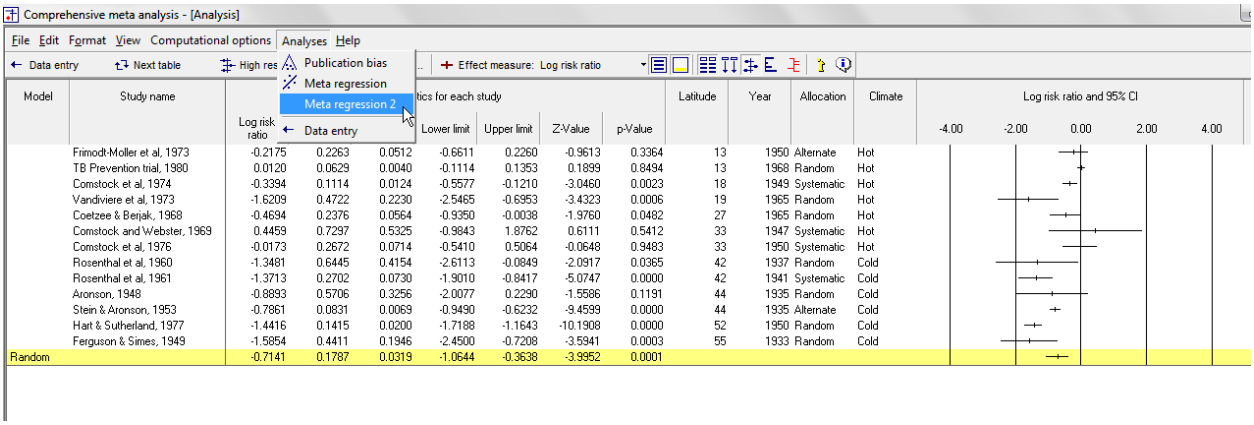


Figure 26

The Interactive Wizard

The program displays the screen shown in Figure 27.

The interactive wizard will walk you through all the steps in running the regression. To display or hide the wizard use the Help menu.

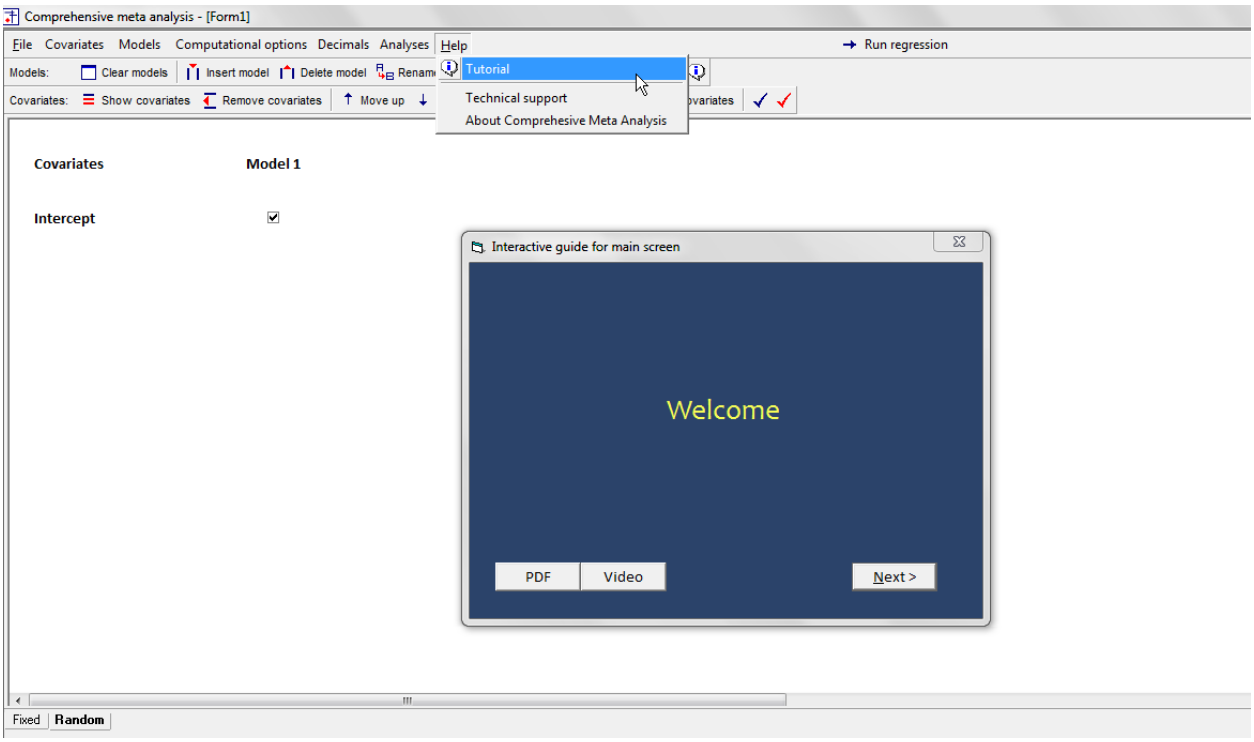


Figure 27

Add covariates to the model

When you initially open the regression module the program displays the following

- The main screen
- A list of available covariates

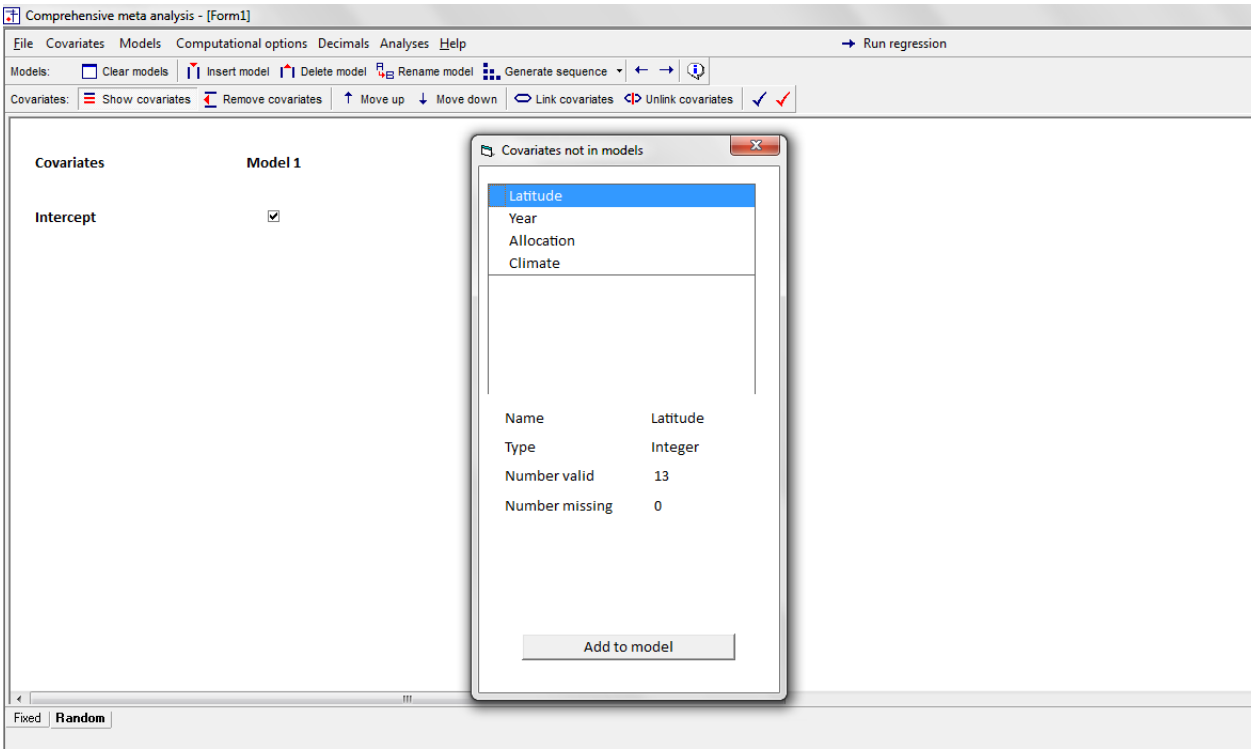


Figure 28

Move the covariates from the wizard onto the main screen.

Add variables in the sequence shown here (allocation, year, latitude) to recreate the example that we use in this text.

- Click “Allocation” on the wizard and then click [Add to model]
- Click “Year” on the wizard and then click [Add to model]
- Click “Latitude” on the wizard and then click [Add to model]

The model is shown in Figure 29

Note that “Allocation” is displayed as two lines, which are linked by a bracket. Since allocation is a categorical variable the program automatically creates and dummy variables to represent allocation. See next chapter for a full discussion.

Tick the check-boxes for all covariates

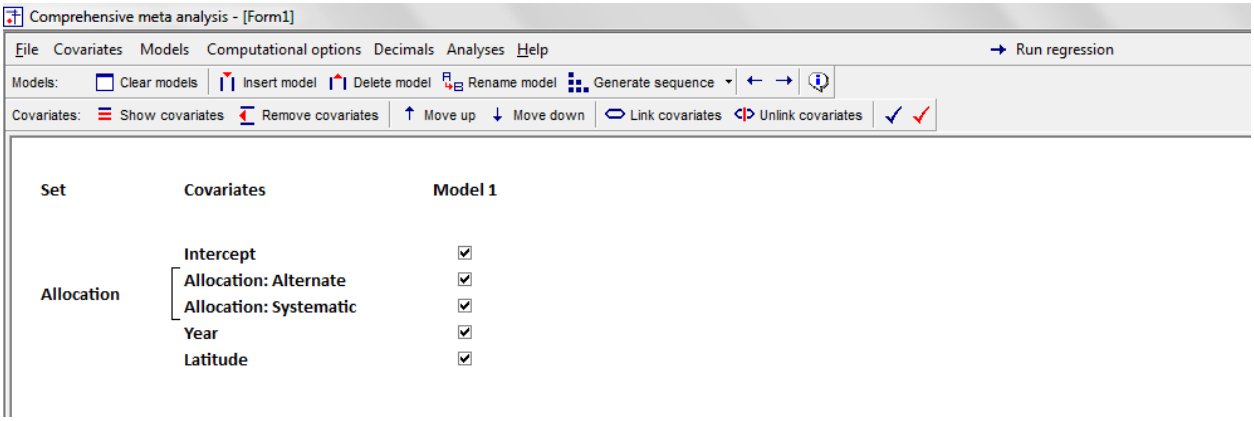


Figure 29

The covariates are controlled by the “Covariates” toolbar

- [Show covariates] shows or hides the wizard
- [Remove covariates] allows you to remove a covariate from the main screen
- [Move up] and [Move down] allow you to edit the sequence of covariates
- The blue and red checks allow you to add (or remove) checks from a series of check-boxes

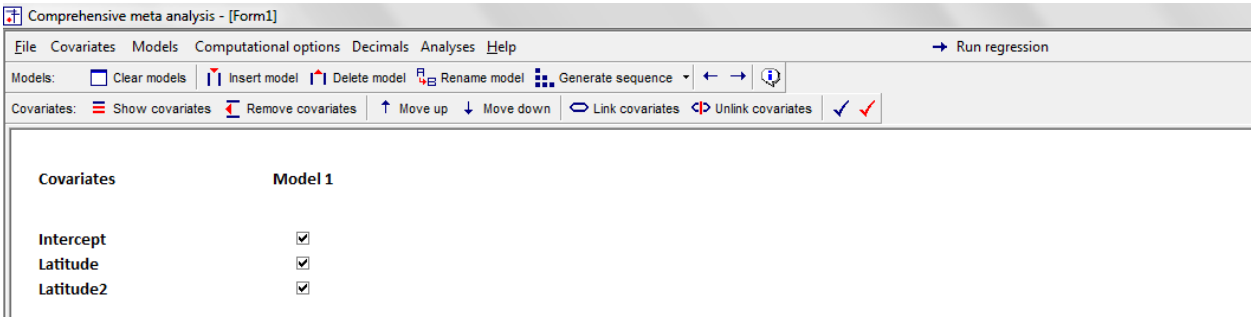
Create a "Set" of covariates (optional)

In regression there are times when we use several covariates to capture a concept. For example

- If we have a categorical covariate with m values, we use (m – 1) covariates to represent this variable in the analysis.
- If we want to assess the relationship between duration of treatment and effect we might include powers of duration such as duration, duration^2, and duration^3 as predictors.
- We may have a series of covariates, such as income and education that, together, are taken to represent the impact of socio-economic status.
- We may have a series of covariates such as dose and duration that, taken together, are intended to represent the intensity of a treatment
- We may have two covariates and also the interaction between, where the three together represent their influence on outcome.

For this example we'll assume that we have an additional covariate, called Latititude2, which is Latitude squared. We want to create a set that incorporates Latitude and Latitude2, and call that set "Latitude Set".

Move Latitude into the model
Move Latitude2 into the model



To create a set covariates

- Ensure that the covariates intended for the set are sequential in the list
- Highlight these covariates [B]

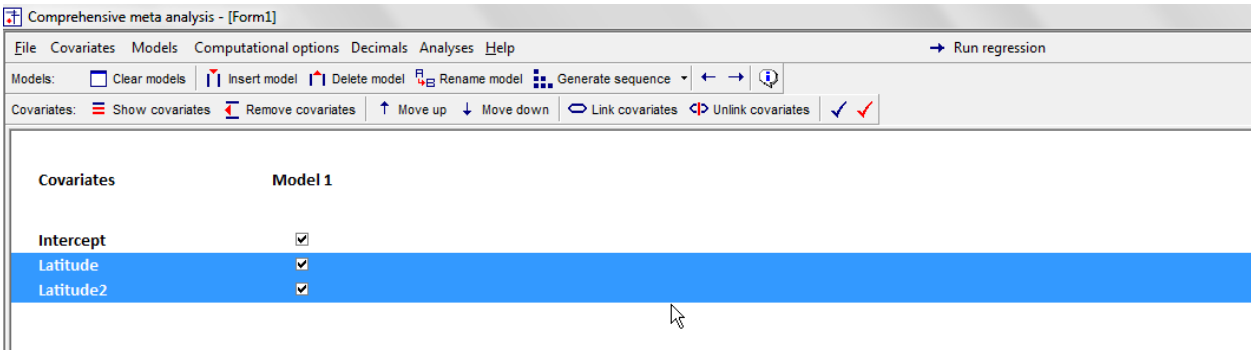


Figure 30

To create the set

- Click [Link Covariates]
- Enter the name Latitude Set and click Ok

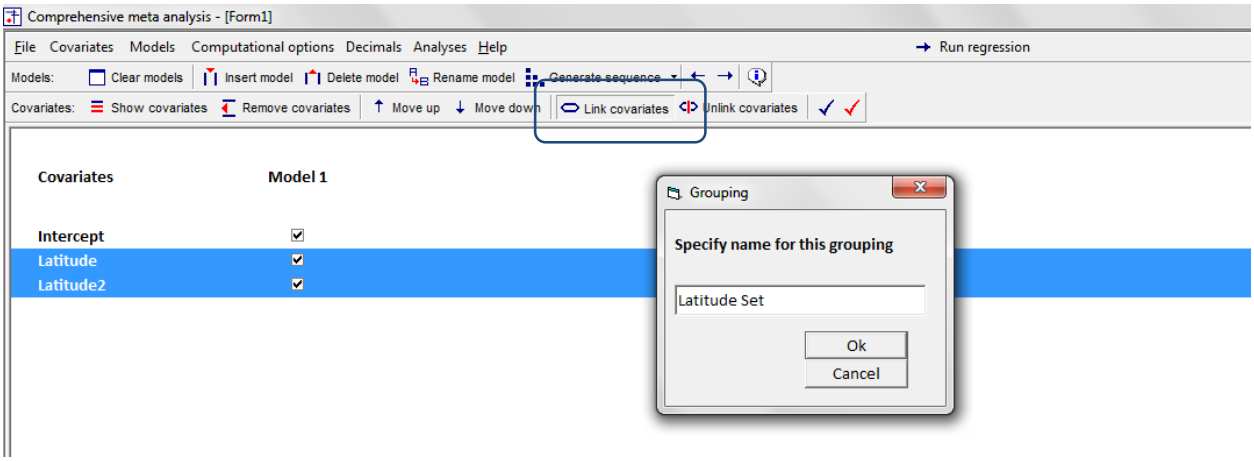


Figure 31

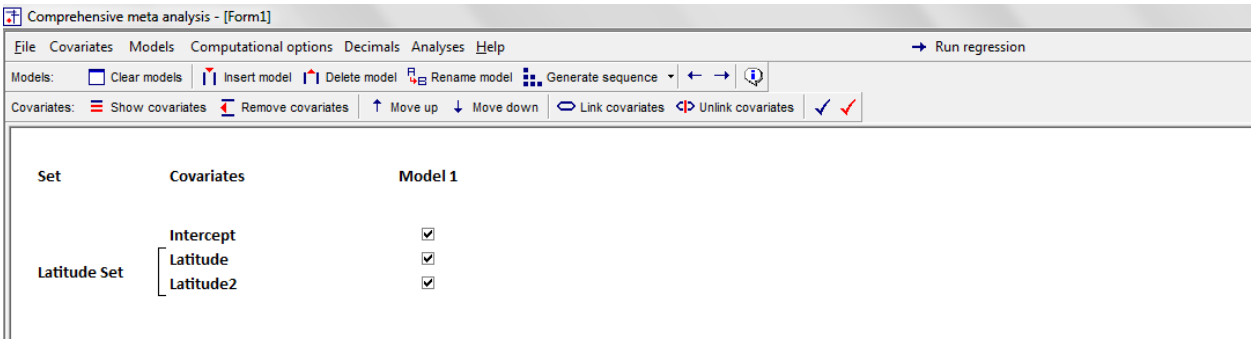
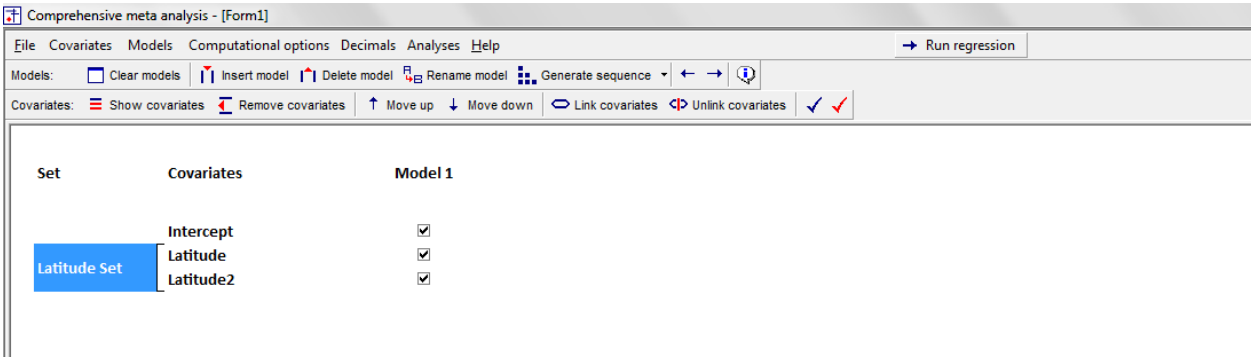


Figure 32

The program has now created a set called “Latitude Set” which includes the two covariates. When you run the regression the program will display statistics for this set.

To remove a set

- Highlight the set's name
- Click Unlink Covariates



Set statistical options

The program allows you to specify various options for the computations

Click Computational options to display this menu

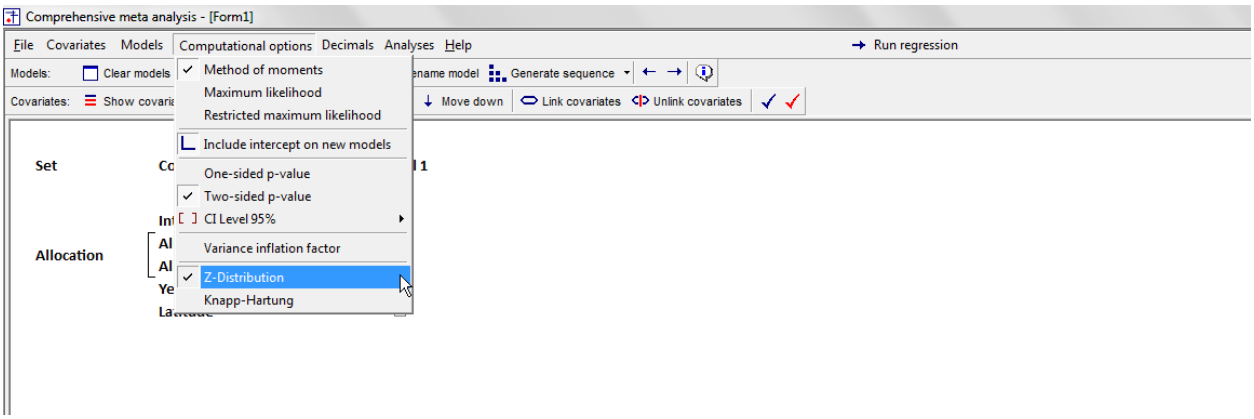


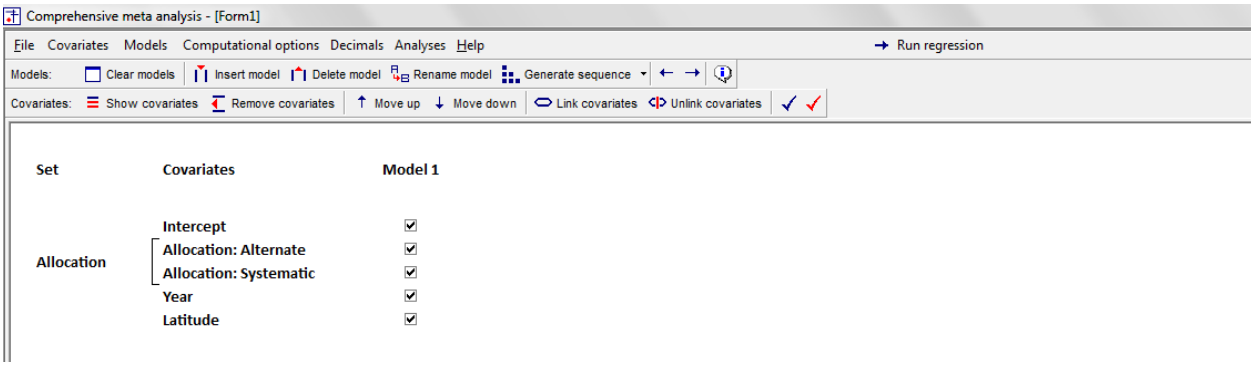
Figure 33

Edit any of the settings, including

- Set the method for estimating T^2
- Use a one-tailed or two-tailed test for *p*-values
- Set the confidence level
- Display the variance inflation factor
- Use the *Z* distribution or the Knapp-Hartung adjustment for p-values and confidence intervals

Run the regression

To run the regression, simply click “Run regression” on the toolbar [A]



STEP 4: NAVIGATE THE RESULTS

Main results screen

After you run the regression

- Click [Main Results] [A]
- Click the desired prediction model [B]
- Click “Fixed” or “Random” at the bottom to select the statistical model.

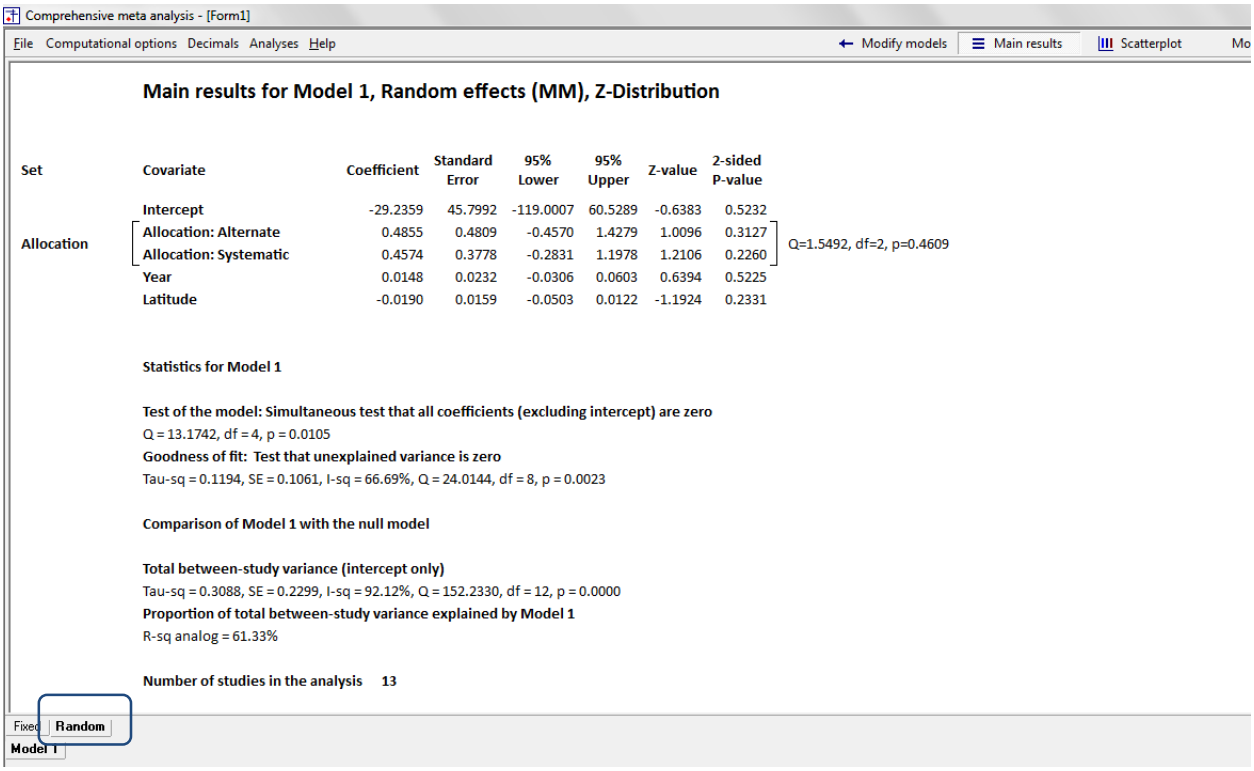


Figure 34

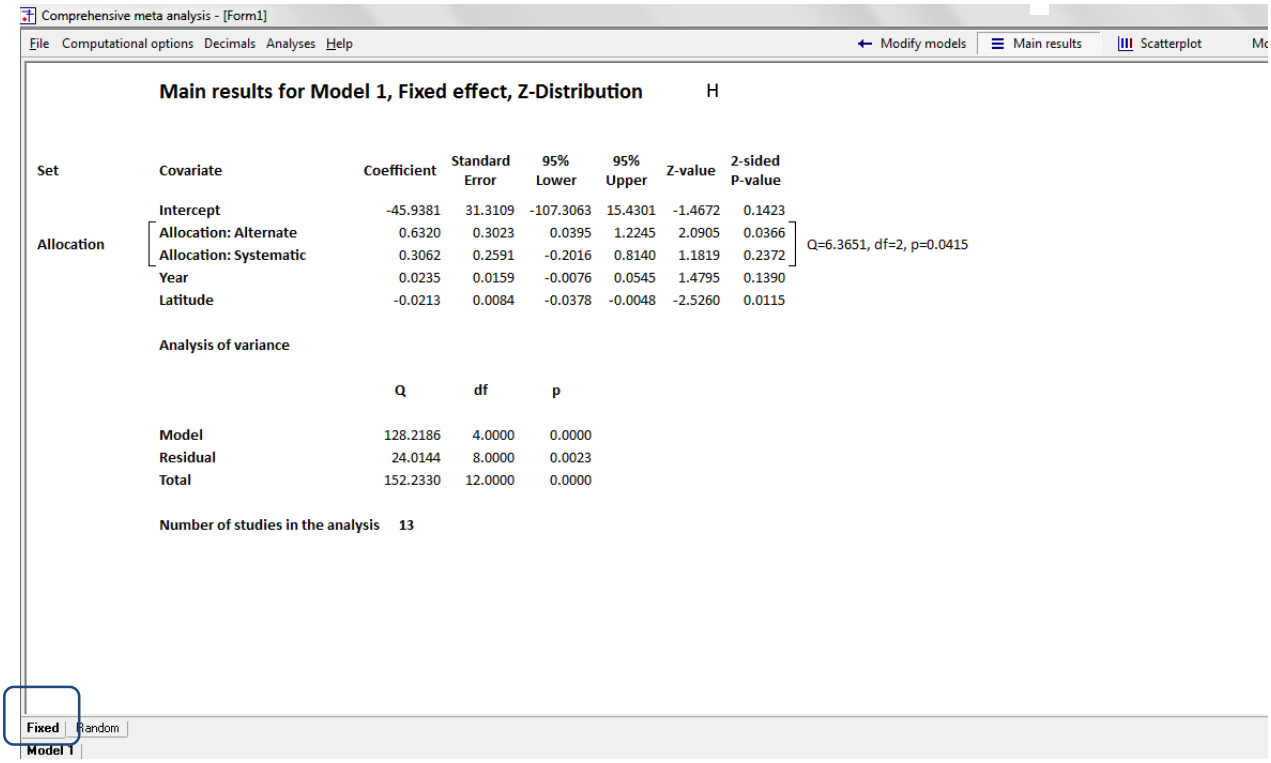
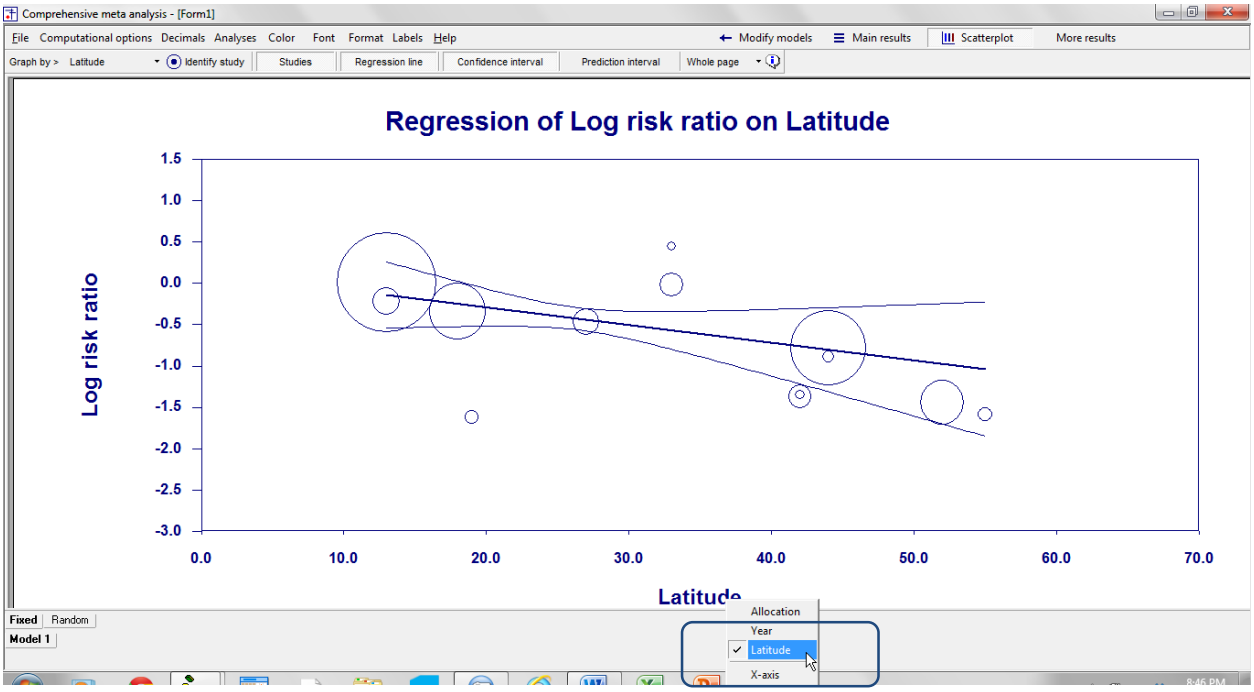

Comprehensive meta analysis - [Form1]
File Computational options Decimals Analyses Help
Modify models
Main results
Scatterplot
Main results for Model 1, Fixed effect, Z-Distribution
H
Set
Covariate
Coefficient
Standard Error
95% Lower
95% Upper
Z-value
2-sided P-value
Intercept -45.9381 31.3109 -107.3063 15.4301 -1.4672 0.1423
Allocation
Allocation: Alternate 0.6320 0.3023 0.0395 1.2245 2.0905 0.0366
Allocation: Systematic 0.3062 0.2591 -0.2016 0.8140 1.1819 0.2372
Q=6.3651, df=2, p=0.0415
Year 0.0235 0.0159 -0.0076 0.0545 1.4795 0.1390
Latitude -0.0213 0.0084 -0.0378 -0.0048 -2.5260 0.0115
Analysis of variance
Q
df
p
Model 128.2186 4.0000 0.0000
Residual 24.0144 8.0000 0.0023
Total 152.2330 12.0000 0.0000
Number of studies in the analysis 13
Fixed
Random
Model 1

Figure 35

Plot

To display the plot

- Click [Scatterplot] on the menu bar to navigate to the plot [A]
- Select “Fixed” or “Random” on the tab [B]
- Select the model (if more than one model had been defined) [C]
- To specify the variable for the X-axis, right-click on the X-axis label [D]



Other screens

To navigate to other tables of results, click “More” and then select any of the following

The following provide additional information about the analysis for the predictive model that is selected at the bottom left of the screen

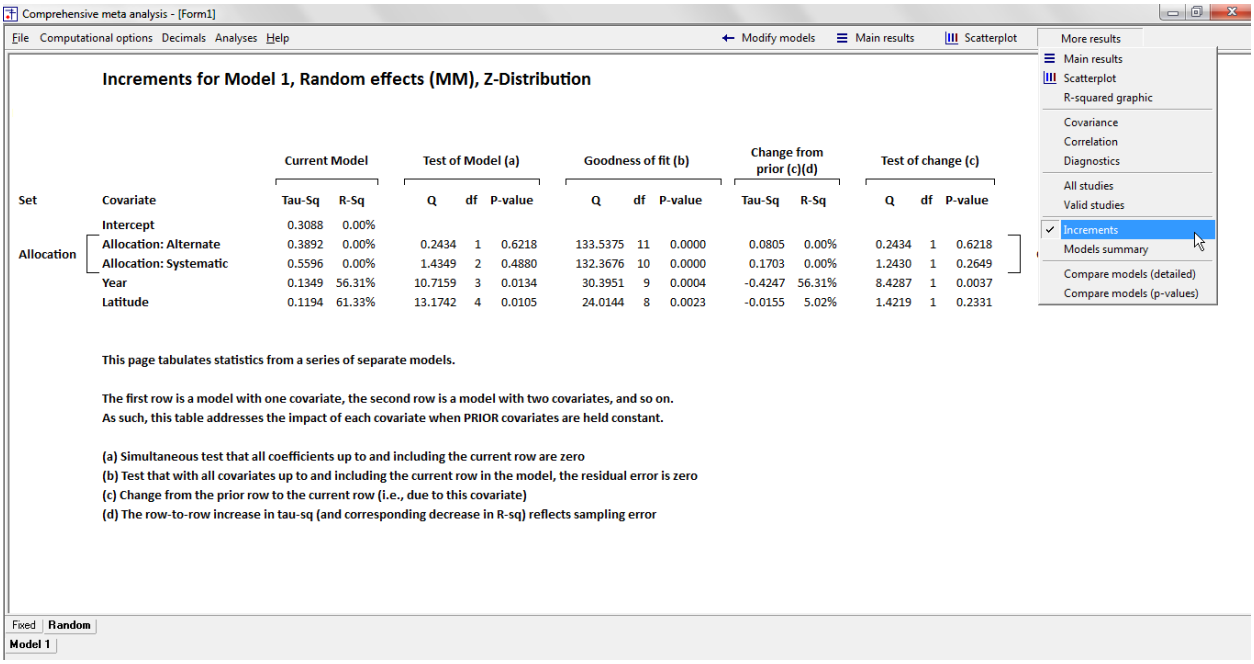
- R-squared graphic
- Covariance
- Correlation
- Diagnostics

The following provide information about the data included in (or excluded from) the analysis

- All studies
- Valid studies

The following provide statistics for and/or compare different models

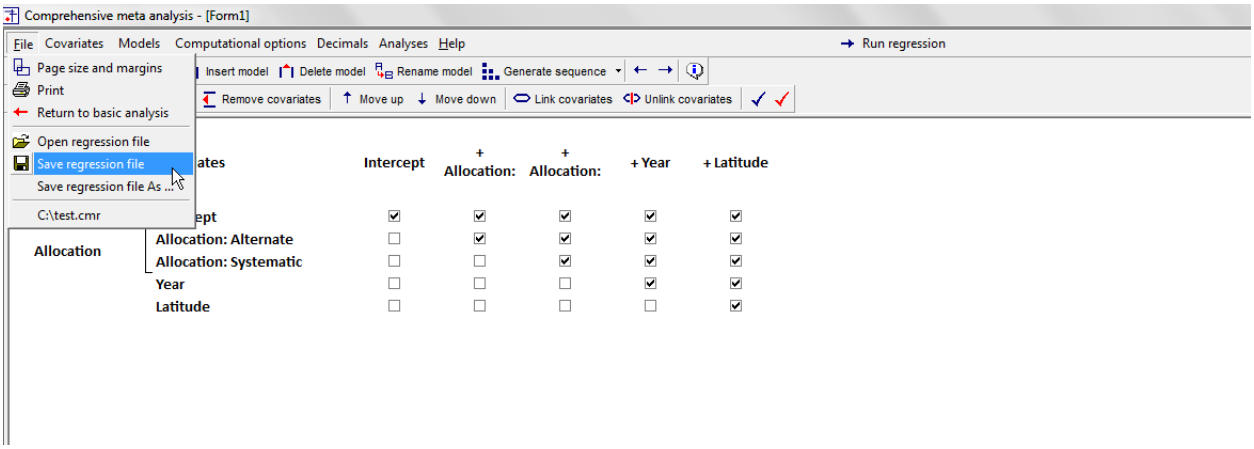
- Increments
- Models summary
- Compare models (detailed)
- Compare models (p-value)



STEP 4: SAVE THE ANALYSIS

Once you've created a meta-regression you can save it using

- File > Save regression file as …
- This will save the regression template with an extension of .cmr.



The .cmr file save the instructions for the analysis, NOT the data. By analogy, programs such as SPSS™, SAS™, and stata™ allow you to save a set of commands in one file and the data in another file. The commands can then be applied to any data file that has the same variables.

The .cmr file saves the following

- The list of covariates
- The list of models
- The check-boxes for each model
- The sets
- The model names

In another session you can return to the regression module and click

- File > Open file

to open this file, and re-run the analysis. This can be with the same dataset as you had used before, or with another dataset. For example, you may return to the data-entry screen and add new studies, or you may return to the main analysis screen and edit the filters, or you may be working with an entirely different data set that has the same variables as the first one. When you open a cmr file the program simply recreates the main MR screen as though you had entered it manually.

The MR file does *not* save the statistical settings that were in place when the file was created. These include the method employed to estimate T^2, the use of Z or Knapp-Hartung, the confidence level, the choice of a one-sided or two-sided test.

Step 5: Export the results

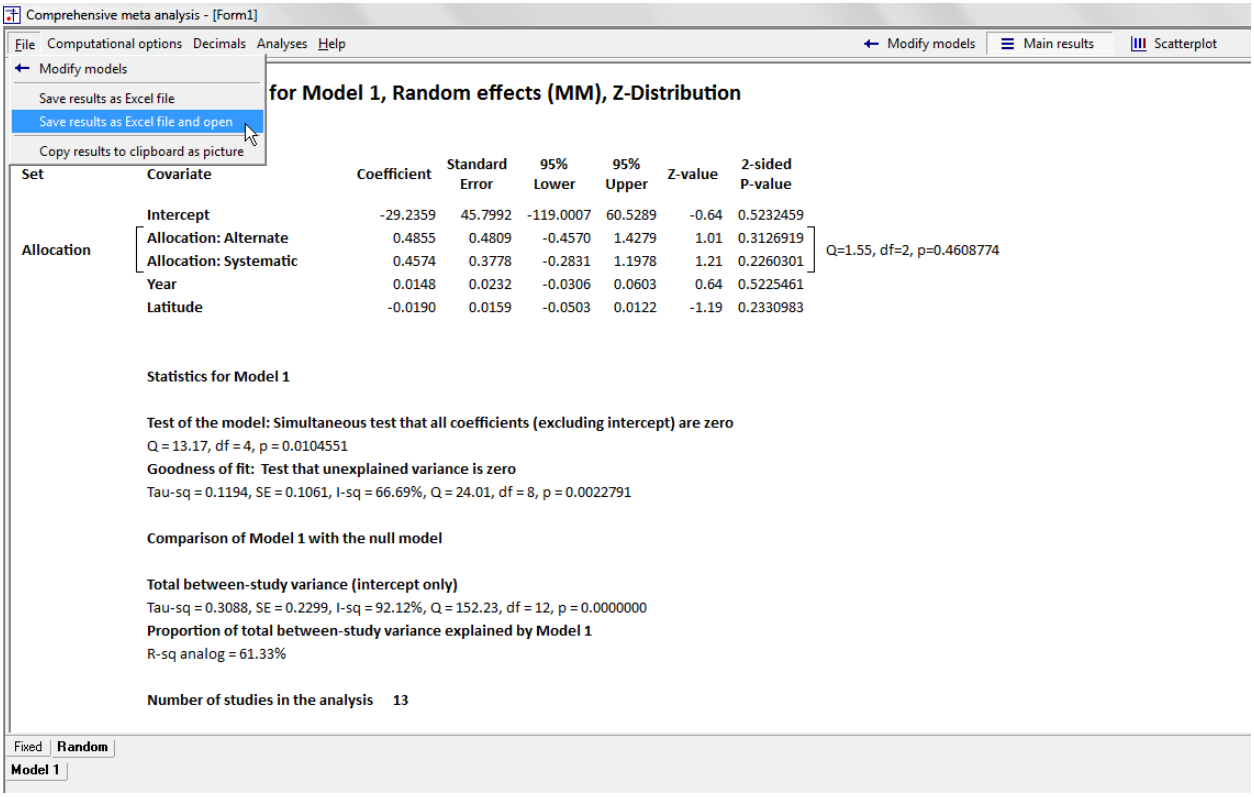
The program offers two options for exporting the results of any analysis.

- Export the results to Excel. Then, you can perform additional computations within Excel, and/or format the results and copy them as a table to other programs

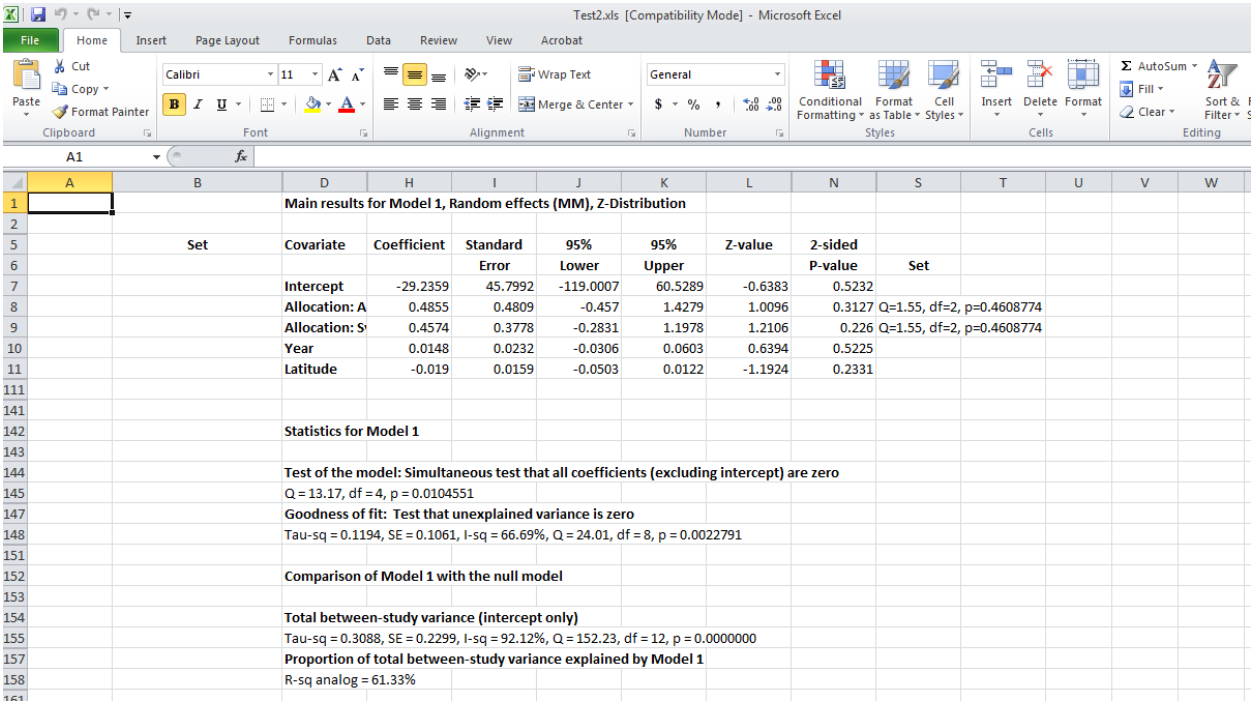
- Copy the results to the clipboard as a picture. Then, paste this picture into Word or any other program.

The example here is for the main analysis screen.

The screen looks like this



- Click File > Export results to Excel and open
- Provide a name for the Excel file



	A	B	D	H	I	J	K	L	N	S	T	U	V	W
1			**Main results for Model 1, Random effects (MM), Z-Distribution**											
2														
5		**Set**	**Covariate**	**Coefficient**	**Standard**	**95%**	**95%**	**Z-value**	**2-sided**					
6					**Error**	**Lower**	**Upper**		**P-value**	**Set**				
7			**Intercept**	-29.2359	45.7992	-119.0007	60.5289	-0.6383	0.5232					
8			**Allocation: A**	0.4855	0.4809	-0.457	1.4279	1.0096	0.3127	Q=1.55, df=2, p=0.4608774				
9			**Allocation: S**	0.4574	0.3778	-0.2831	1.1978	1.2106	0.226	Q=1.55, df=2, p=0.4608774				
10			**Year**	0.0148	0.0232	-0.0306	0.0603	0.6394	0.5225					
11			**Latitude**	-0.019	0.0159	-0.0503	0.0122	-1.1924	0.2331					
111														
141														
142			**Statistics for Model 1**											
143														
144			**Test of the model: Simultaneous test that all coefficients (excluding intercept) are zero**											
145			Q = 13.17, df = 4, p = 0.0104551											
147			**Goodness of fit: Test that unexplained variance is zero**											
148			Tau-sq = 0.1194, SE = 0.1061, I-sq = 66.69%, Q = 24.01, df = 8, p = 0.0022791											
151														
152			**Comparison of Model 1 with the null model**											
153														
154			**Total between-study variance (intercept only)**											
155			Tau-sq = 0.3088, SE = 0.2299, I-sq = 92.12%, Q = 152.23, df = 12, p = 0.0000000											
157			**Proportion of total between-study variance explained by Model 1**											
158			R-sq analog = 61.33%											

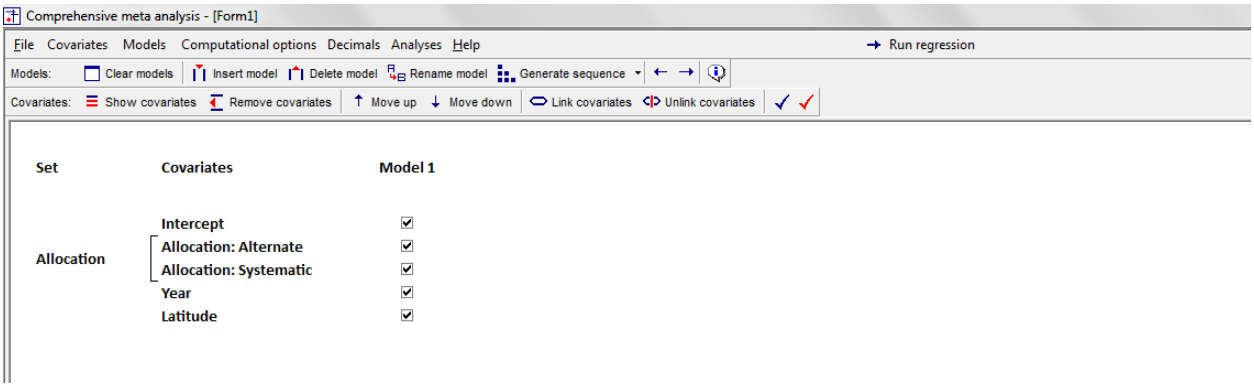
Part 5: Understanding the results

MAIN RESULTS

FIXED-EFFECT ANALYSIS

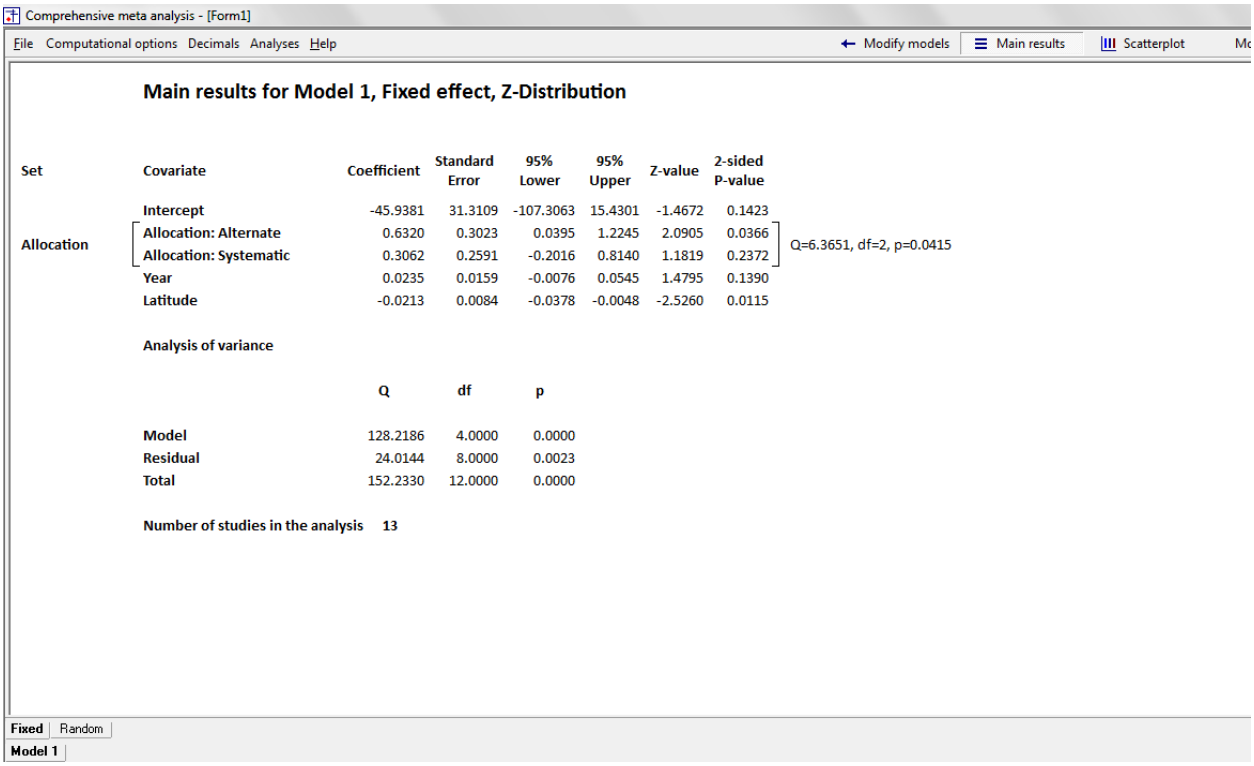
To navigate to this screen

Run the analysis [A]



The toolbar changes as shown here

- Click “Main results”
- Click on the predictive model
- Click “Fixed”



Impact of the full model

Analysis of variance

In this section the total WSS is partitioned into its component parts –

- The WSS *explained* by the covariates (the model)
- The WSS *not explained* by the covariates (the residual).

Model

This is the test that the predictive model explains *any* of the variance in effect size. Put another way, it asks if the dispersion of effects about the regression line smaller when the regression line is based on the covariates rather than based solely on the grand mean. Here, $Q = 128.2186$ with $df = 4$ and $p < 0.0001$, so we conclude that the predictive model explains (at least) some of the variance in effect size.

Residual

This is the test that the data are consistent with the model's assumption of a common effect size for all studies with the same predicted value. The Q value is 24.0144 with $df = 8$ and $p = 0.0023$. We conclude that the data are not consistent with the assumptions of the fixed-effect model.

Total

This is the test that the variance for the full set of studies (with no predictors) is zero. The Q-value is 152.2330 with $df = 12$ and $p < 0.0001$.

Impact of individual covariates

The test of the model is an omnibus test for the full set of covariates. It tells us that the set as a whole is related to effect size. By contrast, the table at the top addresses the impact of each covariate. *In this table, the impact of each covariate is reported with all of the other covariates partialled* out *(or held constant)*.

Since the effect size is the risk ratio, all analyses are carried out in log metric and all coefficients are in the log metric. In this example, virtually all predicted effects are less than zero, so 0 is no effect, -1 is a large effect, and -2 is a very large effect. In this example, therefore, a negative coefficients means that as the covariate gets larger the vaccine is more effective. (The reverse would be true if the predicted values were all positive).

The coefficient for Year is 0.0235, which means that for every increase of one year the log risk ratio will increase by 0.0235 (the vaccine became less effective over time). The corresponding *p*-value is 0.1390.

The coefficient for latitude [I]is −0.0213, which means that for every increase of one unit (degree) in latitude the log risk ratio will decrease by 0.0213 (vaccine is more effective at greater latitudes). The coefficient plus/minus 1.96 times the standard error (0.0084) yields the 95% confidence interval for the coefficient, which is −0.0378 to −0.0048. The coefficient divided by its standard error yields a *Z* value of -2.526, and the corresponding *p*-value of 0.0115. Thus, when year and allocation method are held constant, the relationship between latitude and effect size is statistically significant.

Impact of a set of covariates

The model includes two covariates that have been identified as a set. These are Alternate allocation and Systematic allocation. The test of the set tells us if allocation is related to effect size. In this case Q = 6.3651 with df = 2 and p = 0.0412, and so there is evidence that effect size is related to allocation type. For a more specific analysis we can look at each line within the set, and see that Alternate allocation has a coefficient of 0.6320 (the vaccine is less effective in studies that employed alternate allocation) and a *p*-value of 0.0366. However, as discussed in the chapter on caveats, this finding is almost certainly due to a confound with other factors.

Summary

The total Q of each effect size about the grand mean can be partitioned into its component parts – the Q due to the variation in effect size that can be explained by the covariates, and the part that cannot.

- Model. The Q-value for the model is 128.2186 with df = 4 and $p < 0.0001$, which tells us that effect size is related to at least one of the covariates.
- Residual. The Q-value for the residual is 24.0144 with df = 8 and p = 0.0023, which tells us that the assumptions of the fixed-effect model have been violated.
- Total. The Q-value for the total is 152.23 with df = 12 and $p < 0.0001$, which tells us that that effect sizes vary when we ignore subgroups and work with deviations of all studies from the grand mean.

Impact of individual covariates

The test of the model is an omnibus test for the full set of covariates. It tells us that at least one of the covariates is related to effect size. By contrast, the table at the top addresses the impact of each covariate. *In this table, the impact of each covariate is reported with all of the other covariates partialled out (or held constant)*.

The p-values tell us if there is evidence that the covariate is related to effect size when the other covariates are held constant.

- The *p*-value for allocation is 0.0415. Specifically, with other covariates held constant alternate allocation is associated with a smaller effect size (but see the chapter on caveats).

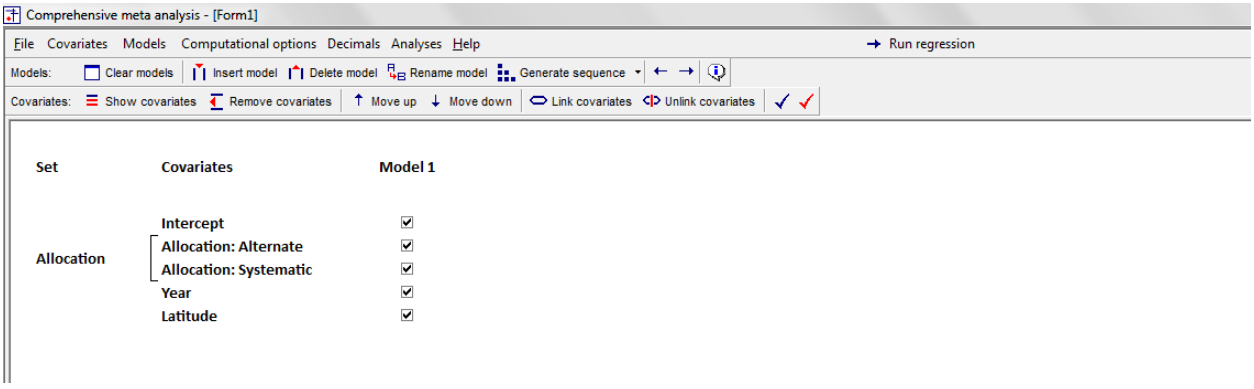
- The *p*-value for year is 0.1390, with the studies that fall further from the equator showing more impact of the vaccine.

- The *p*-value for latitude is 0.0115, with the studies that fall further from the equator showing more impact of the vaccine.

RANDOM-EFFECTS ANALYSIS

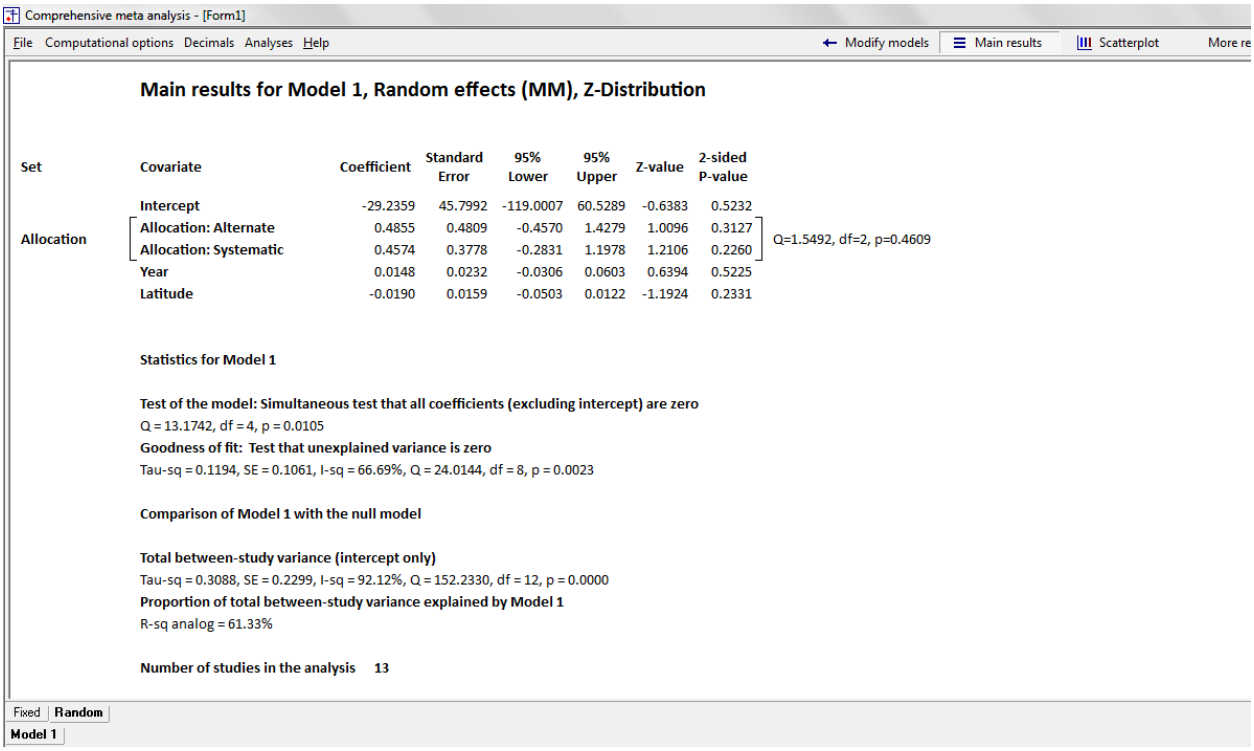
To navigate to this screen

Run the analysis



The toolbar changes as shown here

- Click “Main results”
- Click on the predictive model
- Click “Random”



Test of the model

Is effect size related to the covariates?

The test of the model is a simultaneous test that all covariates (except the intercept) are zero. The *Q*-value is 13.1742 with *df* = 4 and *p* = 0.0105. We reject the null and conclude that at least one of the covariates is related to effect size.

Goodness of fit

Is there any unexplained variance in the true effect sizes?

Immediately above, we saw that the covariates improve our ability to predict that study's effect. But does this information enable us to completely predict that study's effect – do all studies with the same values on all covariates share a common effect size? Or is there variance in true effects among studies with the same predicted value?

The *Q* statistic, based on the deviation of each study from its predicted value, is 24.0144, with 8 *df* and the corresponding *p*-value is 0.0023. This tells us that the true effect size varies from study to study, even for studies that are identical on all covariates. Put another way, the model is incomplete – knowing a study's allocation type, year, and latitude does not allow us to completely predict its effect size.

How much variance is there?

The program shows that T^2, the variance of true effect sizes at any point on the regression line, is 0.1194. It follows that the *T*, the standard deviation of true effect sizes at any point on the regression line is 0.3455. We can use this to get a sense of how closely the true effects at any point on the regression line are (or are not) clustered together.

In Figure 26 we've plotted all 13 studies, the regression line, and a series of normal curves about the regression line. Each normal curve has a height of 1.96 *T* and is centered about some point on the regression line. If the true effects are normally distributed with standard deviation *T*, then 95% of studies with that predicted value will have a true effect size within the range of the normal curve.

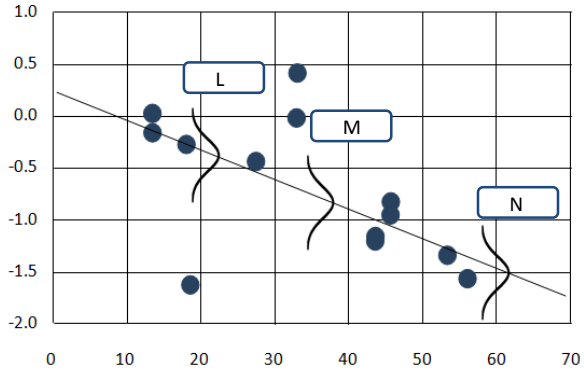


Figure 36

What proportion of the observed variance is true variance?

The variance that cannot be explained by the covariates includes within-study variance (essentially error) and between-study variance (that can be potentially explained by additional study-level covariates). The I^2 statistic is 66.69%, which tells us that 67% of the remaining variance falls into the latter group.

Impact of individual covariates

The test of the model is an omnibus test for the full set of covariates. It tells us that at least one of the covariates is related to effect size. By contrast, the table at the top addresses the impact of each covariate. *In this table, the impact of each covariate is reported with all of the other covariates partialled* out *(or held constant)*.

Since the effect size is the risk ratio, all analyses are carried out in log metric and all coefficients are in the log metric. In this example, virtually all predicted effects are less than zero, so 0 is no effect, -1 is a

large effect, and -2 is a very large effect. In this example, therefore, a negative coefficients means that as the covariate gets larger the vaccine is more effective. (The reverse would be true if the predicted values were all positive).

The model includes two covariates that have been identified as a set. These are Alternate allocation and Systematic allocation [G]. The test of the set tells us if allocation is related to effect size. In this case Q = 1.5402 with df = 2 and p = 0.46, and so there is no evidence that effect size is related to allocation type.

The coefficient for Year [H] is 0.0148, which means that for every increase of one year the log risk ratio will increase by 0.0148 (the vaccine became less effective over time). The corresponding p-value is 0.5225.

The coefficient for latitude [I] is −0.0190, which means that for every increase of one unit (degree) in latitude the log risk ratio will decrease by 0.0190 (vaccine is more effective at greater latitudes). The coefficient plus/minus 1.96 times the standard error (0.0159) yields the 95% confidence interval for the coefficient, which is −0.0503 to 0.0122. The coefficient divided by its standard error yields a Z value of -1.1924, and the corresponding p-value of 0.23. Thus, when year and allocation method are held constant, the relationship between latitude and effect size not is statistically significant.

In this example none of the individual covariates has a p-value less than 0.05. Since the model as a whole is statistically significant, the fact that no covariate is statistically significant probably reflects the fact that some of the covariates are correlated with each other. For example, latitude or year might be statistically significant if entered into the equation alone. However, if the two are correlated with each other and compete for the same variance, neither meets the threshold for statistical significance.

Comparison of Model 1 with the null model

The intent of this display is to report how much variance there is initially (without covariates), and how much variance remains (with covariates). Then, by comparing the two, we can report that the covariates explained some proportion of the initial variance.

Total between-study variance (intercept only)

To get the initial amount of variance we run a regression with no covariates and compute T^2. Here, T^2 is 0.3088, which is the variance of all studies about the grand mean.

Proportion of variance explained

To get the final amount of variance we run a regression with the covariates and compute T^2. This value, reported above as 0.1194, is the variance of studies about their predicted value.

If the initial T^2 is 0.3088 and the remaining T^2 is 0.1194, the difference (0.1894) is the T^2 explained by the model. Then we can compute R^2, the proportion explained by the model, as 0.6133.

The proportion of variance explained is called R^2. To compute R^2 we follow the logic in the preceding paragraph. We compute T^2 with no covariates [J] to provide the value at left. We already have T^2 with covariates [E] to provide the value at right. Then R^2 is computed as (0.3088 – 0.1194)/0.3088 = 0.6133.

We can show this graphically. At left, the normal curve reflects the unexplained variance in effects when the predicted value for each study is the grand mean. At right, the normal curves represent the unexplained variance in effects when the predicted value for each study is based on the regression line. The variance at the right is less than the variance at the left, which tells us that by using these covariates we can reduce the unexplained variance – or (equivalently) explain some of the variance.

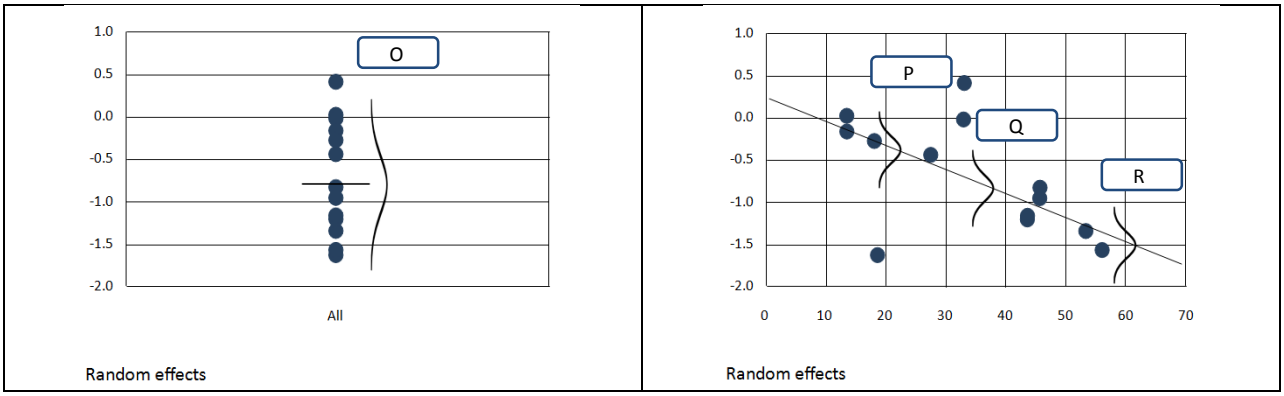


Figure 37

Summary

- The *Q*-Model is 18.85 with *df* = 1 and $p < 0.0001$. This tells us that effect size is related to latitude.

- The *Q*-value for goodness of fit is 30.73 with *df* = 12 and $p = 0.0012$. This tells us that the effect size varies, even within studies at the same latitude.

- The *Q*-total is 152.23 with *df* = 12 and $p < 0.0001$. This tells us that effect sizes vary when we ignore latitude and work with deviations of all studies from the grand mean.

- The *observed* variance in effect sizes is partly due to real differences and partly due to within-study sampling error. When there are no covariates [D] the I^2 value is 92%, which tells us that 92% of the observed variance is real, and may potentially be explained by covariates. When we use these covariates [J] the I^2 value is 66.69%, which tells us that 66.69% of the remaining variance is real, and may potentially be explained by additional covariates.

- The between-study variance is estimated at 0.1194 at any given point on the regression line based on these covariates, as compared to 0.3088 for the regression line based on the grand mean. This corresponds to an R^2 of 61.33%, meaning that some 61% of the true variance in effects can be explained by the covariates.