Example name	Second-hand smoke and lung-cancer
Effect size	Risk ratio
Analysis type	Publication bias
Level	Advanced

Synopsis

Studies compared the risk of lung cancer in non-smokers whose spouse was (or was not) a smoker. The pooled effect showed an increase risk of lung-cancer for the second group, but there was concern that the estimate could have been affected by publication bias.

We use this example to show

- The distinction between a small-study effect and publication bias
- How to use the Fail-Safe N (and why not to)
- How to use the Egger test
- How to use Trim and Fill
- How to use a cumulative analysis to assess the potential impact of bias

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

The accumulated evidence on lung cancer and environmental tobacco smoke

A K Hackshaw, M R Law, N J Wald

See editorial by Davis and p 973

Department of Environmental and Preventive Medicine, Wolfson Institute of Preventive Medicine, St Bartholomew's and Royal London School of Medicine and Dentistry, London EC1M 6BQ

A K Hackshaw, *lecturer*

M R Law, *reader*

N J Wald, *professor*

Correspondence to: Professor Wald njwald@mds.qmw.ac.uk

BMJ 1997;315:980–88

Abstract

Objective: To estimate the risk of lung cancer in lifelong non-smokers exposed to environmental tobacco smoke.

Design: Analysis of 37 published epidemiological studies of the risk of lung cancer (4626 cases) in non-smokers who did and did not live with a smoker. The risk estimate was compared with that from linear extrapolation of the risk in smokers using seven studies of biochemical markers of tobacco smoke intake.

Main outcome measure: Relative risk of lung cancer in lifelong non-smokers according to whether the spouse currently smoked or had never smoked.

Results: The excess risk of lung cancer was 24% (95% confidence interval 13% to 36%) in non-smokers who lived with a smoker ($P < 0.001$). Adjustment for the effects of bias (positive and negative) and dietary confounding had little overall effect; the adjusted excess risk was 26% (7% to 47%). The dose-response

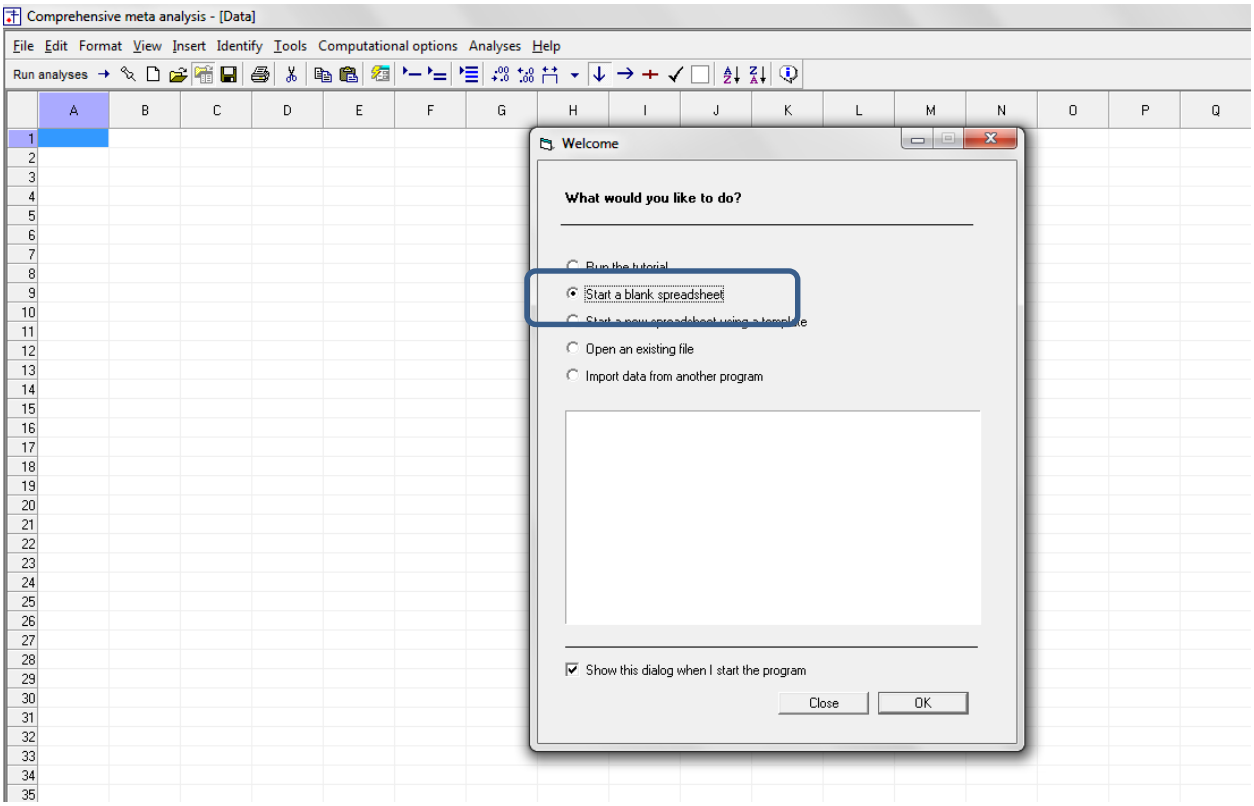
Introduction

Ten years ago scientific committees and national organisations concluded that exposure to environmental tobacco smoke (also called passive smoking) is a cause of lung cancer.[1-4] Substantial additional evidence has since been published, and we report a new analysis. The additional data permit a more precise estimate of the size of the association, with a further assessment of whether it is cause and effect by seeking a dose-response relation and examining whether sources of bias and confounding could account for the association. We also compared the direct estimate of risk from epidemiological studies with that from a low dose linear extrapolation of the risk in smokers using biochemical markers of exposure to tobacco smoke.

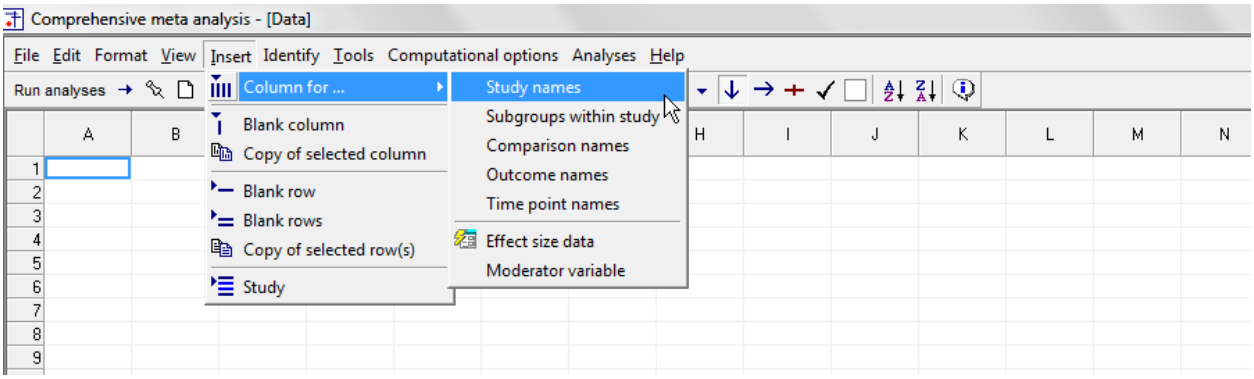
As before,[5] the estimate of effect was the relative risk of lung cancer in lifelong non-smokers according to whether the spouse currently smoked or had never smoked. Spousal exposure is the best available measure: it is well defined and has been validated using biochemical markers.[6-9] It reflects exposure in general

Start the program

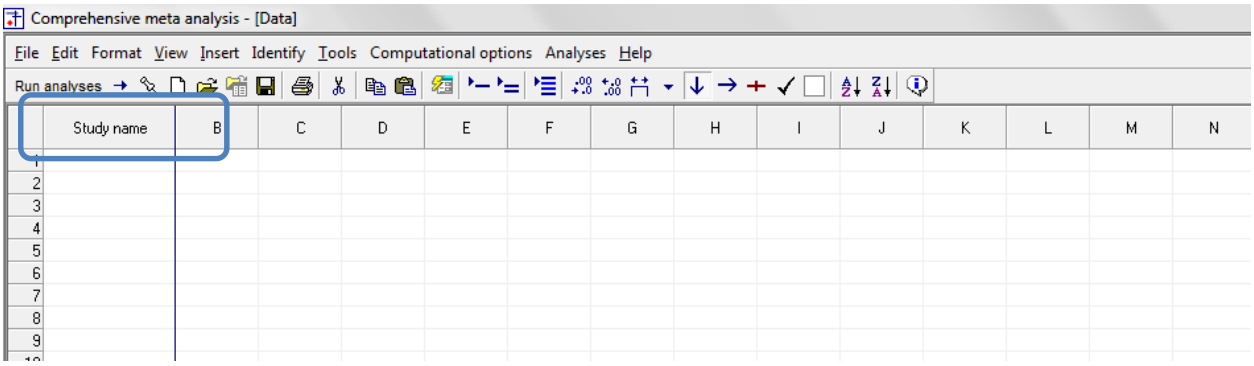
- Select the option [Start a blank spreadsheet]
- Click [Ok]



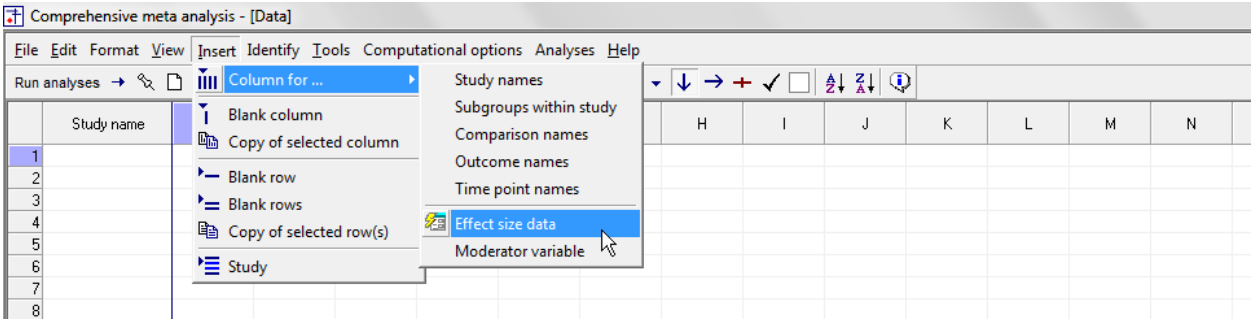
Click Insert > Column for > Study names



The screen should look like this

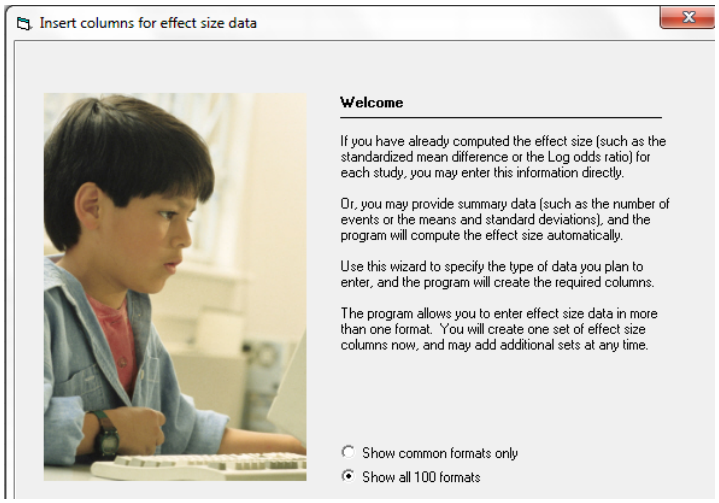


Click Insert > Column for > Effect size data

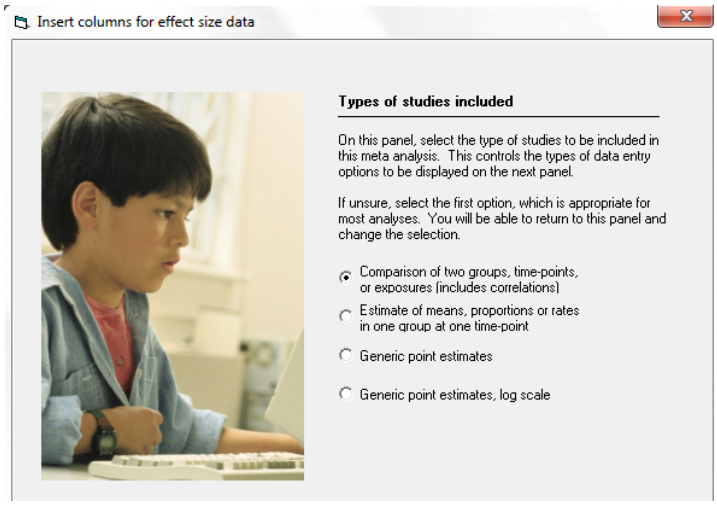


The program displays this wizard

Select [Show all 100 formats]
Click [Next]

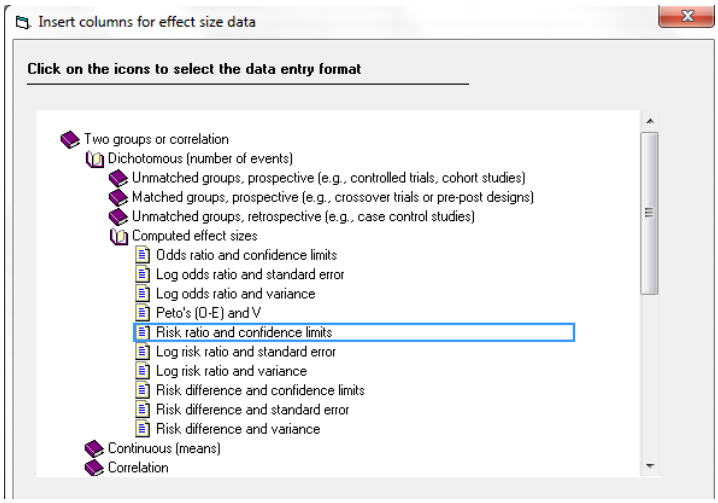


Select [Comparison of two groups…]
Click [Next]

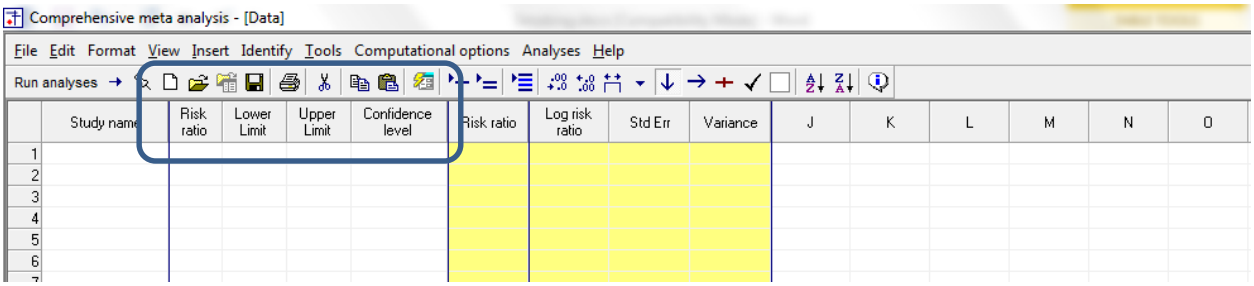


Drill down to

Dichotomous (number of events)
Computed effect sizes
Risk ratio and confidence limits



The screen should look like this



There are three options at this point

- Enter the data directly into CMA
- – or – Open the CMA data file “Smoking.cma”
- – or – Copy the data from Excel “Smoking..xls”

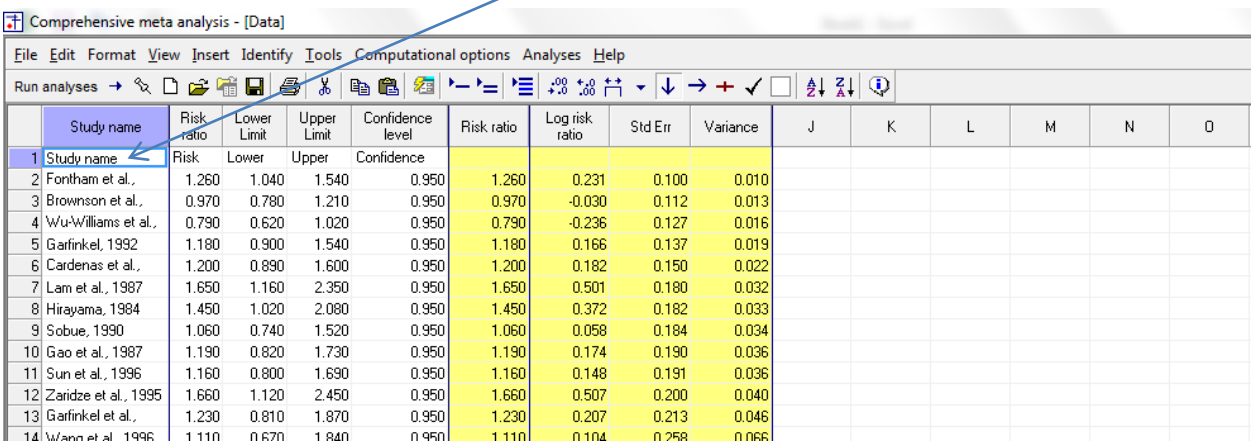
Rather than enter the data directly into CMA we will copy the data from Excel

- Switch to Excel and open the file
- Highlight the rows and columns as shown, and press CTRL-C to copy to clipboard

	A	B	C	D	E	F
1	Study name	Risk ratio	Lower Limit	Upper Limit	Confidence level	
2	Fontham et al., 1994	1.26	1.04	1.54	0.95	
3	Brownson et al., 1992	0.97	0.78	1.21	0.95	
4	Wu-Williams et al., 1990	0.79	0.62	1.02	0.95	
5	Garfinkel, 1992	1.18	0.9	1.54	0.95	
6	Cardenas et al., 1997	1.2	0.89	1.6	0.95	
7	Lam et al., 1987	1.65	1.16	2.35	0.95	
8	Hirayama, 1984	1.45	1.02	2.08	0.95	
9	Sobue, 1990	1.06	0.74	1.52	0.95	
10	Gao et al., 1987	1.19	0.82	1.73	0.95	
11	Sun et al., 1996	1.16	0.8	1.69	0.95	
12	Zaridze et al., 1995	1.66	1.12	2.45	0.95	
13	Garfinkel et al., 1985	1.23	0.81	1.87	0.95	
14	Wang et al., 1996	1.11	0.67	1.84	0.95	
15	Shimizu et al., 1988	1.08	0.64	1.82	0.95	
16	Pershagen et al., 1987	1.03	0.61	1.74	0.95	
17	Koo et al., 1987	1.55	0.9	2.67	0.95	
18	Akiba et al, 1986	1.52	0.87	2.63	0.95	
19	Chan et al, 1982	0.75	0.43	1.3	0.95	
20	Kabat et al., 1995	1.1	0.62	1.96	0.95	
21	Trichopoulos, 1983	2.13	1.19	3.83	0.95	
22	Du et al., 1993	1.19	0.66	2.13	0.95	
23	Kalandidi et al., 1990	1.62	0.9	2.91	0.95	
24	Lam, 1985	2.01	1.09	3.72	0.95	
25	Stockwell et al., 1992	1.6	0.8	3	0.95	
26	Geng et al., 1988	2.16	1.08	4.29	0.95	
27	Liu et al., 1993	1.66	0.73	3.78	0.95	
28	Liu et al., 1991	0.74	0.32	1.69	0.95	
29	Buffler, 1984	0.8	0.34	1.9	0.95	
30	Lee et al., 1986	1.03	0.41	2.55	0.95	
31	Correa, 1983	2.07	0.81	5.25	0.95	
32	Wu et al, 1985	1.2	0.5	3.3	0.95	
33	Humble et al., 1987	2.34	0.81	6.75	0.95	
34	Jockel, 1991	2.27	0.75	6.82	0.95	
35	Kabat, 1984	0.79	0.25	2.45	0.95	
36	Inoue et al., 1988	2.55	0.74	8.78	0.95	
37	Brownson et al., 1987	1.52	0.39	5.96	0.95	
38	Butler, 1988	2.02	0.48	8.56	0.95	

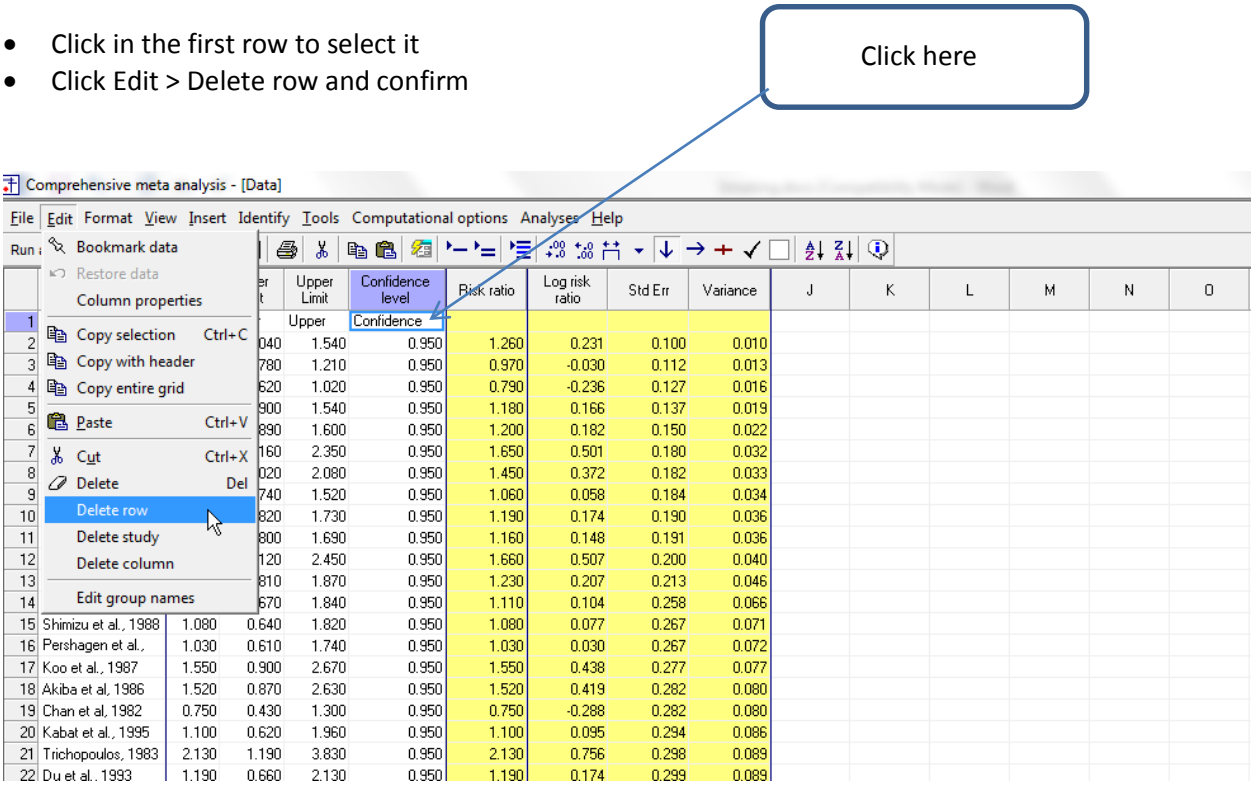
Switch back to CMA

- Click in Cell Study name – 1
- Press [CTRL-V] to paste the data into CMA

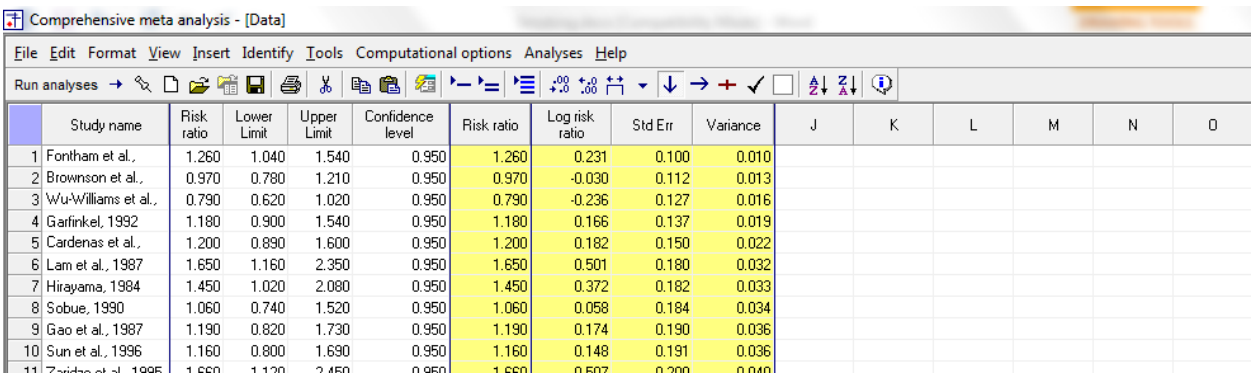


	Study name	Risk ratio	Lower Limit	Upper Limit	Confidence level	Risk ratio	Log risk ratio	Std Err	Variance	J	K	L	M	N	O
1	Study name	Risk	Lower	Upper	Confidence										
2	Fontham et al.,	1.260	1.040	1.540	0.950	1.260	0.231	0.100	0.010						
3	Brownson et al.,	0.970	0.780	1.210	0.950	0.970	-0.030	0.112	0.013						
4	Wu-Williams et al.,	0.790	0.620	1.020	0.950	0.790	-0.236	0.127	0.016						
5	Garfinkel, 1992	1.180	0.900	1.540	0.950	1.180	0.166	0.137	0.019						
6	Cardenas et al.,	1.200	0.890	1.600	0.950	1.200	0.182	0.150	0.022						
7	Lam et al., 1987	1.650	1.160	2.350	0.950	1.650	0.501	0.180	0.032						
8	Hirayama, 1984	1.450	1.020	2.080	0.950	1.450	0.372	0.182	0.033						
9	Sobue, 1990	1.060	0.740	1.520	0.950	1.060	0.058	0.184	0.034						
10	Gao et al., 1987	1.190	0.820	1.730	0.950	1.190	0.174	0.190	0.036						
11	Sun et al., 1996	1.160	0.800	1.690	0.950	1.160	0.148	0.191	0.036						
12	Zaridze et al., 1995	1.660	1.120	2.450	0.950	1.660	0.507	0.200	0.040						
13	Garfinkel et al.,	1.230	0.810	1.870	0.950	1.230	0.207	0.213	0.046						
14	Wang et al., 1996	1.110	0.670	1.840	0.950	1.110	0.104	0.258	0.066						

- Click in the first row to select it
- Click Edit > Delete row and confirm



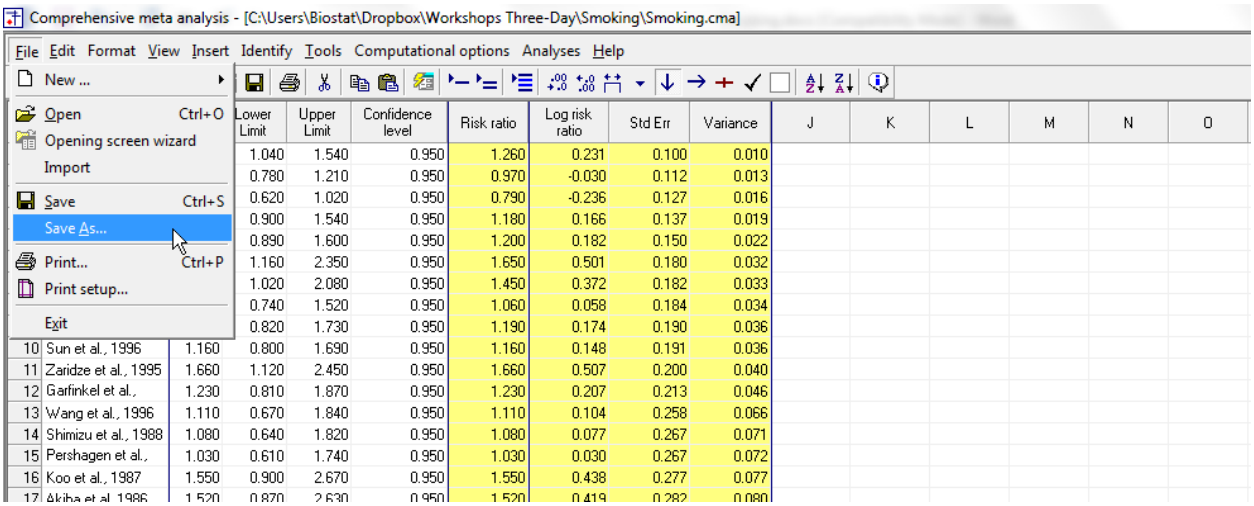
The screen should look like this



	Study name	Risk ratio	Lower Limit	Upper Limit	Confidence level	Risk ratio	Log risk ratio	Std Err	Variance
1	Fontham et al.,	1.260	1.040	1.540	0.950	1.260	0.231	0.100	0.010
2	Brownson et al.,	0.970	0.780	1.210	0.950	0.970	-0.030	0.112	0.013
3	Wu-Williams et al.,	0.790	0.620	1.020	0.950	0.790	-0.236	0.127	0.016
4	Garfinkel, 1992	1.180	0.900	1.540	0.950	1.180	0.166	0.137	0.019
5	Cardenas et al.,	1.200	0.890	1.600	0.950	1.200	0.182	0.150	0.022
6	Lam et al., 1987	1.650	1.160	2.350	0.950	1.650	0.501	0.180	0.032
7	Hirayama, 1984	1.450	1.020	2.080	0.950	1.450	0.372	0.182	0.033
8	Sobue, 1990	1.060	0.740	1.520	0.950	1.060	0.058	0.184	0.034
9	Gao et al., 1987	1.190	0.820	1.730	0.950	1.190	0.174	0.190	0.036
10	Sun et al., 1996	1.160	0.800	1.690	0.950	1.160	0.148	0.191	0.036

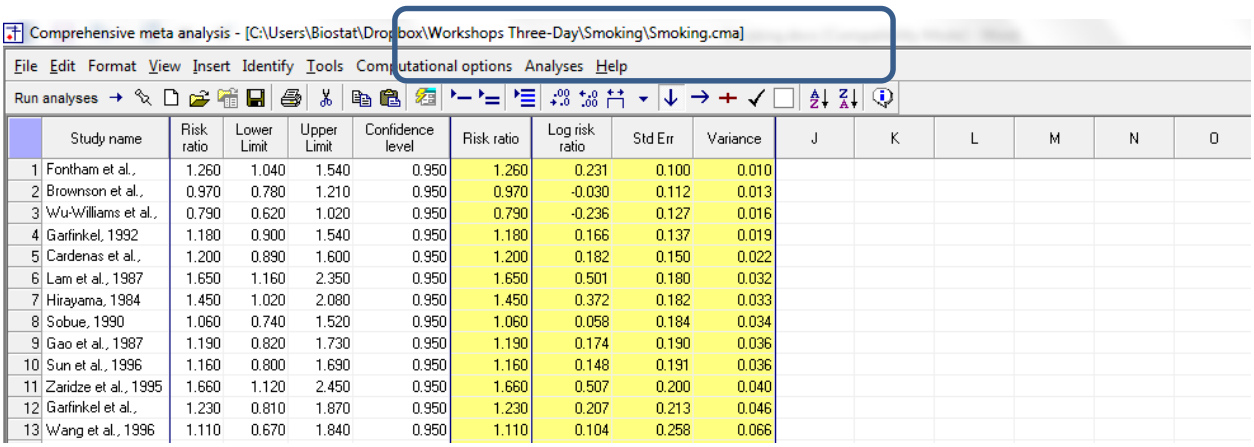
At this point we should save the file

- Click File > Save As …



Note that the file name is now in the header.

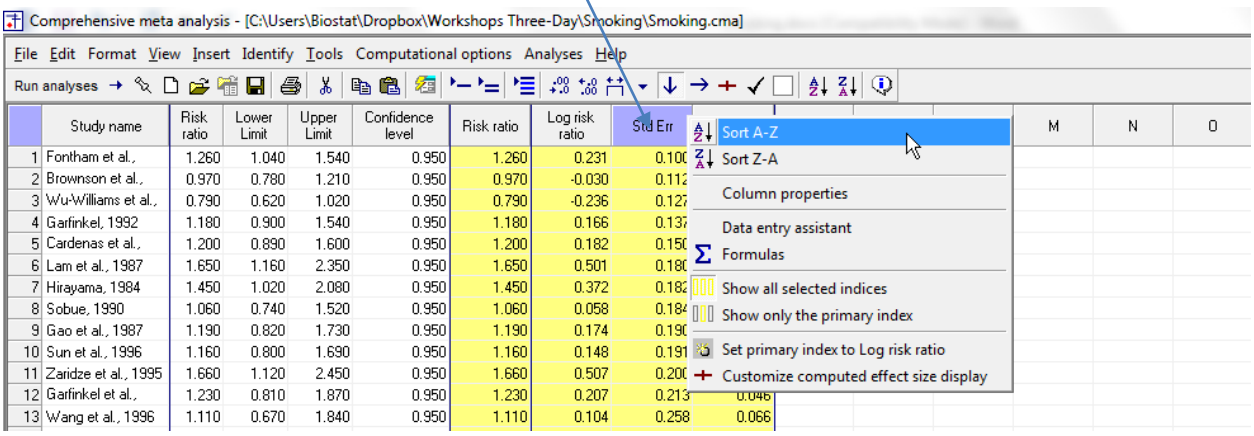
- [Save] will over-write the prior version of this file without warning
- [Save As…] will allow you to save the file with a new name



	Study name	Risk ratio	Lower Limit	Upper Limit	Confidence level	Risk ratio	Log risk ratio	Std Err	Variance
1	Fontham et al.,	1.260	1.040	1.540	0.950	1.260	0.231	0.100	0.010
2	Brownson et al.,	0.970	0.780	1.210	0.950	0.970	-0.030	0.112	0.013
3	Wu-Williams et al.,	0.790	0.620	1.020	0.950	0.790	-0.236	0.127	0.016
4	Garfinkel, 1992	1.180	0.900	1.540	0.950	1.180	0.166	0.137	0.019
5	Cardenas et al.,	1.200	0.890	1.600	0.950	1.200	0.182	0.150	0.022
6	Lam et al., 1987	1.650	1.160	2.350	0.950	1.650	0.501	0.180	0.032
7	Hirayama, 1984	1.450	1.020	2.080	0.950	1.450	0.372	0.182	0.033
8	Sobue, 1990	1.060	0.740	1.520	0.950	1.060	0.058	0.184	0.034
9	Gao et al., 1987	1.190	0.820	1.730	0.950	1.190	0.174	0.190	0.036
10	Sun et al., 1996	1.160	0.800	1.690	0.950	1.160	0.148	0.191	0.036
11	Zaridze et al., 1995	1.660	1.120	2.450	0.950	1.660	0.507	0.200	0.040
12	Garfinkel et al.,	1.230	0.810	1.870	0.950	1.230	0.207	0.213	0.046
13	Wang et al., 1996	1.110	0.670	1.840	0.950	1.110	0.104	0.258	0.066

For reasons that will become clear momentarily, it would be useful to sort the studies with those with the smallest standard error at the top and those with larger standard errors at the bottom.

- Right-click on the “Std Err” column
- Click Sort A-Z
- Save the file

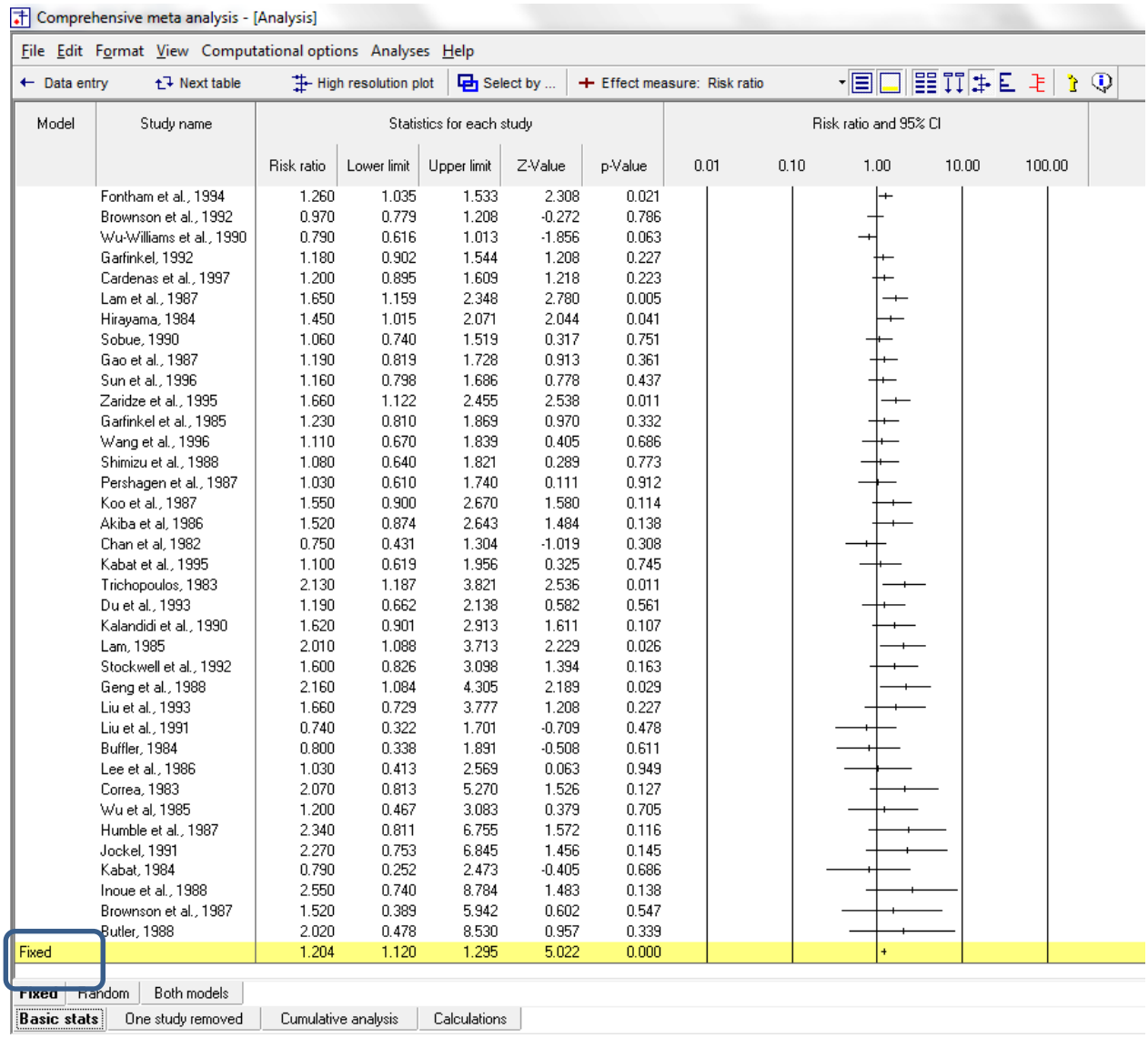


	Study name	Risk ratio	Lower Limit	Upper Limit	Confidence level	Risk ratio	Log risk ratio	Std Err
1	Fontham et al.,	1.260	1.040	1.540	0.950	1.260	0.231	0.100
2	Brownson et al.,	0.970	0.780	1.210	0.950	0.970	-0.030	0.112
3	Wu-Williams et al.,	0.790	0.620	1.020	0.950	0.790	-0.236	0.127
4	Garfinkel, 1992	1.180	0.900	1.540	0.950	1.180	0.166	0.137
5	Cardenas et al.,	1.200	0.890	1.600	0.950	1.200	0.182	0.150
6	Lam et al., 1987	1.650	1.160	2.350	0.950	1.650	0.501	0.180
7	Hirayama, 1984	1.450	1.020	2.080	0.950	1.450	0.372	0.182
8	Sobue, 1990	1.060	0.740	1.520	0.950	1.060	0.058	0.184
9	Gao et al., 1987	1.190	0.820	1.730	0.950	1.190	0.174	0.190
10	Sun et al., 1996	1.160	0.800	1.690	0.950	1.160	0.148	0.191
11	Zaridze et al., 1995	1.660	1.120	2.450	0.950	1.660	0.507	0.200
12	Garfinkel et al.,	1.230	0.810	1.870	0.950	1.230	0.207	0.213
13	Wang et al., 1996	1.110	0.670	1.840	0.950	1.110	0.104	0.258

To run the analysis, click [Run analysis]

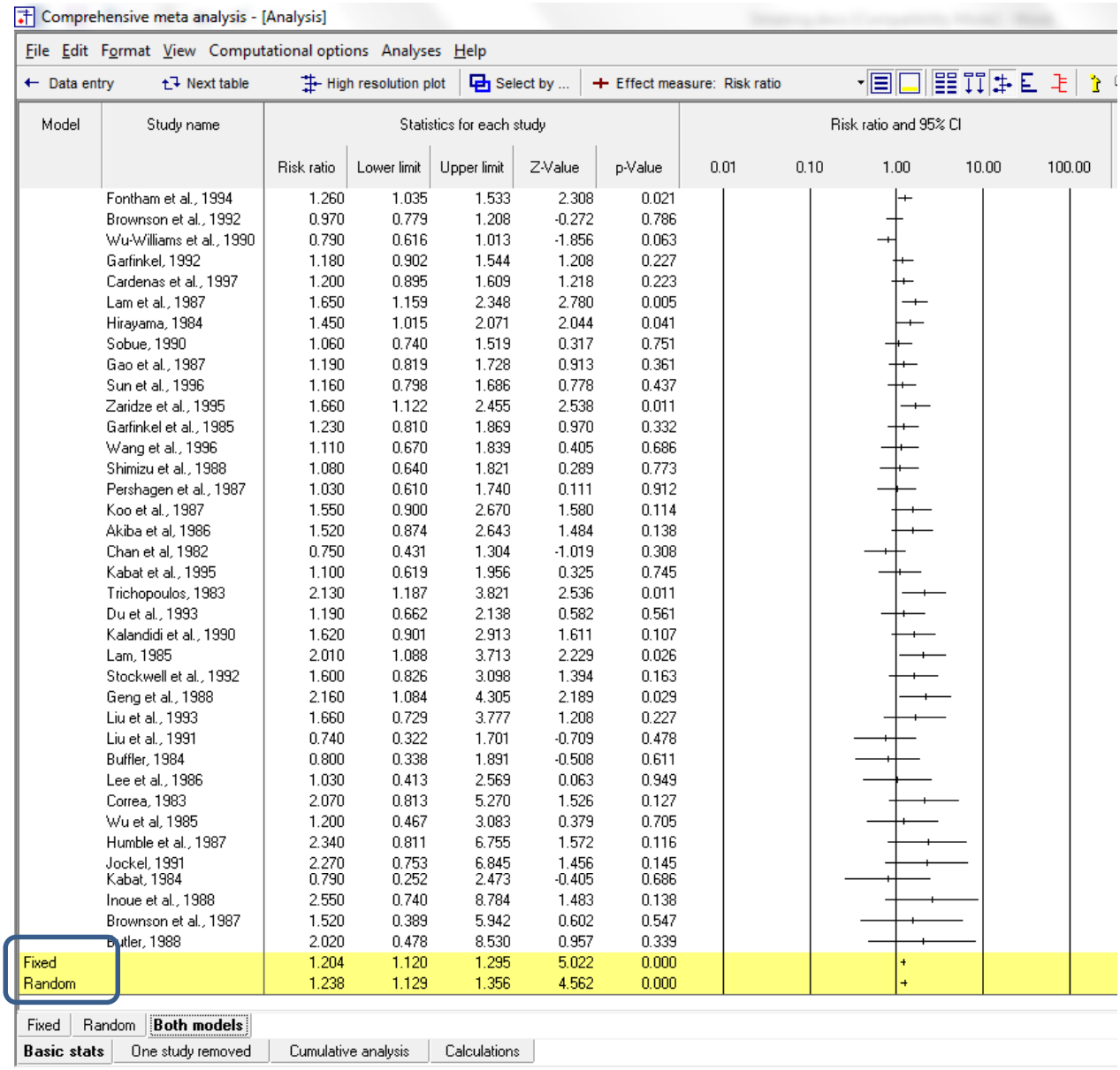
This is the basic analysis screen

Initially, the program displays the fixed-effect analysis. This is indicated by the tab at the bottom and the label in the plot.



Click [Both models]

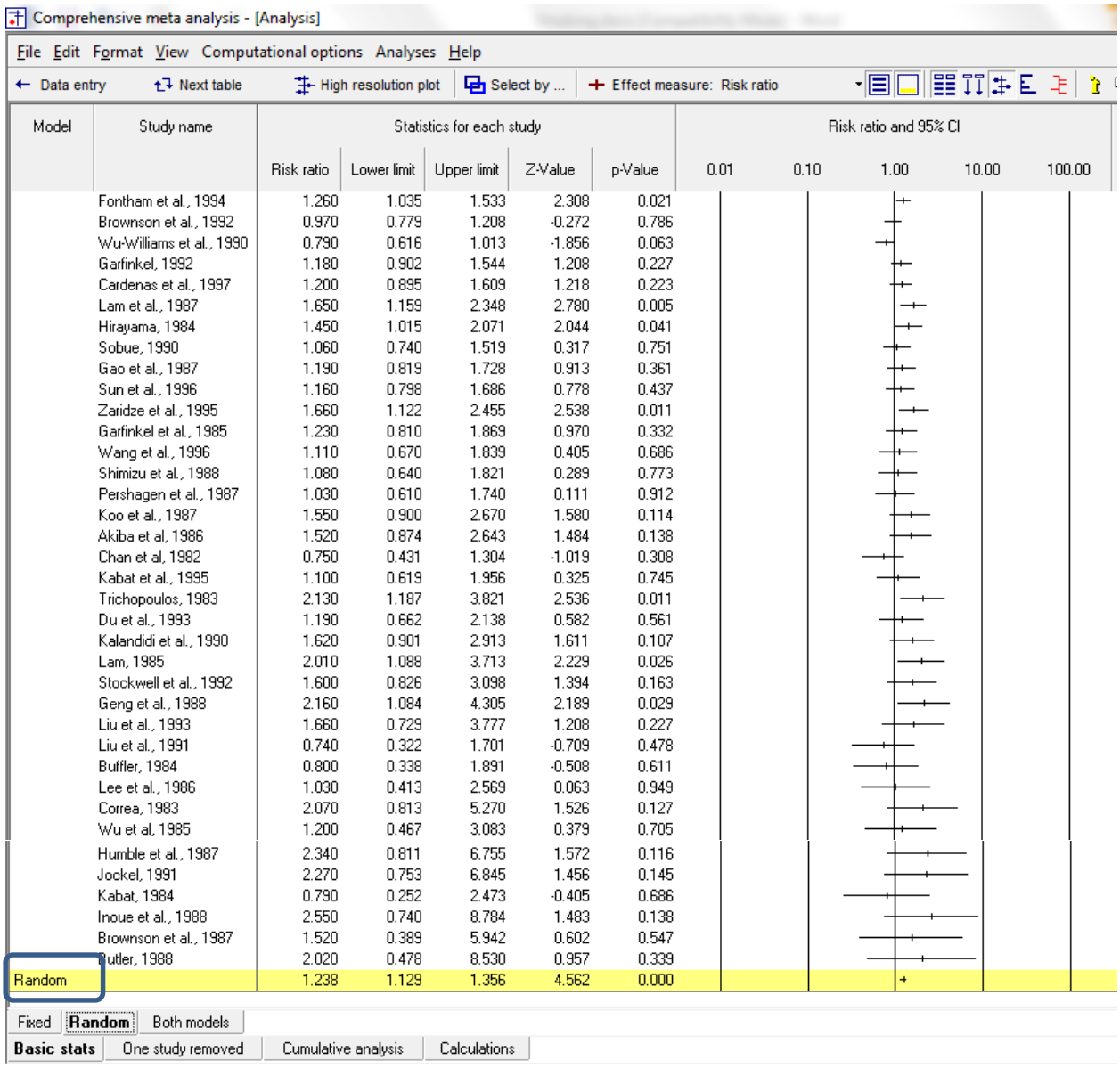
The program displays results for both the fixed-effect and the random-effects analysis.



The random-effects model is a better fit for the way the studies were sampled, and therefore that is the model we will use in the analysis.

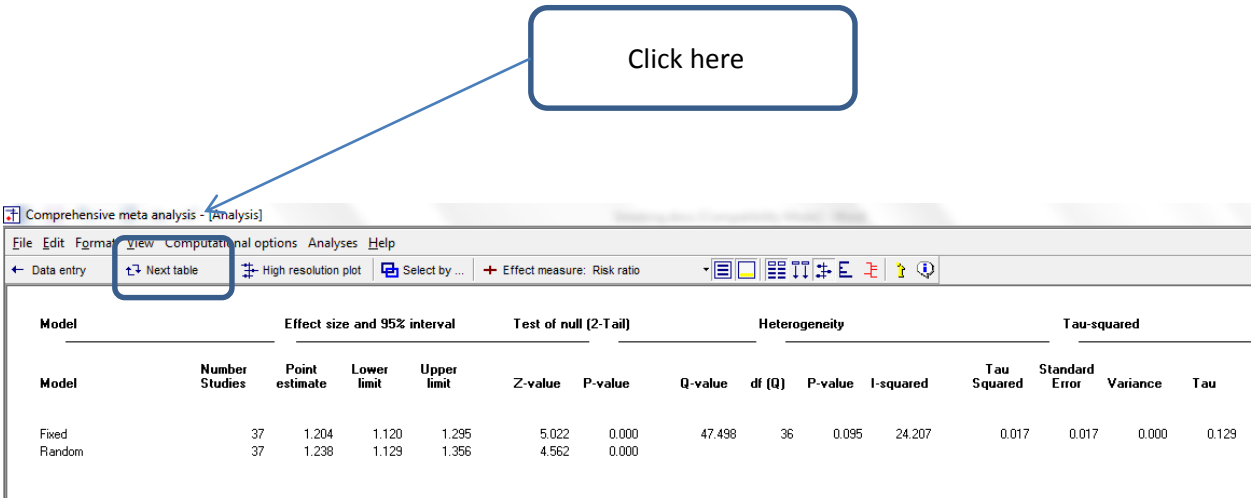
- Click Random on the tab at the bottom

The plot now displays the random-effects analysis alone.



- The summary effect is 1.238 with a CI of 1.129 to 1.356
- The summary effect has a *Z*-value 4.562 and a *p*-value of < 0.001. Thus we can reject the null hypotheses that the true risk ratio is 1.0.

Click [Next table]



Model		Effect size and 95% interval			Test of null (2-Tail)		Heterogeneity				Tau-squared			
Model	Number Studies	Point estimate	Lower limit	Upper limit	Z-value	P-value	Q-value	df (Q)	P-value	I-squared	Tau Squared	Standard Error	Variance	Tau
Fixed	37	1.204	1.120	1.295	5.022	0.000	47.498	36	0.095	24.207	0.017	0.017	0.000	0.129
Random	37	1.238	1.129	1.356	4.562	0.000								

The statistics at the left duplicate those we saw on the prior screen.

- The summary effect is 1.238 with a CI of 1.129 to 1.356.
- The summary effect has a *Z*-value 4.562 and a *p*-value of < 0.001. Thus we can reject the null hypotheses that the true risk ratio is 1.0.
- The *Q*-value is 47.498 with *df*=36 and p=0.095. *Q* reflects the distance of each study from the mean effect (weighted, squared, and summed over all studies). *Q* is always computed using FE weights (which is the reason it is displayed on the "Fixed" row, but applies to both FE and RE analyses.
- If all studies actually shared the same true effect size, the expected value of *Q* would be equal to *df* (which is 36). Here, *Q* is greater than 36, which tells us that there is some excess dispersion in effects, and p is 0.095, which tells us that this amount of excess dispersion is statistically significant using the criterion of 0.10 (which is the suggested criterion for this test).
- T^2 is the estimate of the between-study variance in true effects. This estimate is 0.017. *T* is the estimate of the between-study standard deviation in true effects. This estimate is 0.129. Note that these values are in log units.
- The variance in effect sizes includes both sampling error and variance in the true effect size from study to study. The I^2 value is 24.207, which tell is that about one-fourth of the variance in observed effects reflects differences in true effect sizes.

While it's clear that these studies show that being married to a smoker is associated with an increased risk of lung-cancer, there was concern that the studies in the analysis were not representative of all studies that had been conducted.

Consider the following scenario: 100 researchers run studies similar to the ones in this analysis. Some find an increased risk for second-hand smoke, some find no risk, and some find that second-hand smoke is associated with a *decreased* risk of lung-cancer. (Note that we're talking about the direction of the effect, not necessarily a statistically significant effect).

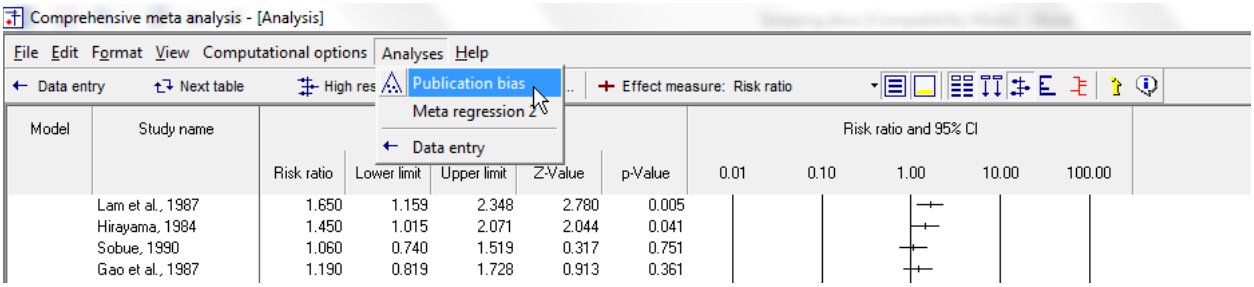
The researchers who find an increased risk proceed to publish, those who find no relationship are less likely to publish, and those who find a protective effect for smoking are even less likely to publish. What happens is that the papers that are published tend to over-estimate the true effect size. These, of course, are the papers that are more likely to find their way into the meta-analysis.

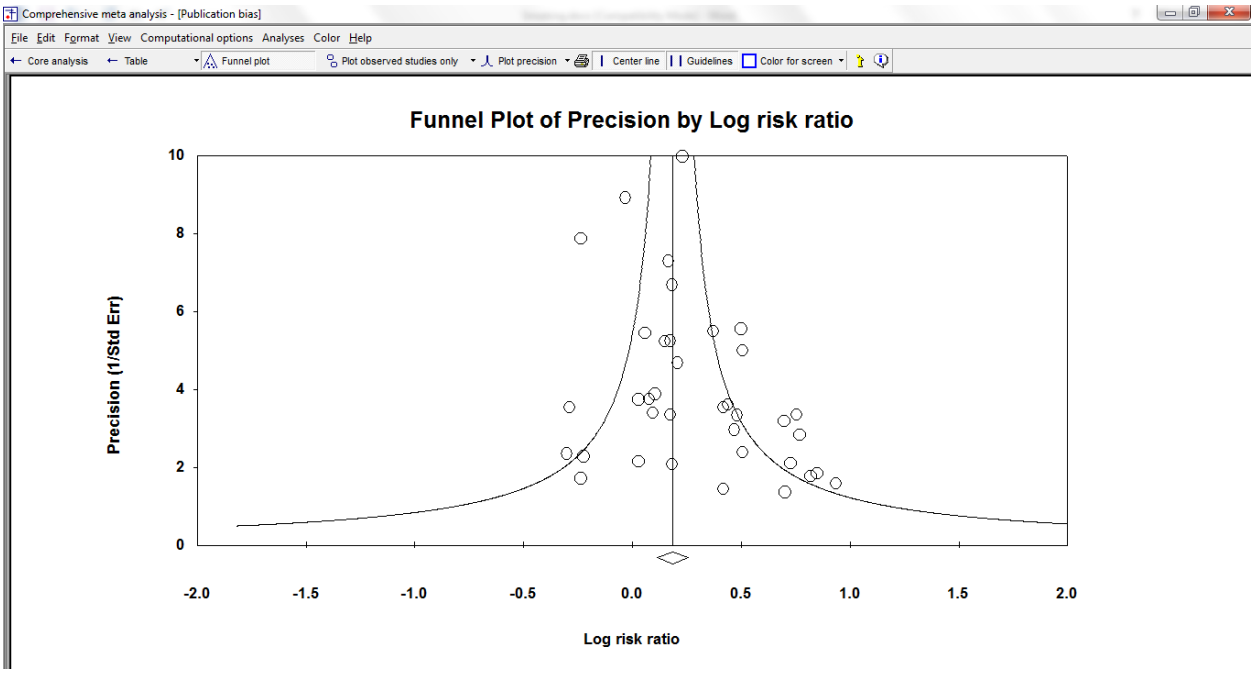
Is it plausible that the true relationship between second-hand smoke and lung-cancer across all the studies that had been performed is nil, and the effect that we see in these studies is entirely an artifact of publication bias?

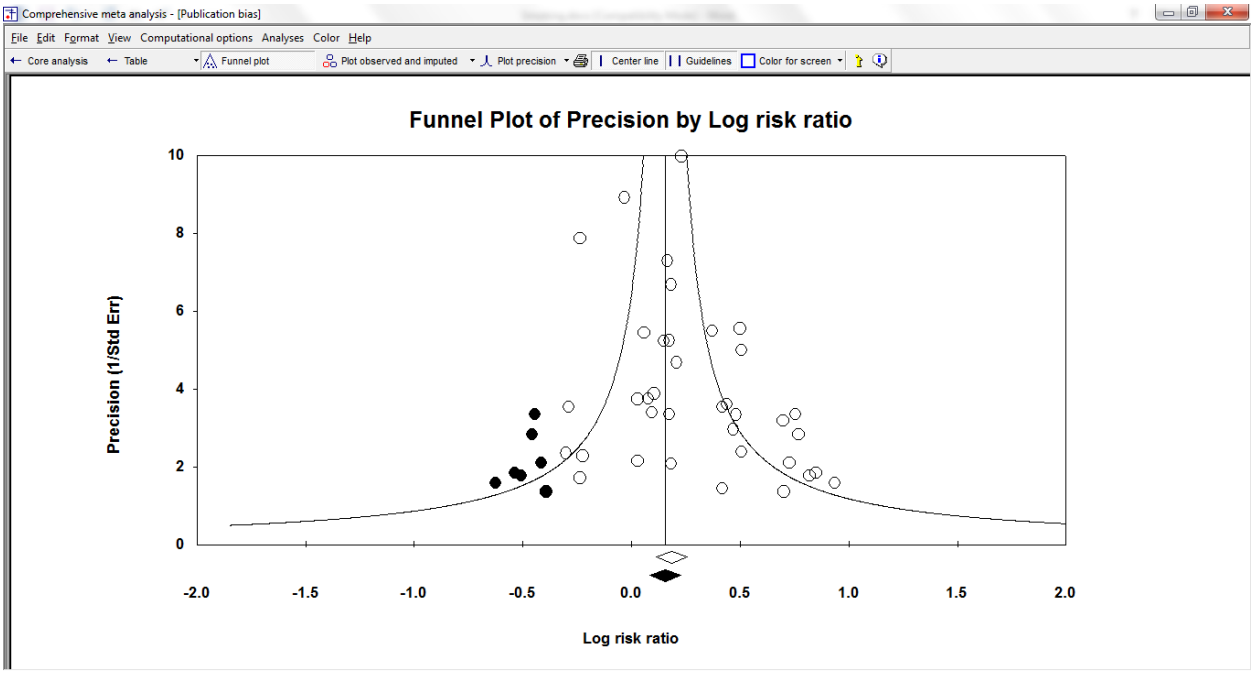
A number of procedures have been developed to test for publication bias and/or assess its potential impact on the results. Here, we outline how to perform these analyses, without fully discussing their meaning. We should note, however, that most of these analyses are based on the assumption that if small studies show a larger effect than large studies, this may be evidence of publication bias. In truth, this *may be* evidence of publication bias but may also be a small study effect, such that the effect is actually larger in smaller studies for legitimate reasons.

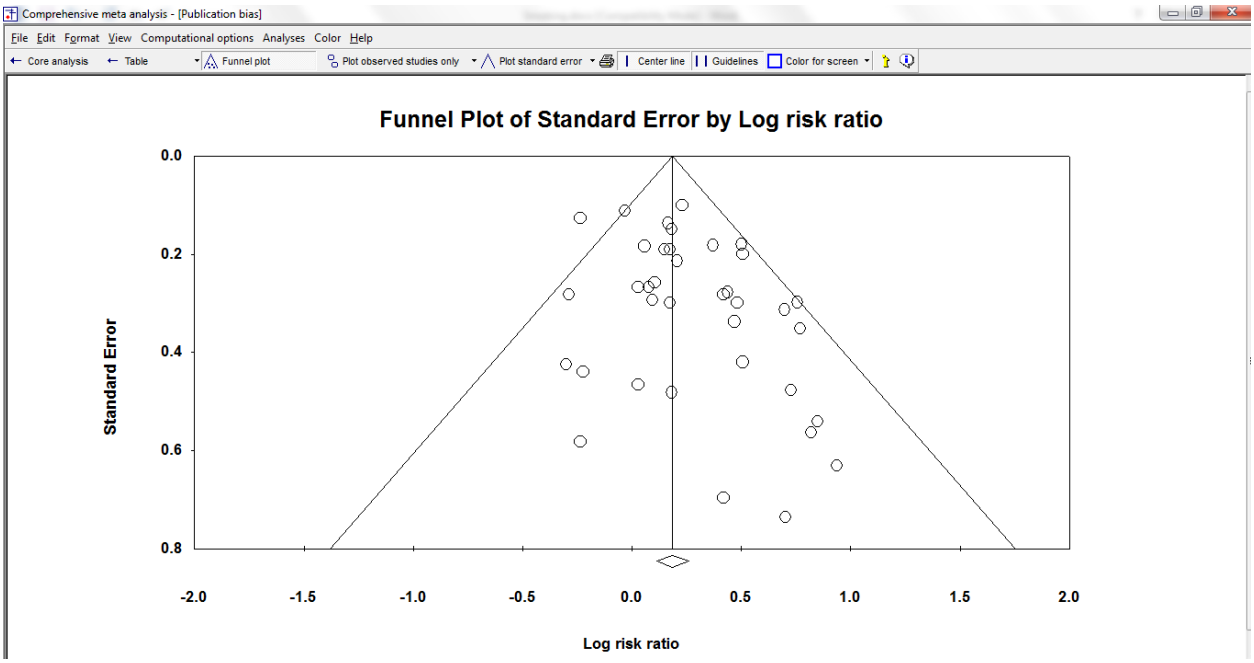
To run the analyses

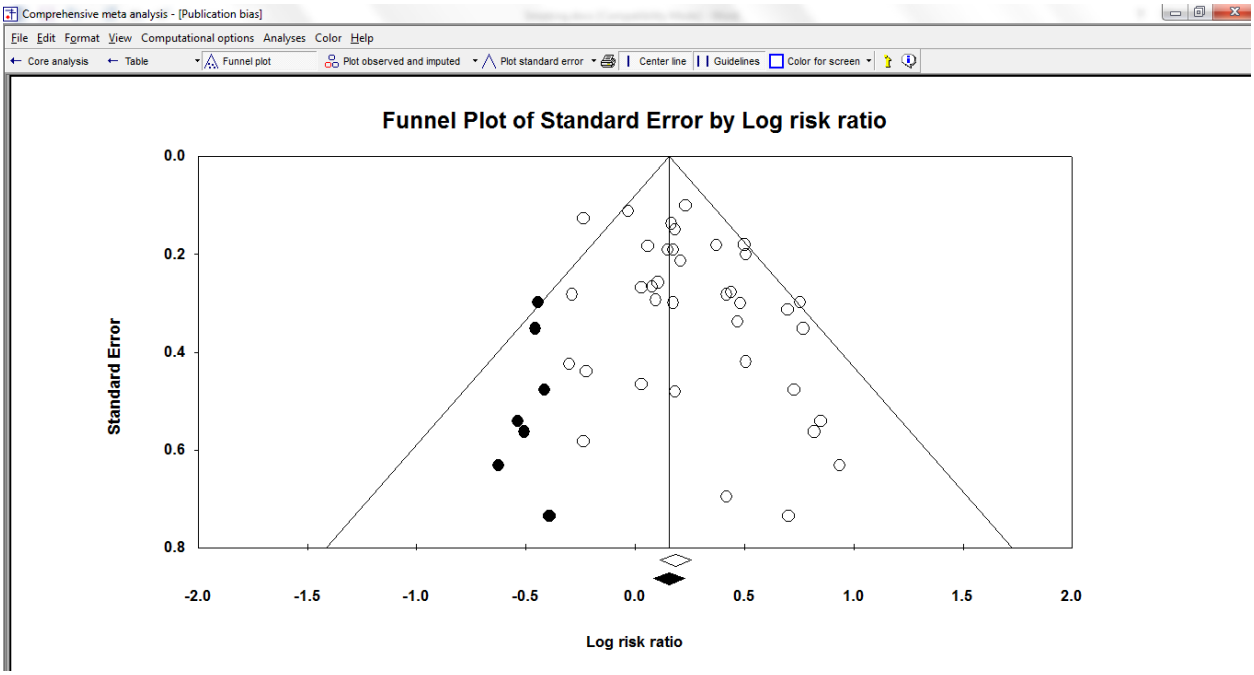
Click Analyses > Publication bias

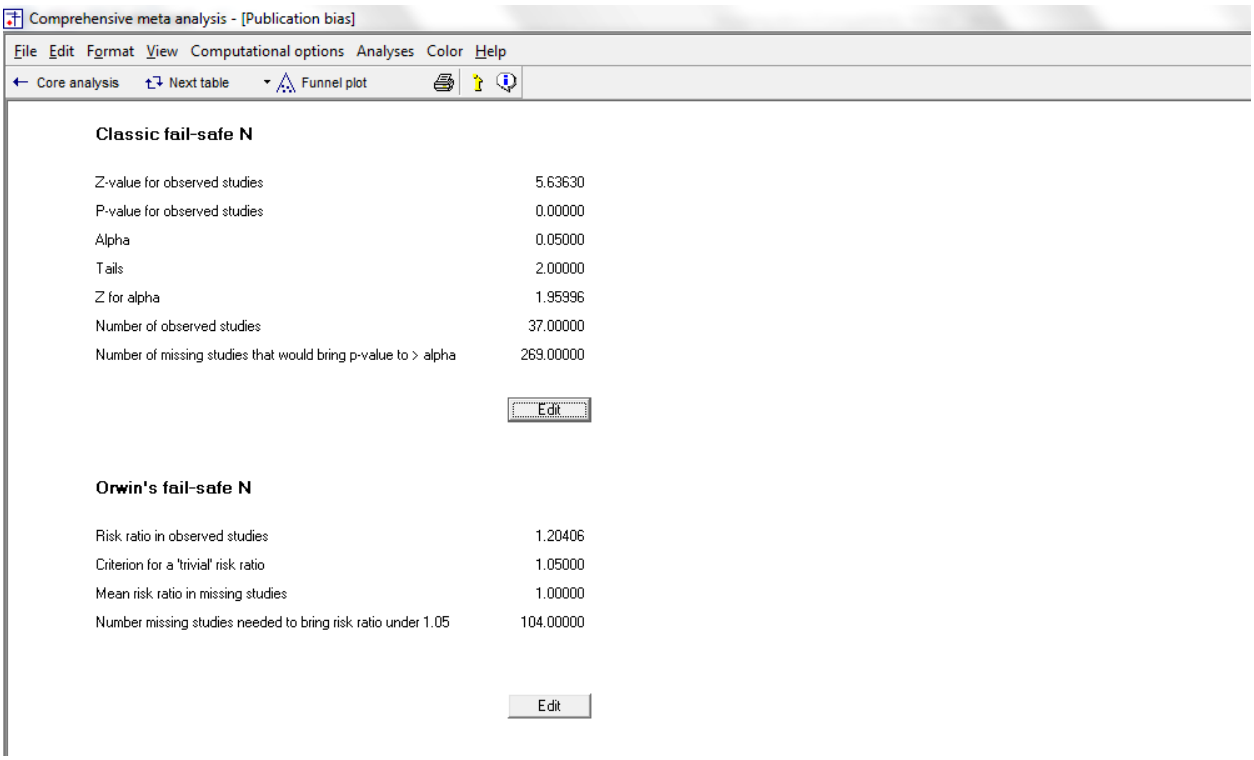


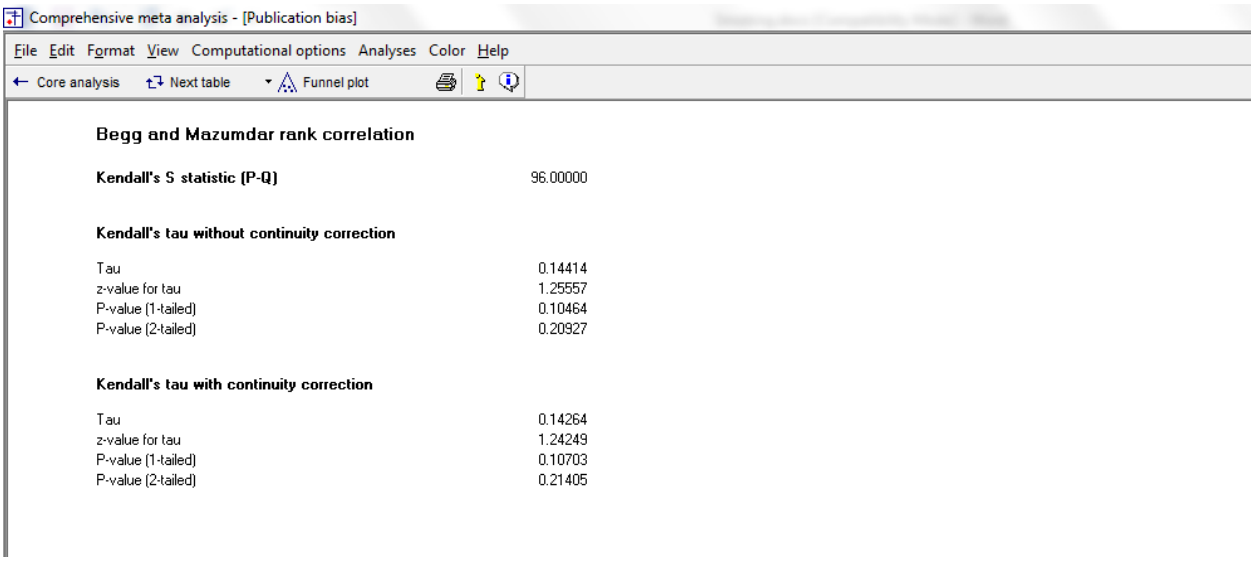

Comprehensive meta analysis - [Publication bias]
File Edit Format View Computational options Analyses Color Help
Core analysis Table Funnel plot Plot observed studies only Plot precision Center line Guidelines Color for screen
Funnel Plot of Precision by Log risk ratio
Precision (1/Std Err)
10
8
6
4
2
0
-2.0
-1.5
-1.0
-0.5
0.0
0.5
1.0
1.5
2.0
Log risk ratio


Comprehensive meta analysis - [Publication bias]
File Edit Format View Computational options Analyses Color Help
Core analysis Table Funnel plot Plot observed and imputed Plot precision Center line Guidelines Color for screen
Funnel Plot of Precision by Log risk ratio
Precision (1/Std Err)
10
8
6
4
2
0
-2.0
-1.5
-1.0
-0.5
0.0
0.5
1.0
1.5
2.0
Log risk ratio


Comprehensive meta analysis - [Publication bias]
File Edit Format View Computational options Analyses Color Help
Core analysis Table Funnel plot Plot observed studies only Plot standard error Center line Guidelines Color for screen
Funnel Plot of Standard Error by Log risk ratio
Standard Error
0.0
0.2
0.4
0.6
0.8
-2.0
-1.5
-1.0
-0.5
0.0
0.5
1.0
1.5
2.0
Log risk ratio


Comprehensive meta analysis - [Publication bias]
File Edit Format View Computational options Analyses Color Help
Core analysis Table Funnel plot Plot observed and imputed Plot standard error Center line Guidelines Color for screen
Funnel Plot of Standard Error by Log risk ratio
Standard Error
0.0
0.2
0.4
0.6
0.8
-2.0
-1.5
-1.0
-0.5
0.0
0.5
1.0
1.5
2.0
Log risk ratio


Comprehensive meta analysis - [Publication bias]
File Edit Format View Computational options Analyses Color Help
Core analysis
Next table
Funnel plot
Classic fail-safe N
Z-value for observed studies 5.63630
P-value for observed studies 0.00000
Alpha 0.05000
Tails 2.00000
Z for alpha 1.95996
Number of observed studies 37.00000
Number of missing studies that would bring p-value to > alpha 269.00000
Edit
Orwin's fail-safe N
Risk ratio in observed studies 1.20406
Criterion for a 'trivial' risk ratio 1.05000
Mean risk ratio in missing studies 1.00000
Number missing studies needed to bring risk ratio under 1.05 104.00000
Edit



Begg and Mazumdar rank correlation

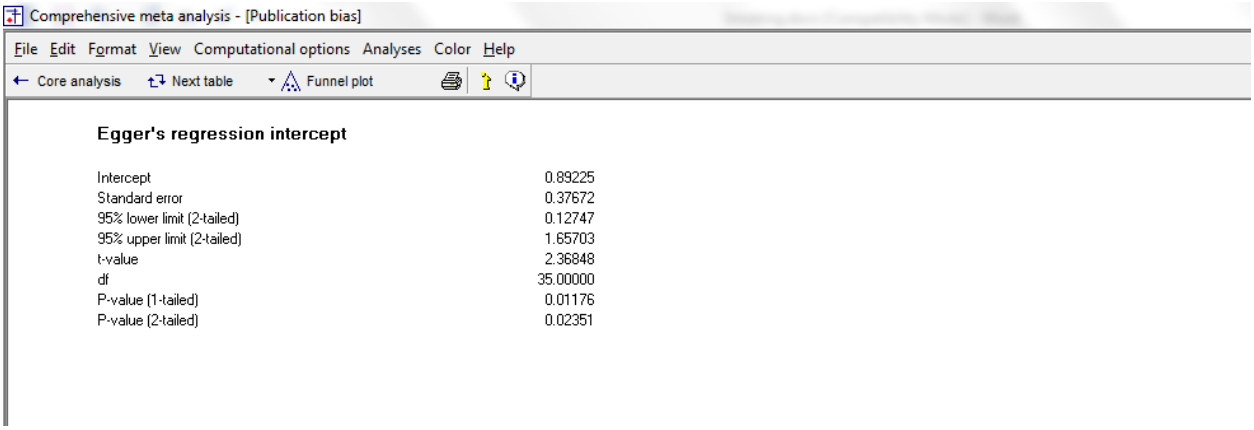
Kendall's S statistic (P-Q)	96.00000

Kendall's tau without continuity correction

Tau	0.14414
z-value for tau	1.25557
P-value (1-tailed)	0.10464
P-value (2-tailed)	0.20927

Kendall's tau with continuity correction

Tau	0.14264
z-value for tau	1.24249
P-value (1-tailed)	0.10703
P-value (2-tailed)	0.21405



Egger's regression intercept

Intercept	0.89225
Standard error	0.37672
95% lower limit (2-tailed)	0.12747
95% upper limit (2-tailed)	1.65703
t-value	2.36848
df	35.00000
P-value (1-tailed)	0.01176
P-value (2-tailed)	0.02351

The program generates the following report

Report starts here

Introduction to publication bias

The basic issue of publication bias is that not all completed studies are published, and the selection process is not random (hence the 'bias'). Rather, studies that report relatively large treatment effects are more likely to be submitted and/or accepted for publication than studies which report more modest treatment effects.

Since the treatment effect estimated from a biased collection of studies would tend to overestimate the true treatment effect, it is important to assess the likley extent of the bias, and its potential impact on the conclusions. This module includes several tools which may be used for this purpose.

Funnel plot

The funnel plot is a plot of a measure of study size (usually standard error or precision) on the vertical axis as a function of effect size on the horizontal axis.

Large studies appear toward the top of the graph, and tend to cluster near the mean effect size. Smaller studies appear toward the bottom of the graph, and (since there is more sampling variation in effect size estimates in the smaller studies) will be dispersed across a range of values.

In the absence of publication bias we would expect the studies to be distributed symmetrically about the combined effect size. By contrast, in the presence of bias, we would expect that the bottom of the plot would show a higher concentration of studies on one side of the mean than the other. This would reflect the fact that smaller studies (which appear toward the bottom) are more likely to be published if they have larger than average effects, which makes them more likely to meet the criterion for statistical significance.

Various statistical procedures can be accessed from the View menu to quantify or augment this display.

The classic fail-Safe N and the Orwin fail-safe N ask if we need to be concerned that the entire observed effect may be an artifact of bias. Rank correlation and regression procedures can test for the presence of bias. Trim and Fill offers a more nuanced perspective, and asks how the effect size would shift if the apparent bias were to be removed.

Classic fail-safe N

One concern of publication bias is that some non-significant studies are missing from our analysis and that these studies, if included, would nullify the observed effect.

Robert Rosenthal suggested that rather than simply speculate about the impact of the missing studies, we compute the number of studies that would be required to nullify the effect. If this number is relatively small then there is indeed cause for concern. However, if this number is large, we can be

confident that the treatment effect, while possibly inflated by the exclusion of some studies, is nevertheless not nil.

He suggested that this analysis be called a 'File-drawer' analysis, file drawers being the presumed location of the missing studies. Harris Cooper proposed the term 'Fail-Safe N', a reference to the number of missing studies that would nullify the effect.

This approach is limited in two important ways. First, it assumes that the effect in the hidden studies is nil, rather than considering the possibility that some of the studies could have shown an effect in the reverse direction. Therefore, the number of studies required to nullify the effect may be smaller than the Fail-Safe N.

Second, and more fundamentally, this approach focuses on statistical significance rather than clinical or substantive significance. That is, it may allow us to assert that the treatment effect is not nil, but does not address the question of whether or not it remains clinically important after the missing studies have been included.

Note also that the fail-safe N algorithm computes a p-value for each study and then combines these p-values. By contrast, the generally accepted approach today (and the one used by this program) is to compute an effect size for each study, combine the effect sizes, and then compute the p-value for the combined effect. The two approaches do not generally yield identical results.

This meta analysis incorporates data from 37 studies, which yield a z-value of 5.63630 and corresponding 2-tailed p-value of 0.00000.

The fail-safe N is 269. This means that we would need to locate and include 269 'null' studies in order for the combined 2-tailed p-value to exceed 0.050.

Put another way, there would be need to be 7.3 missing studies for every observed study for the effect to be nullified.

Orwin fail-safe N

Like the classic fail-safe N, the Orwin fail-safe N addresses the possibility that studies are missing from the analysis and that these studies, if included in the analysis, would shift the effect size toward the null.

Orwin's fail-safe N differs from the classic fail-safe N in two ways.

First, the mean risk ratio in the new (missing) studies can be a value other than the nil value (currently, it is set to 1).

Second, the criterion value is an effect size rather than a p-value. That is, the Orwin fail-safe N is the number of (missing) studies that, when added to the analysis, will bring the combined risk ratio below a specified threshold (currently, 1.05).

The Orwin fail-safe N is 104. This means that we would need to locate 104 studies with mean risk ratio of 1 to bring the combined risk ratio under 1.05

Begg and Mazumdar Rank Correlation Test

The classic case of publication bias is the case depicted by the funnel plot. Large studies tend to be included in the analysis regardless of their treatment effect whereas small studies are more likely to be included when they show a relatively large treatment effect. Under these circumstances there will be an inverse correlation between study size and effect size.

Begg and Mazumdar suggested that this correlation can serve as a test for publication bias. Concretely, they suggest that we compute the rank order correlation (Kendall's tau b) between the treatment effect and the standard error (which is driven primarily by sample size).

This approach is limited in some important ways. A significant correlation suggests that bias exists but does not directly address the implications of this bias. Conversely, a non-significant correlation may be due to low statistical power, and cannot be taken as evidence that bias is absent.

In this case Kendall's tau b (corrected for ties, if any) is 0.14264, with a 1-tailed p-value (recommended) of 0.10703 or a 2-tailed p-value of 0.21405 (based on continuity-corrected normal approximation).

Egger's Test of the Intercept

Egger suggests that we assess this same bias by using precision (the inverse of the standard error) to predict the standardized effect (effect size divided by the standard error). In this equation, the size of the treatment effect is captured by the slope of the regression line (B1) while bias is captured by the intercept (B0).

This approach may offer a number of advantages over the rank correlation approach. Under some circumstances this may be a more powerful test. Additionally, this approach can be extended to include more than one predictor variable, which means that we can simultaneously assess the impact of several factors, including sample size, on the treatment effect.

In this case the intercept (B0) is 0.89225, 95% confidence interval (0.12747, 1.65703), with t=2.36848, df=35. The 1-tailed p-value (recommended) is 0.01176, and the 2-tailed p-value is 0.02351.

Duval and Tweedie's Trim and Fill

If the meta analysis had captured all the relevant studies we would expect the funnel plot to be symmetric. That is, we would expect studies to be dispersed equally on either side of the overall effect. Therefore, if the funnel plot is actually asymmetric, with a relatively high number of small studies (representing a large effect size) falling toward the right of the mean effect and relatively few falling toward the left, we are concerned that these left-hand studies may actually exist, and are missing from the analysis.

Duval and Tweedie developed a method that allows us to impute these studies. That is, we determine where the missing studies are likely to fall, add them to the analysis, and then recompute the combined effect.

The method is known as 'Trim and Fill' as the method initially trims the asymmetric studies from the right-hand side to locate the unbiased effect (in an iterative procedure), and then fills the plot by re-inserting the trimmed studies on the right as well as their imputed counterparts to the left the mean effect.

The program is looking for missing studies based on a fixed effect model, and is looking for missing studies only to the left side of the mean effect (these parameters are set by the user). Using these parameters the method suggests that 7 studies are missing.

Under the fixed effect model the point estimate and 95% confidence interval for the combined studies is 1.20406 (1.11988, 1.29457). Using Trim and Fill the imputed point estimate is 1.16893 (1.08901, 1.25471).

Under the random effects model the point estimate and 95% confidence interval for the combined studies is 1.23770 (1.12934, 1.35645). Using Trim and Fill the imputed point estimate is 1.18862 (1.08023, 1.30789).

To plot the imputed studies click 'Funnel plot' and then select 'Plot observed and imputed' on the toolbar.

An important caveat

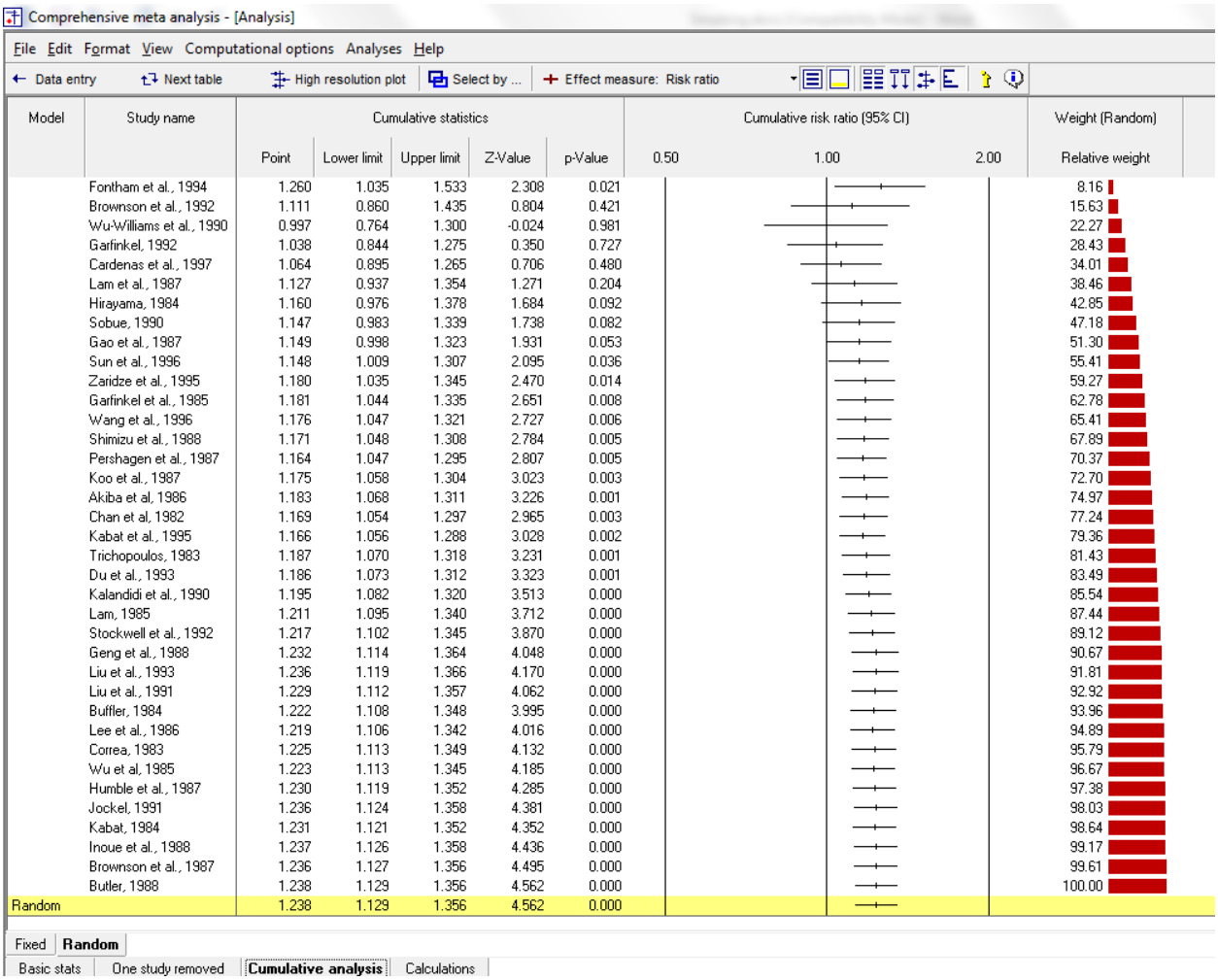
Sterne and Egger note that while the plot and these procedures may detect a relationship between sample size and effect size, they cannot assign a causal mechanism to it.

That is, the effect size MAY be larger in small studies because we retrieved a biased sample of the smaller studies. However, it is also possible that the effect size really IS larger in smaller studies - perhaps because the smaller studies used different populations or different protocols than the larger ones.

Sterne and Egger use the term 'small study effect' to capture these and other potential confounds.

Report ends here

Finally, we can run a cumulative analysis, sorted by standard error



Model	Study name	Cumulative statistics					Cumulative risk ratio (95% CI)	Weight (Random)
		Point	Lower limit	Upper limit	Z-Value	p-Value	0.50 1.00 2.00	Relative weight
	Fontham et al., 1994	1.260	1.035	1.533	2.308	0.021		8.16
	Brownson et al., 1992	1.111	0.860	1.435	0.804	0.421		15.63
	Wu-Williams et al., 1990	0.997	0.764	1.300	-0.024	0.981		22.27
	Garfinkel, 1992	1.038	0.844	1.275	0.350	0.727		28.43
	Cardenas et al., 1997	1.064	0.895	1.265	0.706	0.480		34.01
	Lam et al., 1987	1.127	0.937	1.354	1.271	0.204		38.46
	Hirayama, 1984	1.160	0.976	1.378	1.684	0.092		42.85
	Sobue, 1990	1.147	0.983	1.339	1.738	0.082		47.18
	Gao et al., 1987	1.149	0.998	1.323	1.931	0.053		51.30
	Sun et al., 1996	1.148	1.009	1.307	2.095	0.036		55.41
	Zaridze et al., 1995	1.180	1.035	1.345	2.470	0.014		59.27
	Garfinkel et al., 1985	1.181	1.044	1.335	2.651	0.008		62.78
	Wang et al., 1996	1.176	1.047	1.321	2.727	0.006		65.41
	Shimizu et al., 1988	1.171	1.048	1.308	2.784	0.005		67.89
	Pershagen et al., 1987	1.164	1.047	1.295	2.807	0.005		70.37
	Koo et al., 1987	1.175	1.058	1.304	3.023	0.003		72.70
	Akiba et al, 1986	1.183	1.068	1.311	3.226	0.001		74.97
	Chan et al, 1982	1.169	1.054	1.297	2.965	0.003		77.24
	Kabat et al., 1995	1.166	1.056	1.288	3.028	0.002		79.36
	Trichopoulos, 1983	1.187	1.070	1.318	3.231	0.001		81.43
	Du et al., 1993	1.186	1.073	1.312	3.323	0.001		83.49
	Kalandidi et al., 1990	1.195	1.082	1.320	3.513	0.000		85.54
	Lam, 1985	1.211	1.095	1.340	3.712	0.000		87.44
	Stockwell et al., 1992	1.217	1.102	1.345	3.870	0.000		89.12
	Geng et al., 1988	1.232	1.114	1.364	4.048	0.000		90.67
	Liu et al., 1993	1.236	1.119	1.366	4.170	0.000		91.81
	Liu et al., 1991	1.229	1.112	1.357	4.062	0.000		92.92
	Buffler, 1984	1.222	1.108	1.348	3.995	0.000		93.96
	Lee et al., 1986	1.219	1.106	1.342	4.016	0.000		94.89
	Correa, 1983	1.225	1.113	1.349	4.132	0.000		95.79
	Wu et al, 1985	1.223	1.113	1.345	4.185	0.000		96.67
	Humble et al., 1987	1.230	1.119	1.352	4.285	0.000		97.38
	Jockel, 1991	1.236	1.124	1.358	4.381	0.000		98.03
	Kabat, 1984	1.231	1.121	1.352	4.352	0.000		98.64
	Inoue et al., 1988	1.237	1.126	1.358	4.436	0.000		99.17
	Brownson et al., 1987	1.236	1.127	1.356	4.495	0.000		99.61
	Butler, 1988	1.238	1.129	1.356	4.562	0.000		100.00
Random		1.238	1.129	1.356	4.562	0.000		