CampbellCollaboration

Social Welfare Coordinating Group

Effects of Early, Computerized Brief Interventions on Risky Alcohol Use and Risky Cannabis Use Among Young People

Geir Smedslund, Sabine Wollscheid, Lin Fang, Wendy Nilsen, Asbjørn Steiro and Lillebeth Larun

https://www.campbellcollaboration.org/library/computerised-inteventions-youth-alcohol-cannabis-use.html

The example that follows is based on the systematic review cited above, and published by the Campbell Collaboration.

The example that follows is based on one analysis taken from the review. We chose this analysis because it allows us to illustrate various facets of an analysis (and CMA). It is not intended to be representative of all analyses in the review. For details of the review, please see the original.

In the pages that follow we focus on the meta-analysis (the synthesis of data). The full review can serve as an example of how to search the literature, decide on which studies to include, and the many other parts of a systematic review.

Example name   Prevent alcohol abuse

Effect size   Standardized Mean difference
Analysis type   Basic
Level   Basic

Synopsis

We have a series of studies that evaluated the impact of a brief computer-based intervention on the incidence of alcohol abuse in young people. Since studies used a number of different tests to assess alcohol use, we use the Standardized mean difference (Hedges *g*) as the outcome.

We show

- How to enter data for means and standard deviation in two independent groups
- How to run a random-effects analysis
- How to get a visual sense of the dispersion
- How to see the weight assigned to each study
- How to understand the statistics for the summary effect and the dispersion
- How to create a high-resolution plot

The data for this exercise are located at

www.Common-Mistakes-in-Meta-Analysis.com

Navigate to the folder "Intervention to prevent alcohol abuse"

Download the following

- This PDF
- Prevent Alcohol Abuse.xls [The data for this analysis in an Excel format]
- Prevent Alcohol Abuse-PI.xls [The Excel file for computing prediction intervals]

CMA works with computers that use a period or a comma to indicate decimal places. If you are building the CMA file yourself, CMA will automatically use your computer's setting. If you are opening a CMA file that someone else built, your computer must use the setting that was used when the file was created.

If your computer uses a period to indicate decimals, you may download the file below. Otherwise, please send a note to Biostat100@GMail.com

Prevent Alcohol Abuse.cma [The CMA data file]

To download a trial of CMA go to www.Meta-Analysis.com

If you would like to start with a blank file and enter the data, proceed with the next page

If you would like to simply open the CMA file with the data, proceed with this page
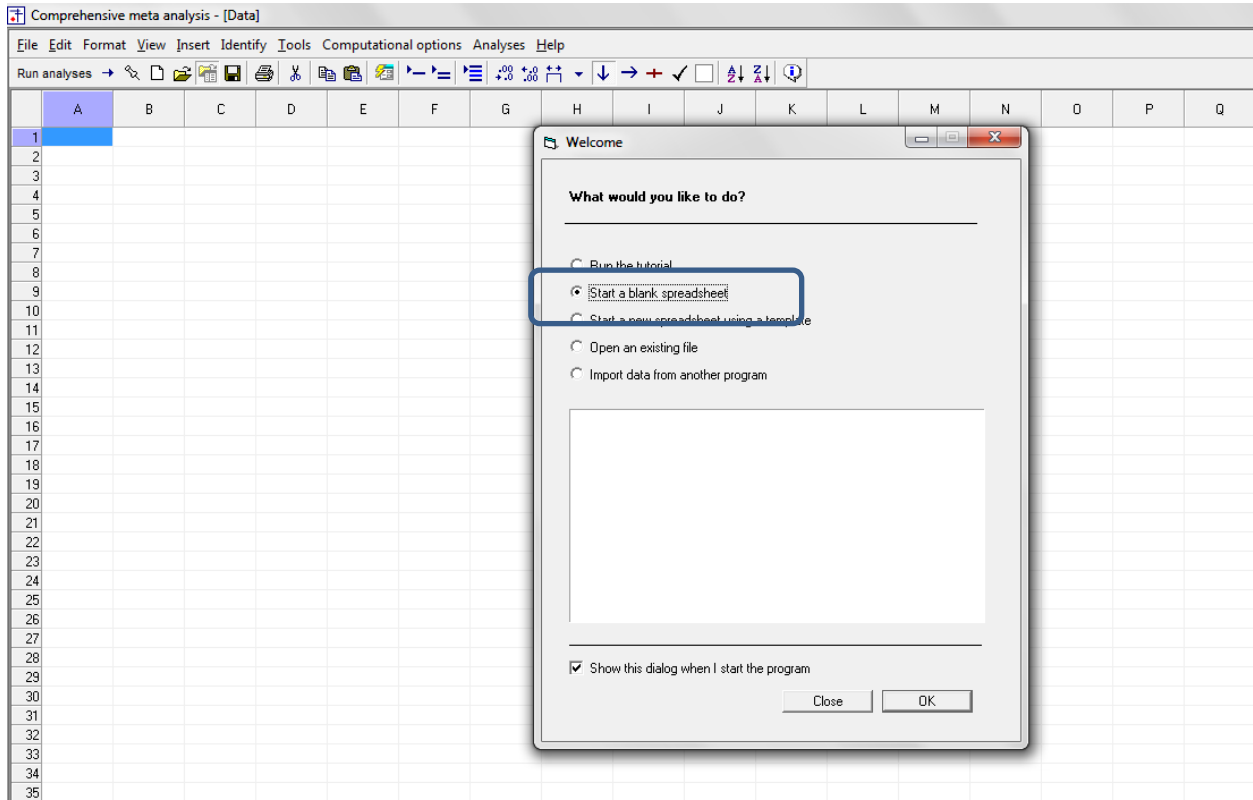
- Start the CMA program
- Click the icon for the Opening Wizard
- Click More Files
- Locate the file Prevent Alcohol Abuse.CMA
- Proceed with the analysis on Page 21
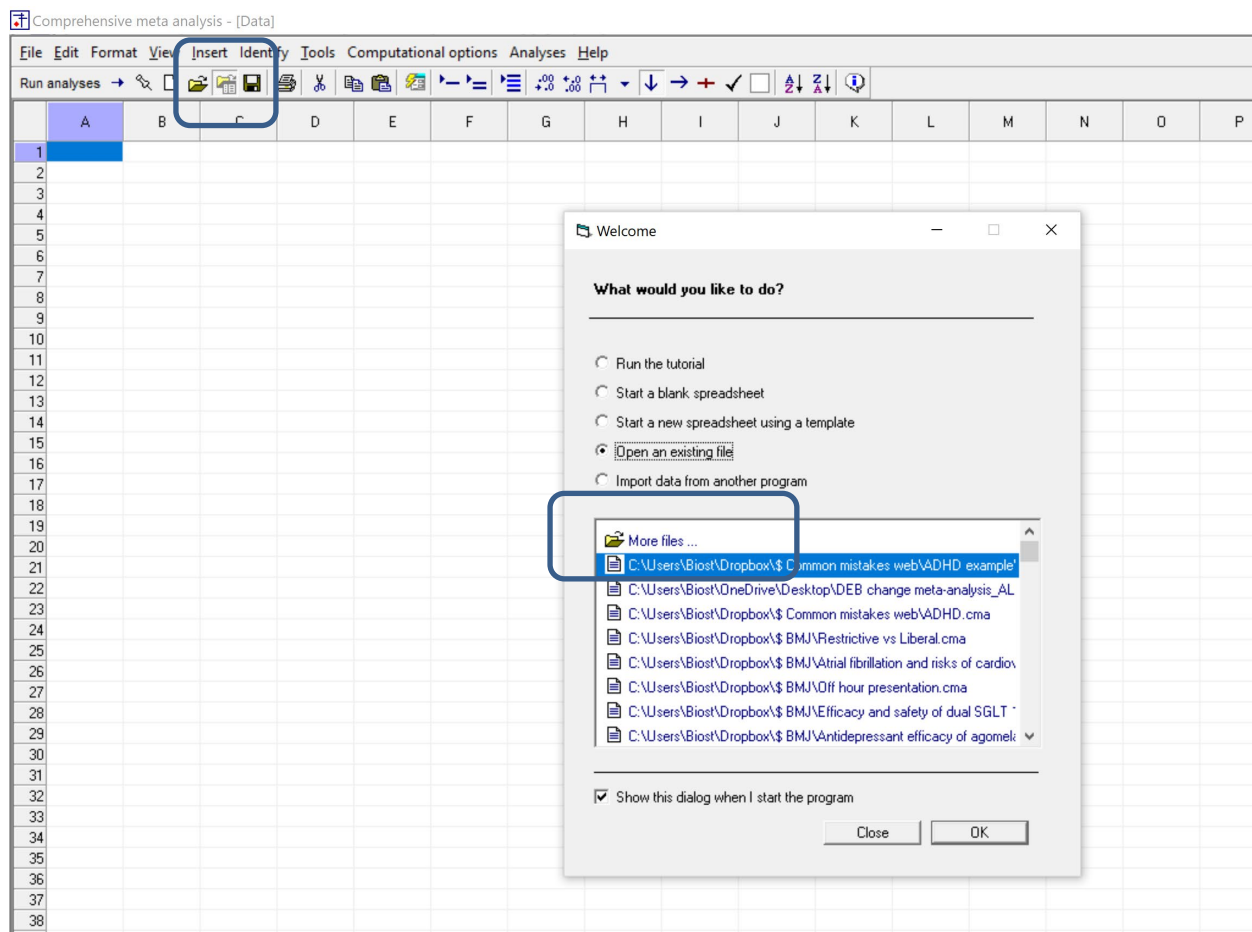
To create a new file from scratch, start here.

Start the program

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

Click Insert > Column for > Study names
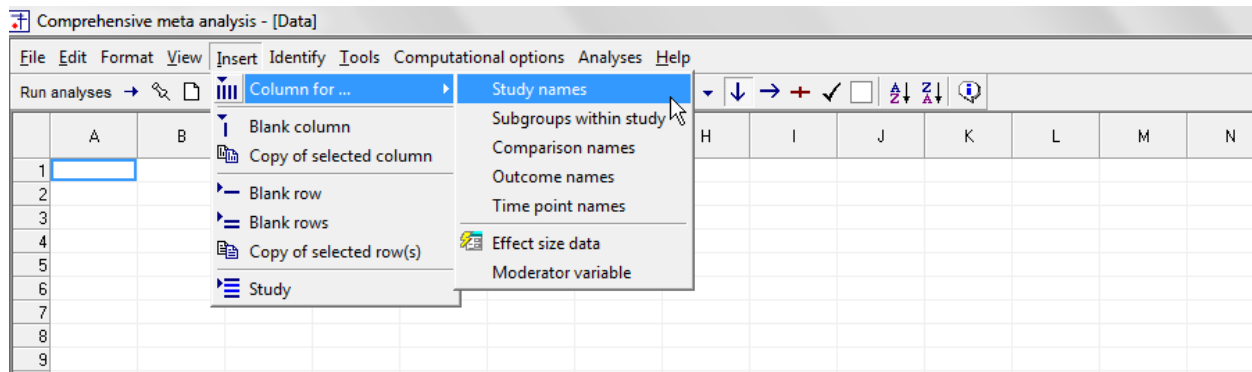
The screen should look like this
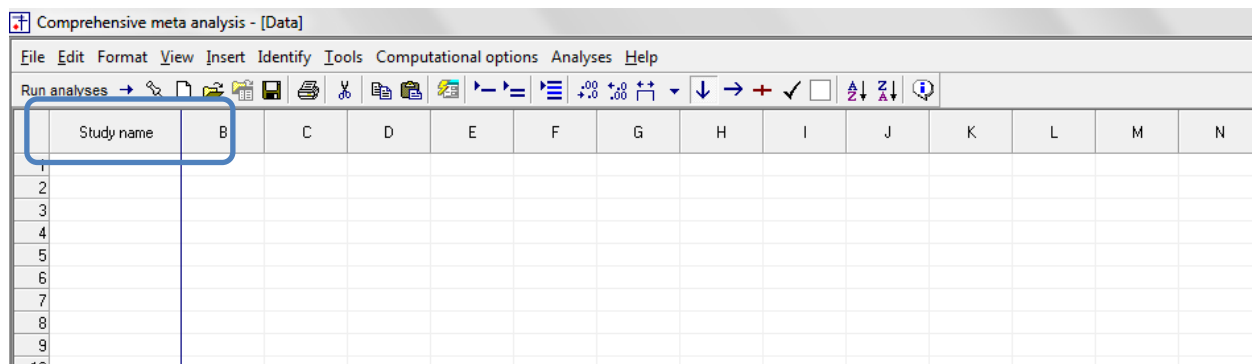
Click Insert > Column for > Effect size data
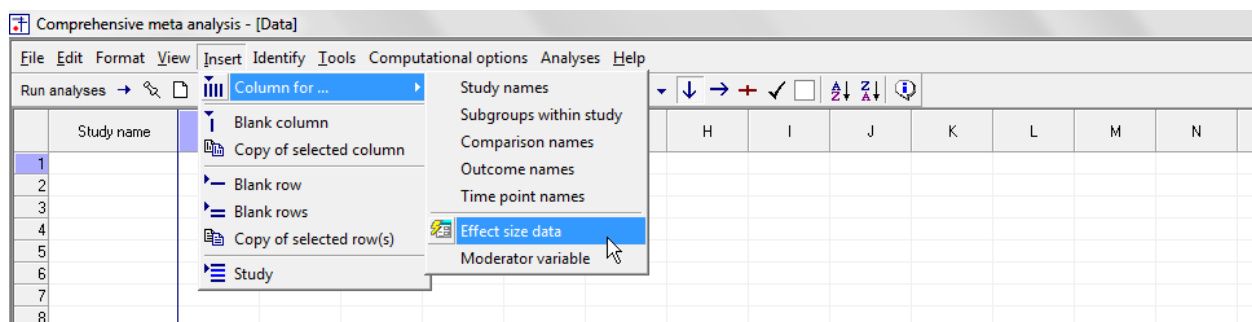
The program displays this wizard

Select [Show all 100 formats]
Click [Next]
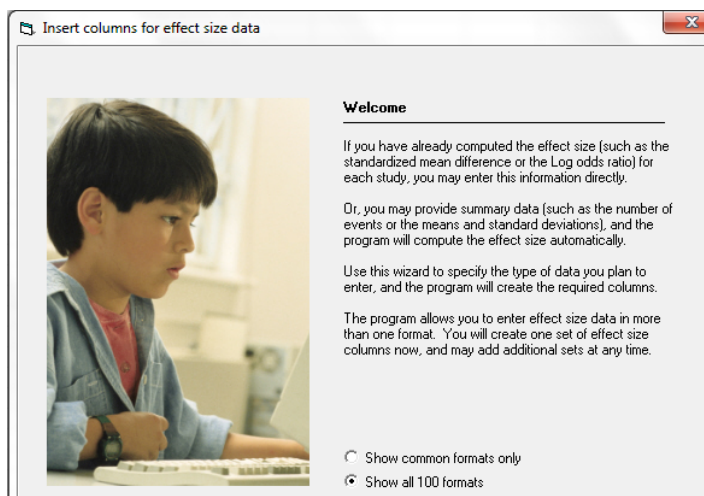
Select [Comparison of two groups…]
Click [Next]
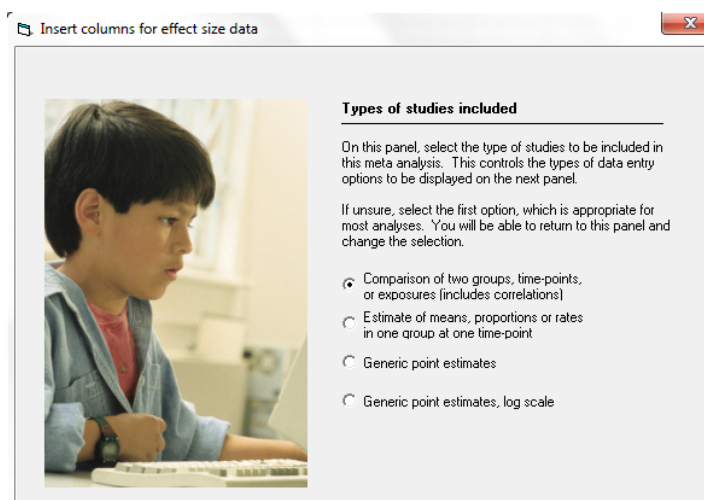
Drill down to

Continuous (means)
Unmatched groups, post-data only
Mean, SD and sample size in each group
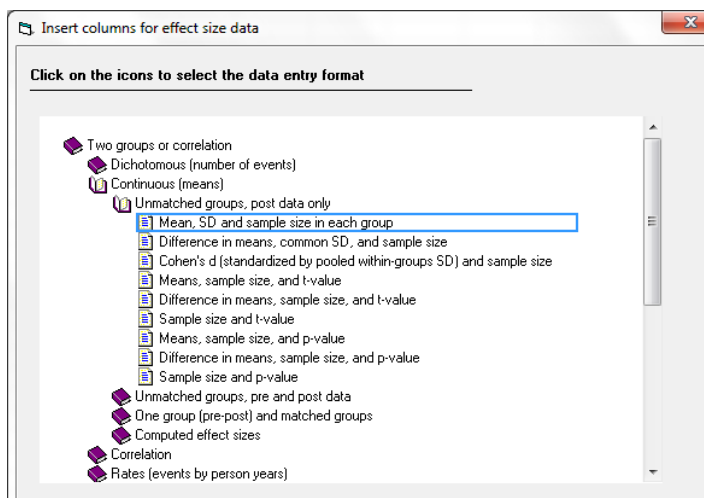
The program displays this wizard

Enter the following labels into the wizard

- First group > Trained
- Second group > Control

Click [Ok] and the program will copy the names into the grid

There are three options at this point

- Enter the data directly into CMA
- Copy the data from Excel "Preventing Alcohol Abuse.xls"

Here, we'll show how to copy the data from Excel

- Switch to Excel and open the file
- Highlight the rows and columns as shown (Columns A to G), and press CTRL-C to copy to clipboard
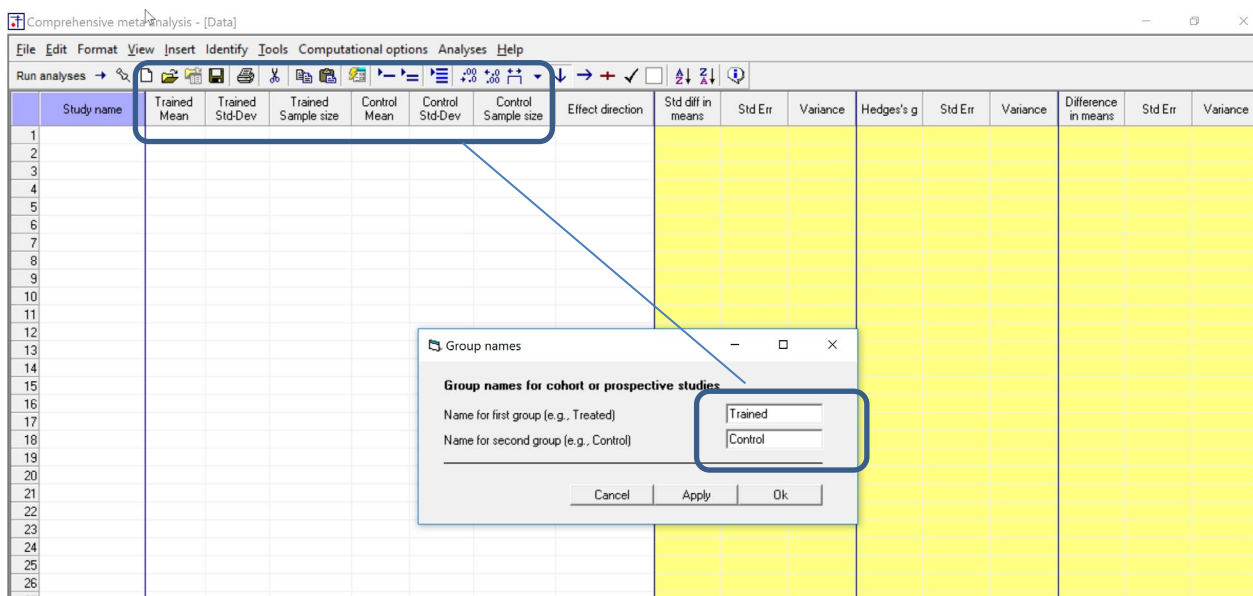
	A	B	C	D	E	F	G	H
		Tx Mean	Tx SD	Tx N	Ctrl Mean	Ctrl SD	Ctrl N	
1	Study	Tx Mean	Tx SD	Tx N	Ctrl Mean	Ctrl SD	Ctrl N	
2	Bendtsen, 2015	113.4000	81.1000	402	120.8000	86.4000	529	
3	Butler, 2009	12.1500	4.7652	30	17.2300	4.7421	26	
4	Cunningham, 2015	3.4800	3.4000	117	3.8400	3.1800	120	
5	Hester, 2012	13.2000	21.9000	42	23.1000	30.2000	39	
6	Kypri, 2008	41.5000	23.7500	138	50.0000	27.1900	146	
7	LaBrie, 2013	7.5000	7.3000	160	9.4000	10.2000	164	
8	Lewis, 2014	7.9100	8.5200	119	9.3100	8.4100	121	
9	Mason, 2014	8.3000	9.9000	8	3.6000	5.2000	10	
10	Palfai, 2011	4.3900	3.0300	57	4.3700	3.2300	62	
11	Ridout, 2014	19.7900	22.6700	47	32.5800	34.6400	48	
12	Suffoletto, 2012	3.1000	1.6000	14	3.8000	1.9000	13	
13	Suffoletto, 2014	3.5000	1.9933	384	4.0000	2.7576	185	
14	Voogt, 2013a	12.2000	15.1000	318	11.7000	14.0000	291	
15	Voogt, 2013b	21.6000	20.6000	456	22.5000	19.8000	451	
16	Witkiewitz, 2014	4.8300	2.5900	32	6.0500	2.8800	29	
17								

J18

- Switch to CMA
- Click in cell Study-name 1
- Press [CTRL-V] to paste the data
- The screen should look like this

Click here

Comprehensive meta analysis - [Data]

File Edit Format View Insert Identify Tools Computational options Analyses Help

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	Variance	Hedges's g	Std Err	Variance	Difference in means	Std Err	Variance
1	Study	Tx Mean	Tx SD	Tx N	Ctrl Mean	Ctrl SD	Ctrl N										
2	Bendtsen, 2015	113.400	81.100	402	120.800	86.400	529										
3	Butler, 2009	12.150	4.765	30	17.230	4.742	26										
4	Cunningham, 2015	3.480	3.400	117	3.840	3.180	120										
5	Hester, 2012	13.200	21.900	42	23.100	30.200	39										
6	Kypri, 2008	41.500	23.750	138	50.000	27.190	146										
7	LaBrie, 2013	7.500	7.300	160	9.400	10.200	164										
8	Lewis, 2014	7.910	8.520	119	9.310	8.410	121										
9	Mason, 2014	8.300	9.900	8	3.600	5.200	10										
10	Palfai, 2011	4.390	3.030	57	4.370	3.230	62										
11	Ridout, 2014	19.790	22.670	47	32.580	34.640	48										
12	Suffoletto, 2012	3.100	1.600	14	3.800	1.900	13										
13	Suffoletto, 2014	3.500	1.993	384	4.000	2.758	185										
14	Voogt, 2013a	12.200	15.100	318	11.700	14.000	291										
15	Voogt, 2013b	21.600	20.600	456	22.500	19.800	451										
16	Witkiewitz, 2014	4.830	2.590	32	6.050	2.880	29										
17																	
18																	

After checking that the data has been copied correctly, we can delete Row 1

- Click anywhere in Row 1
- Select Edit > Delete row, and confirm
- Your screen should look like this

Comprehensive meta analysis - [Data]

File Edit Format View Insert Identify Tools Computational options Analyses Help

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	Variance	Hedges's g	Std Err	Variance	Difference in means	Std Err	Variance
1	Bendtsen, 2015	113.400	81.100	402	120.800	86.400	529										
2	Butler, 2009	12.150	4.765	30	17.230	4.742	26										
3	Cunningham, 2015	3.480	3.400	117	3.840	3.180	120										
4	Hester, 2012	13.200	21.900	42	23.100	30.200	39										
5	Kypri, 2008	41.500	23.750	138	50.000	27.190	146										
6	LaBrie, 2013	7.500	7.300	160	9.400	10.200	164										
7	Lewis, 2014	7.910	8.520	119	9.310	8.410	121										
8	Mason, 2014	8.300	9.900	8	3.600	5.200	10										
9	Palfai, 2011	4.390	3.030	57	4.370	3.230	62										
10	Ridout, 2014	19.790	22.670	47	32.580	34.640	48										
11	Suffoletto, 2012	3.100	1.600	14	3.800	1.900	13										
12	Suffoletto, 2014	3.500	1.993	384	4.000	2.758	185										
13	Voogt, 2013a	12.200	15.100	318	11.700	14.000	291										
14	Voogt, 2013b	21.600	20.600	456	22.500	19.800	451										
15	Witkiewitz, 2014	4.830	2.590	32	6.050	2.880	29										
16																	

We need to enter a value for "Effect Direction"

Enter "Auto" for each study.

Click here

<Auto> tells the program to take the first mean minus the second. Since all studies used scales that worked in the same direction (where a higher score reflects more knowledge), this is appropriate here.

Comprehensive meta analysis - [Data]

File Edit Format View Insert Identify Tools Computational options Analyses Help

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	Variance	Hedges's g	Std Err	Variance	Difference in means	Std Err	Variance
1	Bendtsen, 2015	113.400	81.100	402	120.800	86.400	529	Auto	-0.088	0.066	0.004	-0.088	0.066	0.004	-7.400	5.568	31.003
2	Butler, 2009	12.150	4.765	30	17.230	4.742	26	Auto	-1.068	0.286	0.082	-1.054	0.282	0.080	-5.080	1.274	1.623
3	Cunningham, 2015	3.480	3.400	117	3.840	3.180	120	Auto	-0.109	0.130	0.017	-0.109	0.130	0.017	-0.360	0.428	0.183
4	Hester, 2012	13.200	21.900	42	23.100	30.200	39	Auto	-0.378	0.224	0.050	-0.374	0.222	0.049	-9.900	5.831	34.003
5	Kypri, 2008	41.500	23.750	138	50.000	27.190	146	Auto	-0.332	0.120	0.014	-0.331	0.119	0.014	-8.500	3.037	9.221
6	LaBrie, 2013	7.500	7.300	160	9.400	10.200	164	Auto	-0.214	0.111	0.012	-0.213	0.111	0.012	-1.900	0.988	0.975
7	Lewis, 2014	7.910	8.520	119	9.310	8.410	121	Auto	-0.165	0.129	0.017	-0.165	0.129	0.017	-1.400	1.093	1.194
8	Mason, 2014	8.300	9.900	8	3.600	5.200	10	Auto	0.617	0.485	0.236	0.587	0.462	0.214	4.700	3.615	13.070
9	Palfai, 2011	4.390	3.030	57	4.370	3.230	62	Auto	0.006	0.184	0.034	0.006	0.182	0.033	0.020	0.575	0.331
10	Ridout, 2014	19.790	22.670	47	32.580	34.640	48	Auto	-0.436	0.208	0.043	-0.432	0.206	0.042	-12.790	6.020	36.241
11	Suffoletto, 2012	3.100	1.600	14	3.800	1.900	13	Auto	-0.400	0.389	0.151	-0.388	0.377	0.142	-0.700	0.674	0.455
12	Suffoletto, 2014	3.500	1.993	384	4.000	2.758	185	Auto	-0.220	0.090	0.008	-0.220	0.090	0.008	-0.500	0.203	0.041
13	Voogt, 2013a	12.200	15.100	318	11.700	14.000	291	Auto	0.034	0.081	0.007	0.034	0.081	0.007	0.500	1.183	1.400
14	Voogt, 2013b	21.600	20.600	456	22.500	19.800	451	Auto	-0.045	0.066	0.004	-0.045	0.066	0.004	-0.900	1.342	1.801
15	Witkiewitz, 2014	4.830	2.590	32	6.050	2.880	29	Auto	-0.447	0.260	0.067	-0.441	0.256	0.066	-1.220	0.700	0.490
16																	
17																	
18																	
19																	

The program displays three effect sizes, as follows

Difference in means

This is the raw mean difference. When we can use this index, it's often a good choice, since it provides an intuitive measure of the effect. For example, if all studies reported the number of drinks per day, this might have been a useful index. In this case, however, we cannot use this index because different studies used different measures to assess outcome. For example, some recorded the number of drinks over a given time period, while others recorded the number of binge-drink episodes over a time period. Some used a scale that took account of both amounts and frequency. And so on. Additionally, some studies recorded behavior over a week while others recorded behavior over a month.

Std Diff in means

This is the standardized difference in means, sometimes called Cohen's d. It's defined as the raw difference in means divided by the standard deviation (computed within groups and pooled). This index transforms the outcomes for all studies onto a common metric. If one study reported the impact of treatment on the number of drinks, and another reported the impact of treatment on the number of binge-drinking events, these are both converted to a common metric, so they can be used in the same analysis.

Hedges' g

The standardized mean difference (above) has a slight bias – the effect size tends to be slightly larger (in absolute value) than the true value. The bias is small, and this is not generally a cause for concern. However, when we have the sample size (as we do here) we can remove the bias and create an adjusted value. This is also the standardized mean difference, but to distinguish between the two, we call the adjusted value Hedges' g.

In this analysis we will be using Hedges' g. To save space on the screen we will hide the raw difference and d, leaving only g on the display

- Right-click in any yellow column
- Click "Customize computed effect size display"

In the wizard,

- Select "Hedges' *g*" in the drop-down box
- Un-check "Std diff in means"
- Un-check "Difference in means"
- Put a check-mark adjacent to "Also show standard error"
- Put a check-mark adjacent to "Also show variance"
- Click [Ok]

The screen should look like this

Comprehensive meta analysis - [Data]

File Edit Format View Insert Identify Tools Computational options Analyses Help

Run analyses → ⚒ ☐ ☞ 📑 ⊟ ⏳ ✂ 📋 📋 🗔 ┉┉ ┉= ┇≡ .00 +.0 ↔ ▼ ↓ → ✛ ✓ ☐ ⓐ↓ ⓩ↓ ⓘ

	Study name	Treated Mean	Treated Std-Dev	Treated Sample size	Control Mean	Control Std-Dev	Control Sample size	Effect direction	Hedges's g	Std Err	Variance	L	M
1	Bendtsen, 2015	113.400	81.100	402	120.800	86.400	529	Auto	-0.088	0.066	0.004		
2	Butler, 2009	12.150	4.765	30	17.230	4.742	26	Auto	-1.054	0.282	0.080		
3	Cunningham, 2015	3.480	3.400	117	3.840	3.180	120	Auto	-0.109	0.130	0.017		
4	Hester, 2012	13.200	21.900	42	23.100	30.200	39	Auto	-0.374	0.222	0.049		
5	Kypri, 2008	41.500	23.750	138	50.000	27.190	146	Auto	-0.331	0.119	0.014		
6	LaBrie, 2013	7.500	7.300	160	9.400	10.200	164	Auto	-0.213	0.111	0.012		
7	Lewis, 2014	7.910	8.520	119	9.310	8.410	121	Auto	-0.165	0.129	0.017		
8	Mason, 2014	8.300	9.900	8	3.600	5.200	10	Auto	0.587	0.462	0.214		
9	Palfai, 2011	4.390	3.030	57	4.370	3.230	62	Auto	0.006	0.182	0.033		
10	Ridout, 2014	19.790	22.670	47	32.580	34.640	48	Auto	-0.432	0.206	0.042		
11	Suffoletto, 2012	3.100	1.600	14	3.800	1.900	13	Auto	-0.388	0.377	0.142		
12	Suffoletto, 2014	3.500	1.993	384	4.000	2.758	185	Auto	-0.220	0.090	0.008		
13	Voogt, 2013a	12.200	15.100	318	11.700	14.000	291	Auto	0.034	0.081	0.007		
14	Voogt, 2013b	21.600	20.600	456	22.500	19.800	451	Auto	-0.045	0.066	0.004		
15	Witkiewitz, 2014	4.830	2.590	32	6.050	2.880	29	Auto	-0.441	0.256	0.066		
16													
17													
18													

Click File > Save As and save the file

File Edit Format View Insert Identify Tools Computational options Analyses Help

Menu			Treated Std-Dev	Treated Sample size	Control Mean	Control Std-Dev	Control Sample size	Effect direction	Hedges's g	Std Err	Variance	L	M	N
New …	►		81.100	402	120.800	86.400	529	Auto	-0.088	0.066	0.004			
Open	Ctrl+O		4.765	30	17.230	4.742	26	Auto	-1.054	0.282	0.080			
Opening screen wizard			3.400	117	3.840	3.180	120	Auto	-0.109	0.130	0.017			
Import			21.900	42	23.100	30.200	39	Auto	-0.374	0.222	0.049			
Save	Ctrl+S		23.750	138	50.000	27.190	146	Auto	-0.331	0.119	0.014			
Save As…			7.300	160	9.400	10.200	164	Auto	-0.213	0.111	0.012			
Print…	Ctrl+P		8.520	119	9.310	8.410	121	Auto	-0.165	0.129	0.017			
Print setup…			9.900	8	3.600	5.200	10	Auto	0.587	0.462	0.214			
Exit			3.030	57	4.370	3.230	62	Auto	0.006	0.182	0.033			
10 Ridout, 2014	19.790		22.670	47	32.580	34.640	48	Auto	-0.432	0.206	0.042			
11 Suffoletto, 2012	3.100		1.600	14	3.800	1.900	13	Auto	-0.388	0.377	0.142			
12 Suffoletto, 2014	3.500		1.993	384	4.000	2.758	185	Auto	-0.220	0.090	0.008			
13 Voogt, 2013a	12.200		15.100	318	11.700	14.000	291	Auto	0.034	0.081	0.007			
14 Voogt, 2013b	21.600		20.600	456	22.500	19.800	451	Auto	-0.045	0.066	0.004			
15 Witkiewitz, 2014	4.830		2.590	32	6.050	2.880	29	Auto	-0.441	0.256	0.066			
16														
17														

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

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.
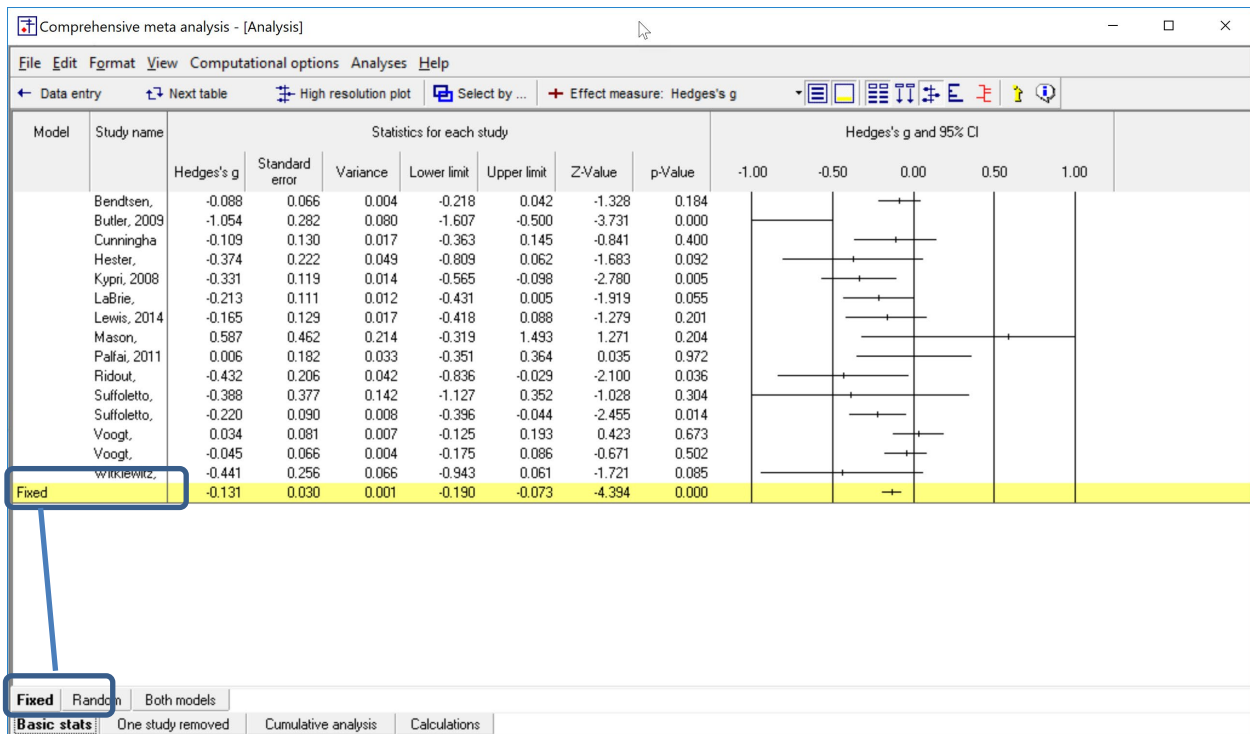
The fixed-effect model is appropriate when all studies are based on the same population, and are identical in all material respects. That's rarely the case, and is clearly not the case here.

Rather, we want to use the random-effects model. This model assumes that the studies in the analysis have been selected from a universe of studies that meet some criteria (as detailed by the description of Populations, Interventions, Outcomes, and Comparison groups in the full paper). The model also assumes that the effect sizes in these studies are a random selection of the effect sizes in the universe.

To switch to the random-effects model, click "Random" at the bottom of the screen.

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   ▾ 🔳 🔳 🔠 🔢 🔀 E 🔧 | ↑ ⓘ

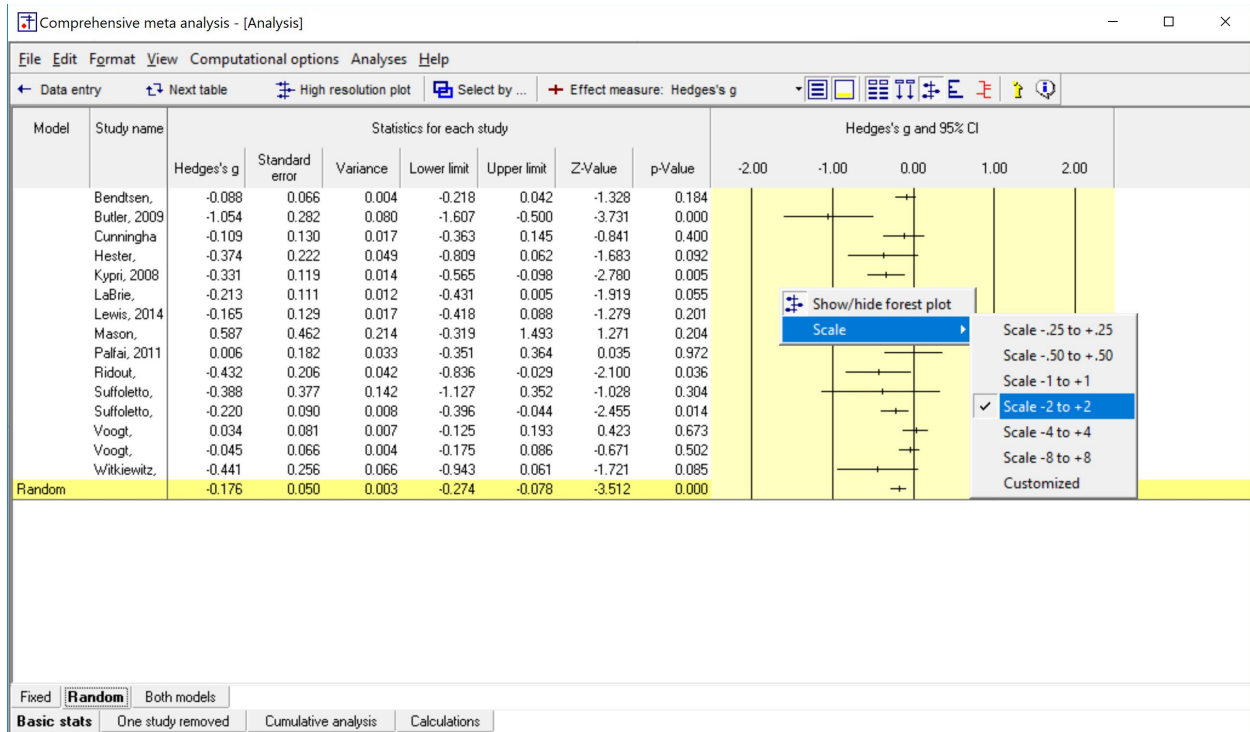| Model	Study name	Statistics for each study							Hedges's g and 95% CI				
		Hedges's g	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value	-1.00	-0.50	0.00	0.50	1.00
	Bendtsen,	-0.088	0.066	0.004	-0.218	0.042	-1.328	0.184					
	Butler, 2009	-1.054	0.282	0.080	-1.607	-0.500	-3.731	0.000					
	Cunningha	-0.109	0.130	0.017	-0.363	0.145	-0.841	0.400					
	Hester,	-0.374	0.222	0.049	-0.809	0.062	-1.683	0.092					
	Kypri, 2008	-0.331	0.119	0.014	-0.565	-0.098	-2.780	0.005					
	LaBrie,	-0.213	0.111	0.012	-0.431	0.005	-1.919	0.055					
	Lewis, 2014	-0.165	0.129	0.017	-0.418	0.088	-1.279	0.201					
	Mason,	0.587	0.462	0.214	-0.319	1.493	1.271	0.204					
	Palfai, 2011	0.006	0.182	0.033	-0.351	0.364	0.035	0.972					
	Ridout,	-0.432	0.206	0.042	-0.836	-0.029	-2.100	0.036					
	Suffoletto,	-0.388	0.377	0.142	-1.127	0.352	-1.028	0.304					
	Suffoletto,	-0.220	0.090	0.008	-0.396	-0.044	-2.455	0.014					
	Voogt,	0.034	0.081	0.007	-0.125	0.193	0.423	0.673					
	Voogt,	-0.045	0.066	0.004	-0.175	0.086	-0.671	0.502					
	Witkiewitz,	-0.441	0.256	0.066	-0.943	0.061	-1.721	0.085					
Random		-0.176	0.050	0.003	-0.274	-0.078	-3.512	0.000					

Fixed   **Random**   oth models
Basic stats   One study removed   Cumulative analysis   Calculations
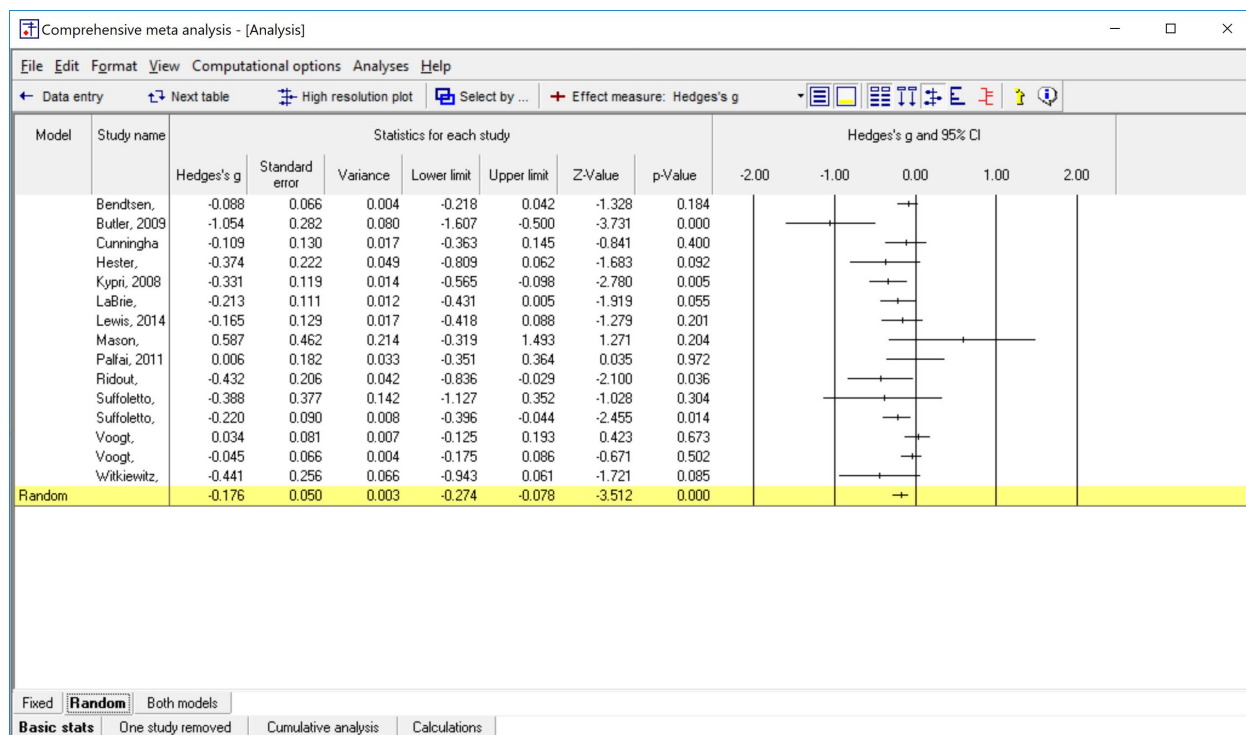
Next, we may want to modify the scale for the forest plot

At the moment, the scale goes from -1 to +1 and some studies (e.g., Butler) are not being displayed.

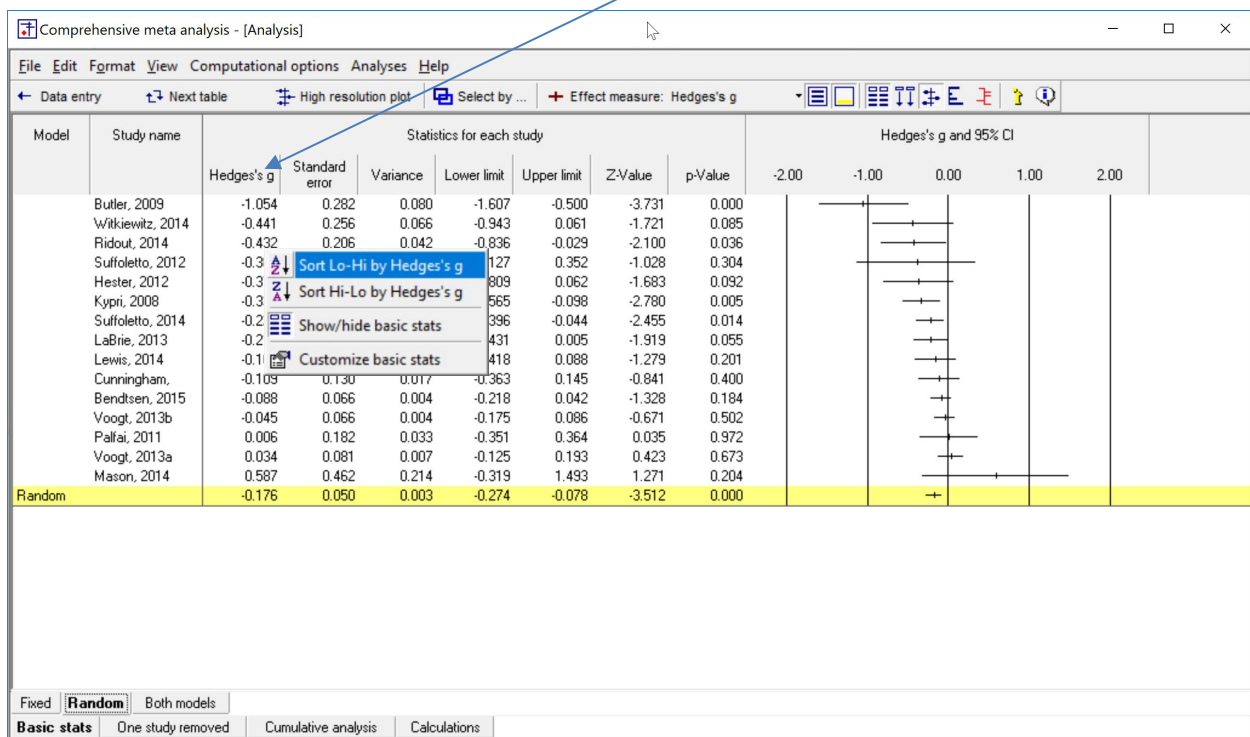- Right-click on the forest plot > Scale > Scale -2 to +2

The plot should look like this

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	Hedges's g	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value
	Bendtsen,	-0.088	0.066	0.004	-0.218	0.042	-1.328	0.184
	Butler, 2009	-1.054	0.282	0.080	-1.607	-0.500	-3.731	0.000
	Cunningha	-0.109	0.130	0.017	-0.363	0.145	-0.841	0.400
	Hester,	-0.374	0.222	0.049	-0.809	0.062	-1.683	0.092
	Kypri, 2008	-0.331	0.119	0.014	-0.565	-0.098	-2.780	0.005
	LaBrie,	-0.213	0.111	0.012	-0.431	0.005	-1.919	0.055
	Lewis, 2014	-0.165	0.129	0.017	-0.418	0.088	-1.279	0.201
	Mason,	0.587	0.462	0.214	-0.319	1.493	1.271	0.204
	Palfai, 2011	0.006	0.182	0.033	-0.351	0.364	0.035	0.972
	Ridout,	-0.432	0.206	0.042	-0.836	-0.029	-2.100	0.036
	Suffoletto,	-0.388	0.377	0.142	-1.127	0.352	-1.028	0.304
	Suffoletto,	-0.220	0.090	0.008	-0.396	-0.044	-2.455	0.014
	Voogt,	0.034	0.081	0.007	-0.125	0.193	0.423	0.673
	Voogt,	-0.045	0.066	0.004	-0.175	0.086	-0.671	0.502
	Witkiewitz,	-0.441	0.256	0.066	-0.943	0.061	-1.721	0.085
Random		-0.176	0.050	0.003	-0.274	-0.078	-3.512	0.000

Hedges's g and 95% CI: -2.00 -1.00 0.00 1.00 2.00

Fixed Random Both models
Basic stats   One study removed   Cumulative analysis   Calculations

To get a better sense of the dispersion we can sort the studies based on their effect sizes.

To sort by any column, right-click on that column (Right-click here) and click <Sort>

Model	Study name	Hedges's g	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value
	Butler, 2009	-1.054	0.282	0.080	-1.607	-0.500	-3.731	0.000
	Witkiewitz, 2014	-0.441	0.256	0.066	-0.943	0.061	-1.721	0.085
	Ridout, 2014	-0.432	0.206	0.042	-0.836	-0.029	-2.100	0.036
	Suffoletto, 2012	-0.3				127	0.352	-1.028	0.304
	Hester, 2012	-0.3				809	0.062	-1.683	0.092
	Kypri, 2008	-0.3				565	-0.098	-2.780	0.005
	Suffoletto, 2014	-0.2				396	-0.044	-2.455	0.014
	LaBrie, 2013	-0.2				431	0.005	-1.919	0.055
	Lewis, 2014	-0.1				418	0.088	-1.279	0.201
	Cunningham,	-0.109	0.130	0.017	-0.363	0.145	-0.841	0.400
	Bendtsen, 2015	-0.088	0.066	0.004	-0.218	0.042	-1.328	0.184
	Voogt, 2013b	-0.045	0.066	0.004	-0.175	0.086	-0.671	0.502
	Palfai, 2011	0.006	0.182	0.033	-0.351	0.364	0.035	0.972
	Voogt, 2013a	0.034	0.081	0.007	-0.125	0.193	0.423	0.673
	Mason, 2014	0.587	0.462	0.214	-0.319	1.493	1.271	0.204
Random		-0.176	0.050	0.003	-0.274	-0.078	-3.512	0.000

(Context menu shown: Sort Lo-Hi by Hedges's g / Sort Hi-Lo by Hedges's g / Show/hide basic stats / Customize basic stats)

The studies have been coded so that a negative difference means that the intervention reduced the amount of alcohol use.

The mean effect size is −0.176, but this is only an estimate of the true mean for the 15 studies in the analysis, and THAT (the true mean for these 15 studies) is only an estimate of the true mean for all studies in the relevant universe (the universe from which these studies were sampled.
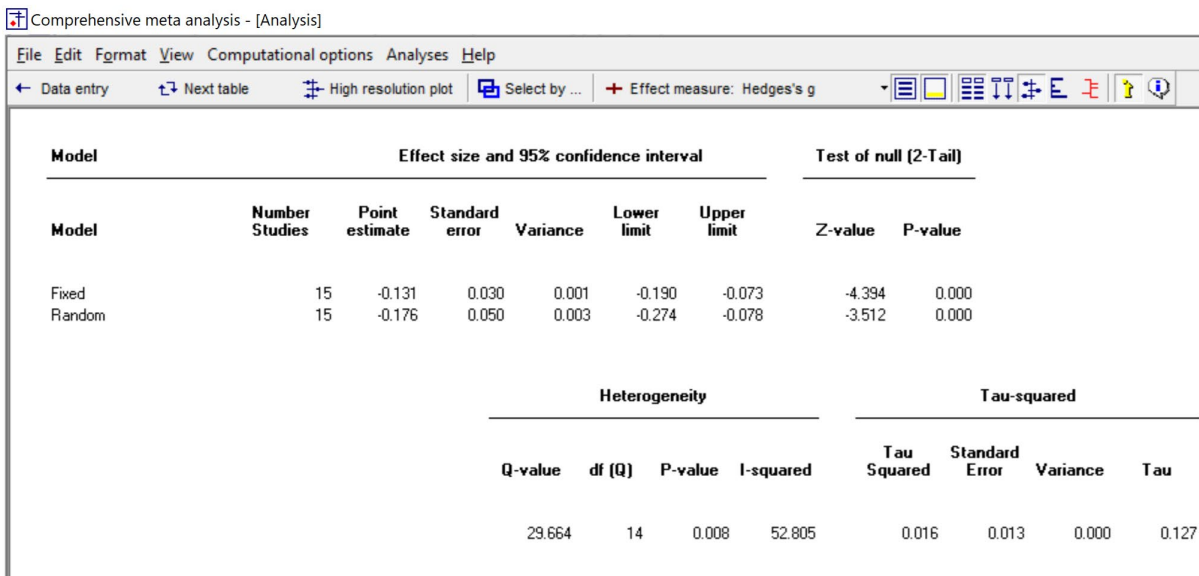
The confidence interval is −0.274 to −0.078. In the universe from which these studies were sampled, the true *mean* probably falls somewhere in that range.

People who are familiar with the standardized mean difference will have an intuitive sense of what this effect size means (and most would feel that it's a relatively small effect). For those who are not familiar with this effect size, we can provide an analog in raw units. Specifically, if we are looking at the number of drinks per week, and the standard deviation is 5.0 drinks, this would correspond to a decrease of around (−0.176 x 5.0) one drink per week. Or, if we are looking at the number of binges over two months and the standard deviation is 4.0, this would correspond to a decrease of (−0.176 x 4.0) less than one binge over that time span.

Heterogeneity

To this point we've addressed the mean effect size. We need to also consider how much the effects vary from study to study.

To see the relevant statistics, click <Next Table>

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		Effect size and 95% confidence interval						Test of null (2-Tail)		
Model	Number Studies	Point estimate	Standard error	Variance	Lower limit	Upper limit		Z-value	P-value	
Fixed	15	-0.131	0.030	0.001	-0.190	-0.073		-4.394	0.000	
Random	15	-0.176	0.050	0.003	-0.274	-0.078		-3.512	0.000	

	Heterogeneity				Tau-squared			
	Q-value	df (Q)	P-value	I-squared	Tau Squared	Standard Error	Variance	Tau
	29.664	14	0.008	52.805	0.016	0.013	0.000	0.127

The relevant statistics are as follows. Q is 29.664 with 14 degrees of freedom and a p-value of 0.008. The I^2 statistic is 52.805%. T^2 (Tau-squared) is 0.016, and T (Tau) is 0.127.

I'll start by presenting these statistics in the way that they are typically presented. Then I'll explain why these statistics are largely irrelevant, and suggest a better approach.

The Q-value is the sum of squared deviations of all effect-sizes from the mean effect size. This value is on a standardized scale, so that a large deviation gets more weight if the estimate is precise, and less weight if the estimate is imprecise. Under the null hypothesis that the true effect size is precisely the same in all studies (and that all of the variation in observed effects is due to sampling error), the expected value of Q would be equal to the degrees of freedom (the number of studies minus 1).

In this analysis, Q is 29.664 with 14 degrees of freedom. Under the null hypothesis that the treatment has precisely the same impact in all studies, Q would follow a chi-squared distribution with 14 df. So by referring to a table of chi-square we can determine that the p-value for a test of this null hypothesis is 0.008. We reject the null hypothesis, and conclude that the treatment is more effective in some populations than in others.

There is a widespread belief that the I^2 statistic tells us how much the effect size varies across studies. This is a fundamental misunderstanding of what this statistic means.

In fact, I^2 tells us what proportion of the variance in observed effects is due to variance in true effects rather than sampling error. As such, it provides some context for understanding the forest plot. If I^2 is low (near zero), then most of the variance in the forest plot is due to sampling error. If we could somehow plot the variance of true effects (which is what we are about), most of the variance would disappear. Conversely, if I^2 is high (say, more than 75%) then most of the variance in the forest plot is due to variance in true effects. If we could somehow plot the variance of true effects (which is what we are about), most of the variance would remain.

In this case, I^2 is around 53%, so the plot of observed effects gives us an exaggerated idea of how the true effects are actually distributed.

The variance of true effects (T^2) is 0.016, and the standard deviation of true effects (T) is 0.127.

All of these statistics are accurate. But (with the exception of T) they all are largely irrelevant. When we ask about heterogeneity, we intend to ask "Over what range do the effects vary", and none of these statistics directly addresses this question.

The p-value provides a test of the null hypothesis that the effects are all the same. There is not really any reason to test this hypothesis, since common sense tells us that the intervention will not have precisely the same impact in all populations. In any event, the p-value only tells us that the effects vary *at all* – it does not tell us *how much* they vary.

Similarly, I^2 tells us that most of the variation that we see in the observed effects reflects variation in true effects, rather than sampling error. This (along with the plot) provides some idea of the variation, but it's a rough idea, at best. Based on I^2 and the plot, it's unlikely that any two people will arrive at the same estimate of the dispersion.

T^2 is the variance of true effects. As such, it plays the same role as the variance of scores in a primary study. It's an important statistic, but does not directly address the question that we care about.

Finally, we have T, the standard deviation of true effects. This statistic is the one that most directly speaks to the dispersion of effects. If we had a precise estimate of the mean effect size and a precise estimate of T, and we assume that the effects are normally distributed about the mean, then the true effect size in most studies would fall roughly within two T on either side of the mean.

As a practical matter, the estimate of the mean and the estimate of T are subject to error. Therefore, we apply a modified version of this formula to compute the prediction interval. This formula is programmed into the Excel file Prediction_Intervals

	AutoSave ● off	🖫 ⤺ ⤻ ▾	prediction_intervals.xls - Compatibility Mode - Excel			
	File	Home	Insert	Page Layout	Formulas	Data	Review	View	ACROBAT	♀ Tell me what you want to do		

C8   fx   0.0161418126565087

	A	B	C	D	F
1		**Prediction interval for D, d, g, RD**			
2					
3		Number of studies	15		
6		Mean effect (random effect weights)	-0.1761		
7		Upper limit of mean effect	-0.0778		
8		Tau-squared	0.0161		
10					
11		Prediction interval			
12		Mean	-0.1761		
13		Prediction interval (95%) lower limit	-0.4712		
14		Prediction interval (95%) upper limit	0.1190		
15					
16					

Open the Excel file

- At the bottom, select the tab for [Means]
- Copy the four values needed for the blue cells, as shown here.
- The Excel file displays the prediction interval as -0.4712 to +0.1190

This tells us that there will be some populations where the intervention reduces alcohol use by nearly one-half a standard deviation. And some where it increases alcohol use slightly.

Explore the study weights

We might want to explore the study weights, to see how much weight each study contributed to the mean effect size.

Click the button for Show Weights
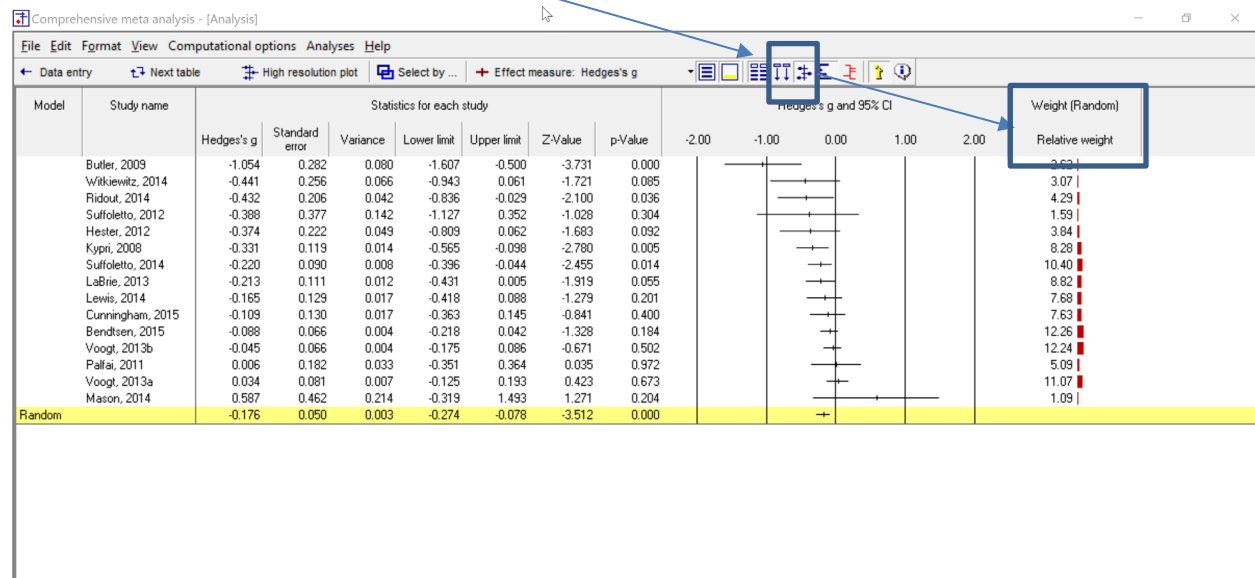
The program displays the relative weight assigned to each study, along with a bar graph on a scale of 0 to 1 corresponding to the percentage of weight (these sum to 1.0).

Right-click on the column of weights to customize the display. For example, you may elect to display the raw weight rather than the relative weight. Or (as shown here), you may elect to scale relative to the maximum weight as shown here. The scale will then go (in this case) from 0 to 0.13 rather than 0 to 1, which makes it easier to distinguish among the study weights.

Model	Study name	Hedges's g	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value	Relative weight
	Butler, 2009	-1.054	0.282	0.080	-1.607	-0.500	-3.731	0.000	2.62
	Witkiewitz, 2014	-0.441	0.256	0.066	-0.943	0.061	-1.721	0.085	3.07
	Ridout, 2014	-0.432	0.206	0.042	-0.836	-0.029	-2.100	0.036	4.29
	Suffoletto, 2012	-0.388	0.377	0.142	-1.127	0.352	-1.028	0.304	1.59
	Hester, 2012	-0.374	0.222	0.049	-0.809	0.062	-1.683	0.092	3.84
	Kypri, 2008	-0.331	0.119	0.014	-0.565	-0.098	-2.780	0.005	8.28
	Suffoletto, 2014	-0.220	0.090	0.008	-0.396	-0.044	-2.455	0.014	10.40
	LaBrie, 2013	-0.213	0.111	0.012	-0.431	0.005	-1.919	0.055	8.82
	Lewis, 2014	-0.165	0.129	0.017	-0.418	0.088	-1.279	0.201	7.68
	Cunningham, 2015	-0.109	0.130	0.017	-0.363	0.145	-0.841	0.400	7.63
	Bendtsen, 2015	-0.088	0.066	0.004	-0.218	0.042	-1.328	0.184	12.26
	Voogt, 2013b	-0.045	0.066	0.004	-0.175	0.086	-0.671	0.502	12.24
	Palfai, 2011	0.006	0.182	0.033	-0.351	0.364	0.035	0.972	5.09
	Voogt, 2013a	0.034	0.081	0.007	-0.125	0.193	0.423	0.673	11.07
	Mason, 2014	0.587	0.462	0.214	-0.319	1.493	1.271	0.204	1.09
Random		-0.176	0.050	0.003	-0.274	-0.078	-3.512	0.000	

We can see here that there are some ten studies that share the bulk of the weight, with the other five playing a relatively minor role.

My take-home woud be that the analysis is not driven by one or two studies, which gives me confidence that the effects are reasonably robust. It's also reassuring that the two outliers (one at either end) get relatively little weight in the analysis.

We want to create a high-resolution plot. On the current screen we will specify which columns to display or hide. Then, we'll move to the high-resolution screen.

In the high-resolution plot we will use a box for each effect size, and the size of the box will reflect that study's weight in the analysis. Therefore, we want to hide the column of weight. To hide the column of weights, click here
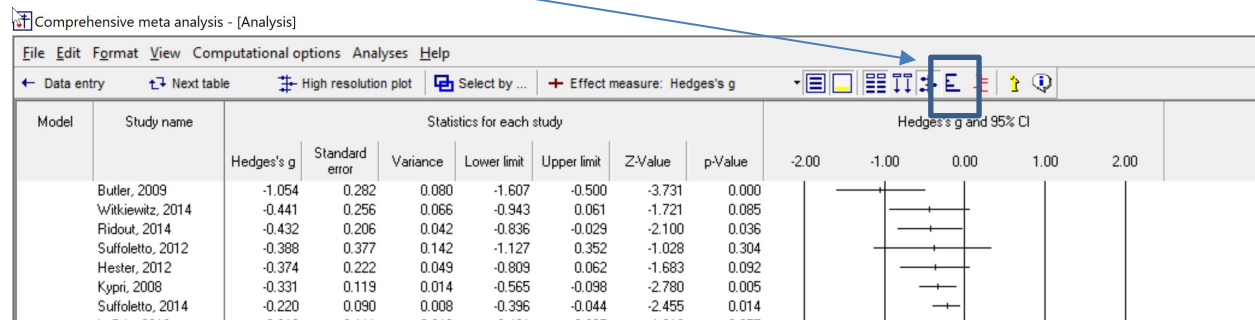
Right-click on the statistics, then select Customize basic stats. This will allow us to choose which columns to include in the high-resolution plot
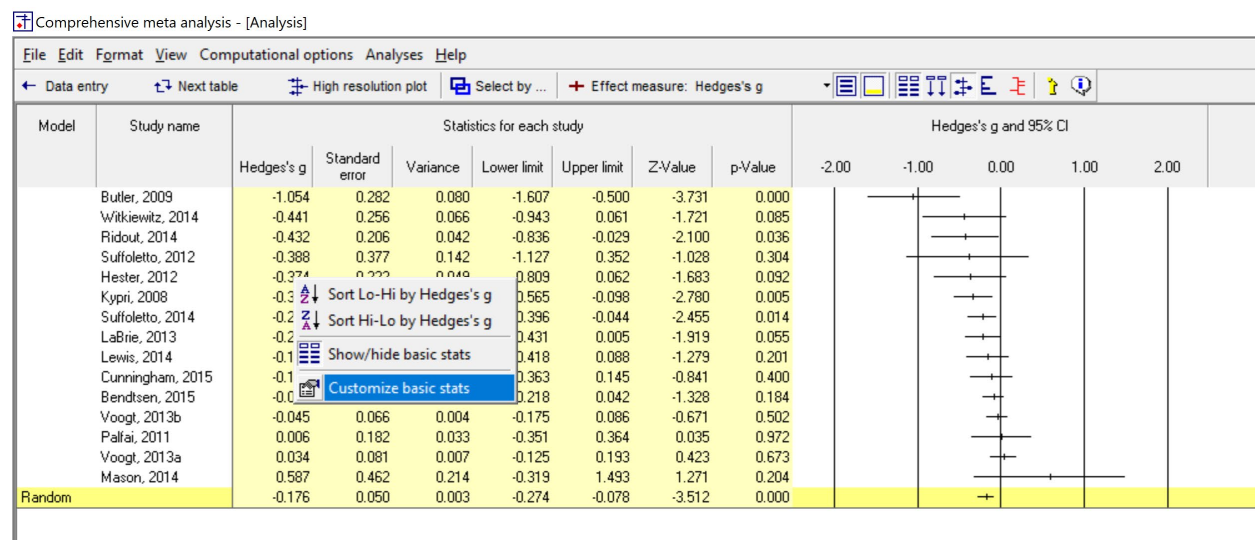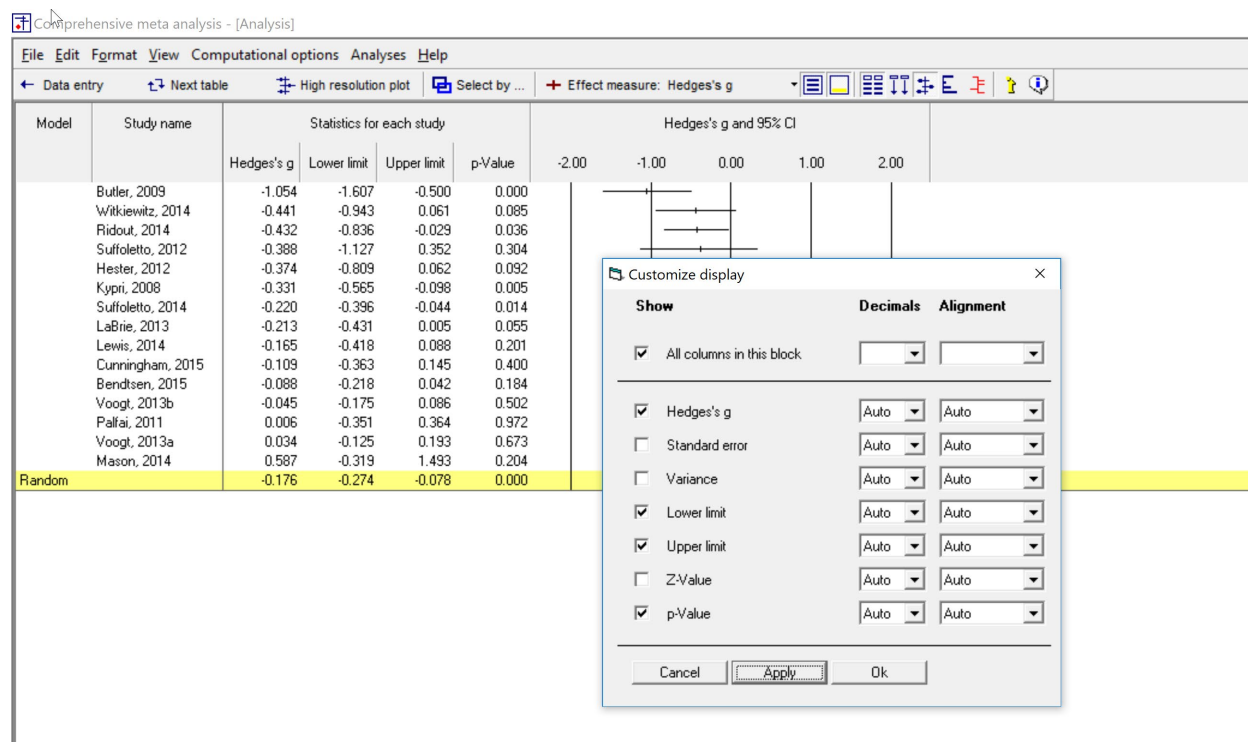
I've chose to keep the mean, lower and upper limit, and p-value. I've removed the standard error, variance, and *Z*-value. This will make the plot more readable.

Click <High resolution plot>
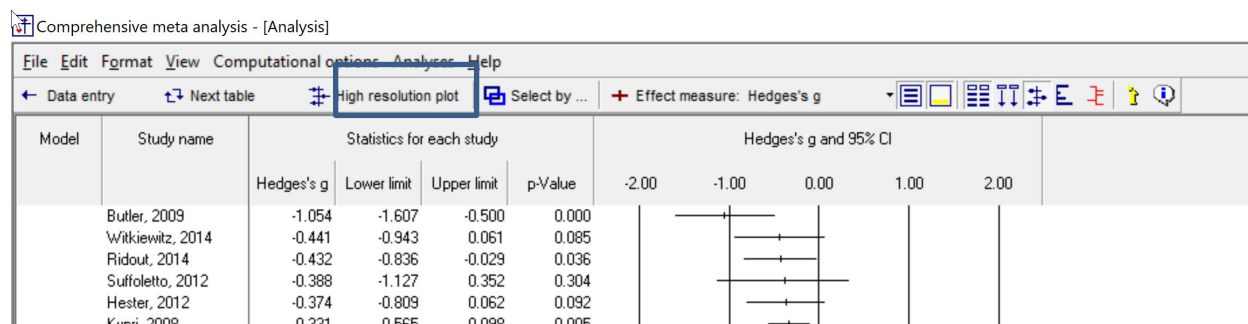
The program displays the plot.

Initialize the plot

Click Reset All to set all the options to their defaults
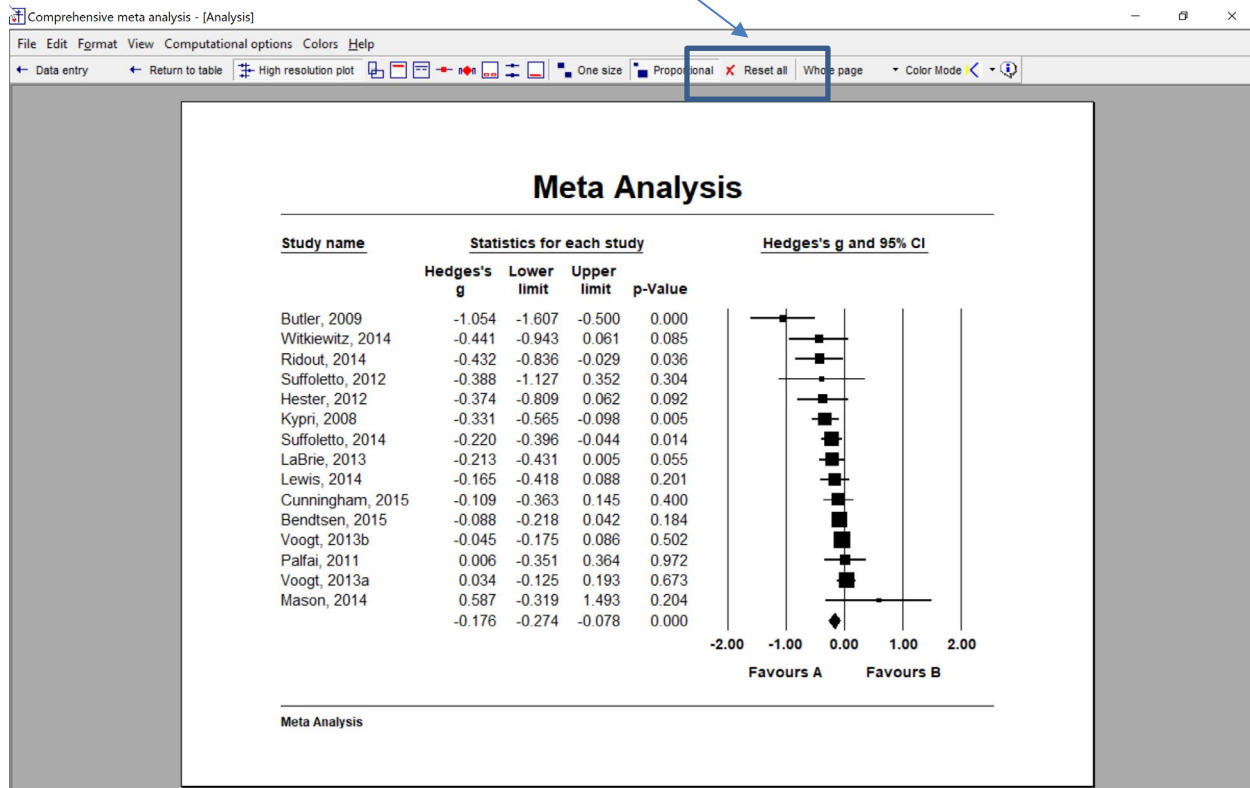
Modify the spacing

Right-click on the forest plot, and select <Spacing and forest plot width>
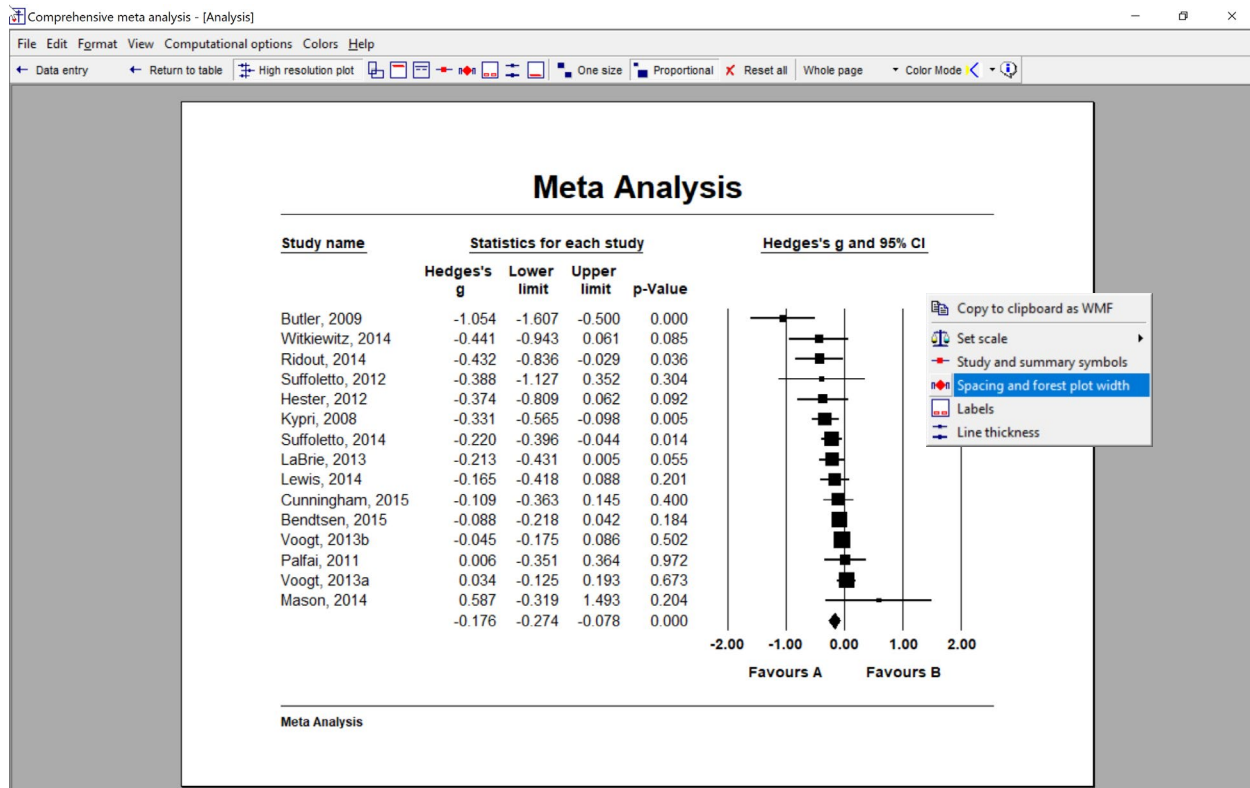
Adjust the width of the forest plot

- Select the tab for Right buffer
- Click Remove.
- This will remove space to the right of the plot

- Select the tab for Forest plot width
- Click the <Single up Arrow> twice
- This will expand the space alloted to the plot
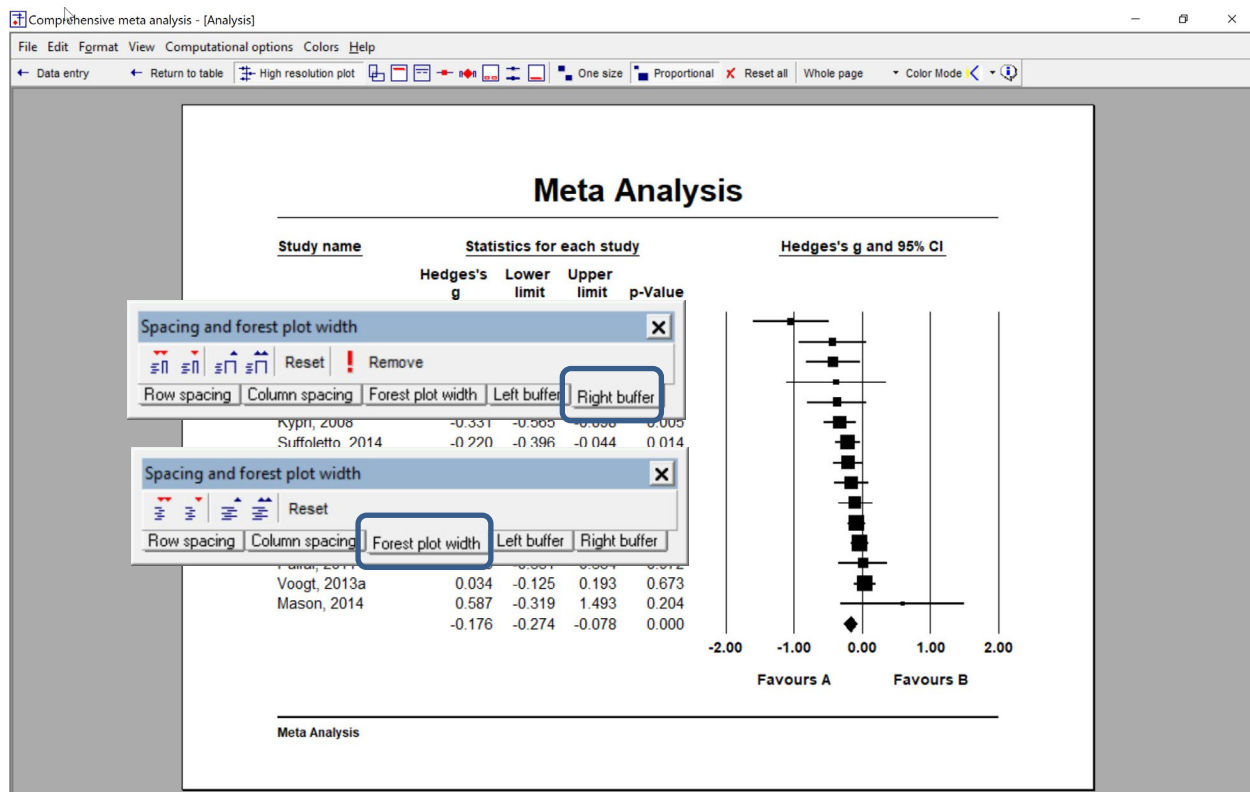
Edit the title and labels

- Right-click on the Title (or click Edit > Header on the menu).
- Enter a new title
- Click <Apply>
- Click Font and increase the font size by one <Up>

- Right-click on the Labels (or click Edit > Labels on the menu).
- Enter new labels <Favors Treatment> <Favors Control>
- Click <Apply>
- Click Font and increase the font size by one <Up>
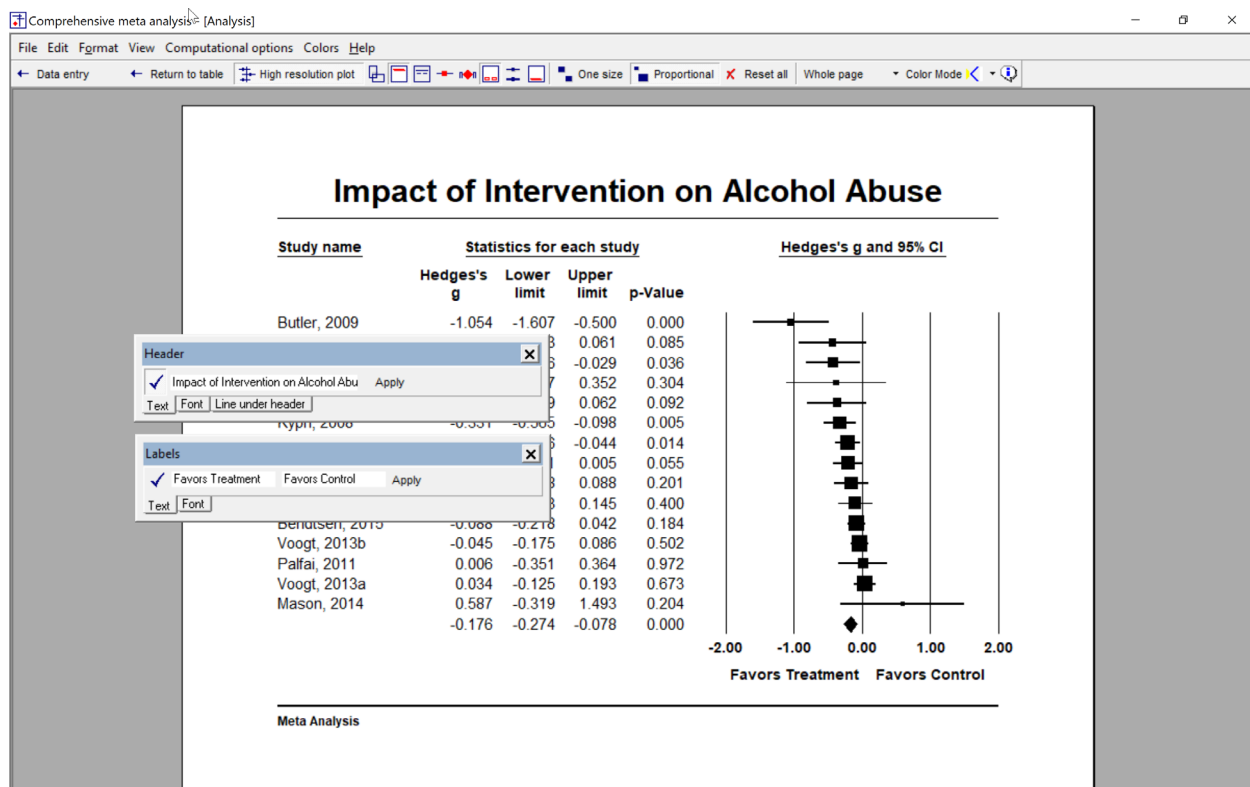
Edit the footer

- Right-click on the Title (or click Edit > Footer on the menu).
- Enter a new footer
- Click <Apply>
- Click Font and increase the font size by one <Up>
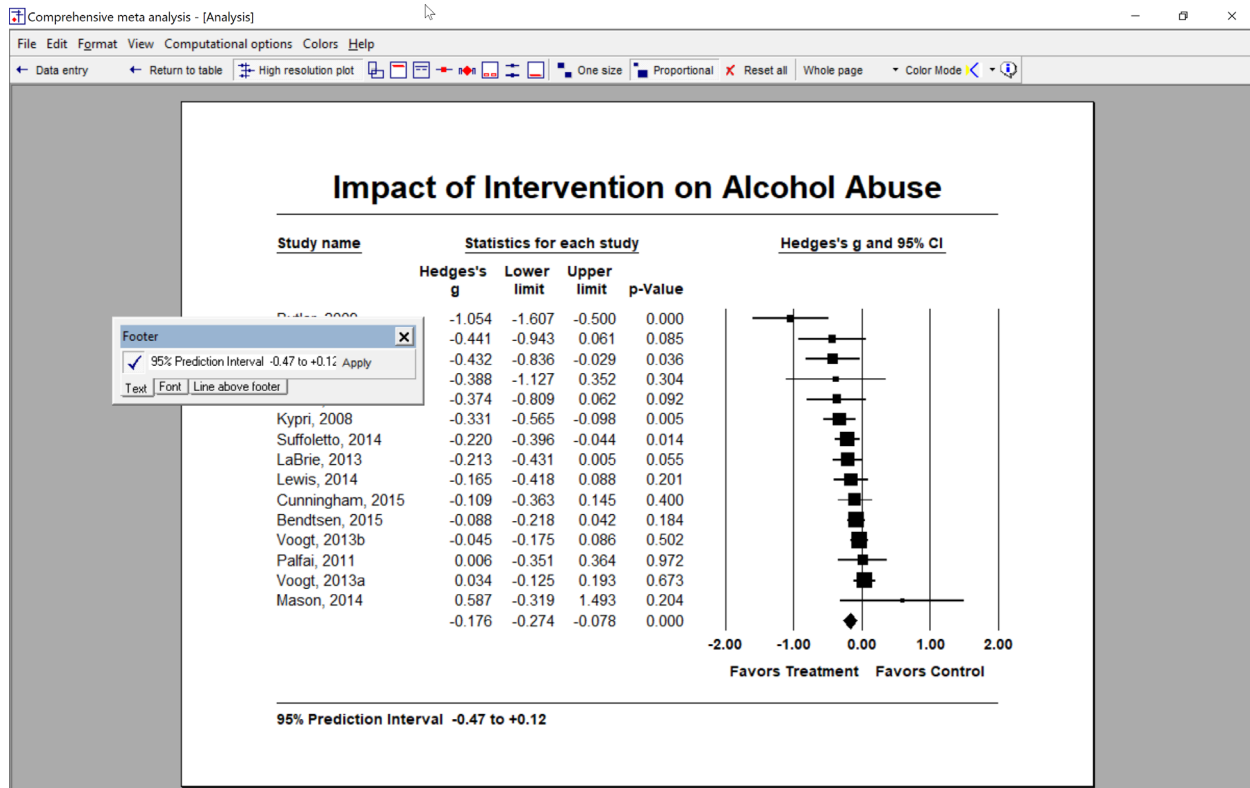
Export to Word

- Click File > Export to Word
- The program created a Word file and inserts the plot
- Look for the Word icon on your taskbar
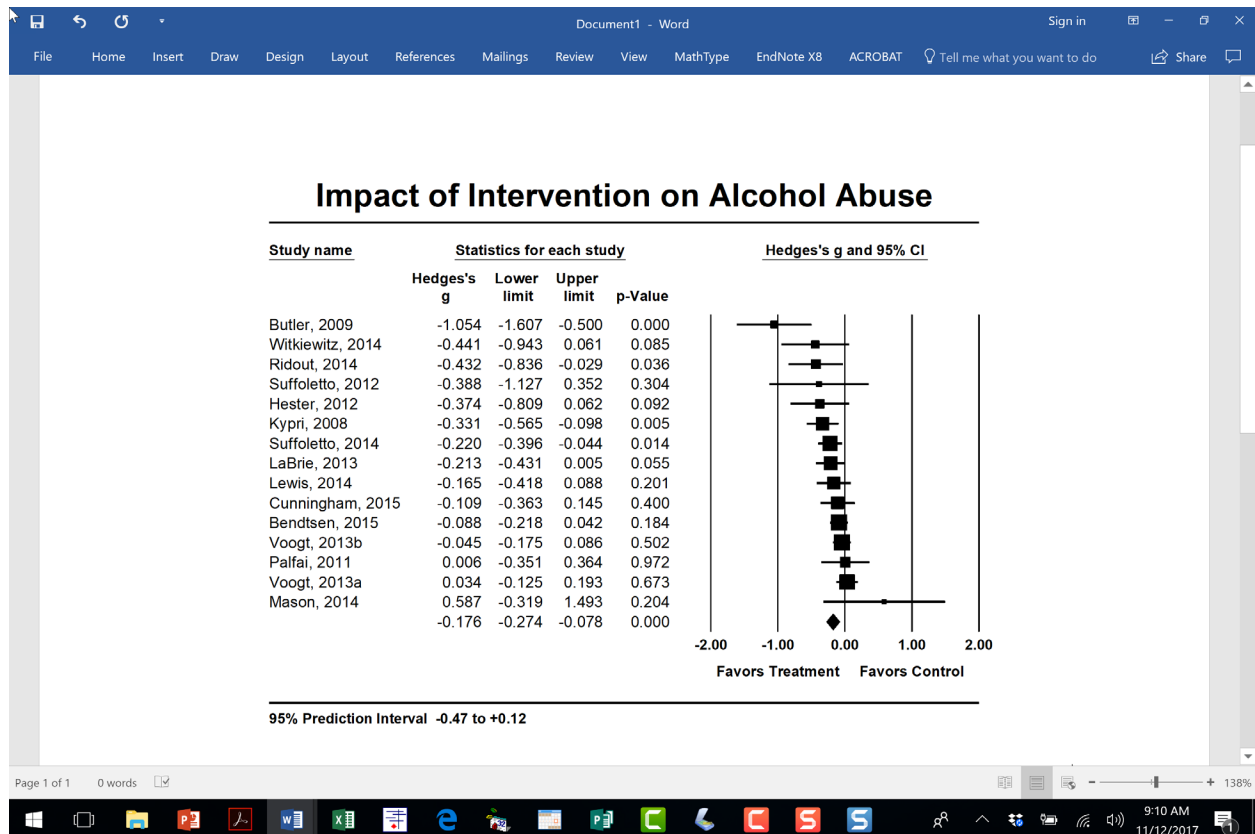
Export to PowerPoint

- If you're in Word, return to CMA
- Select colors for slides
- Click File > Export to PowerPoint
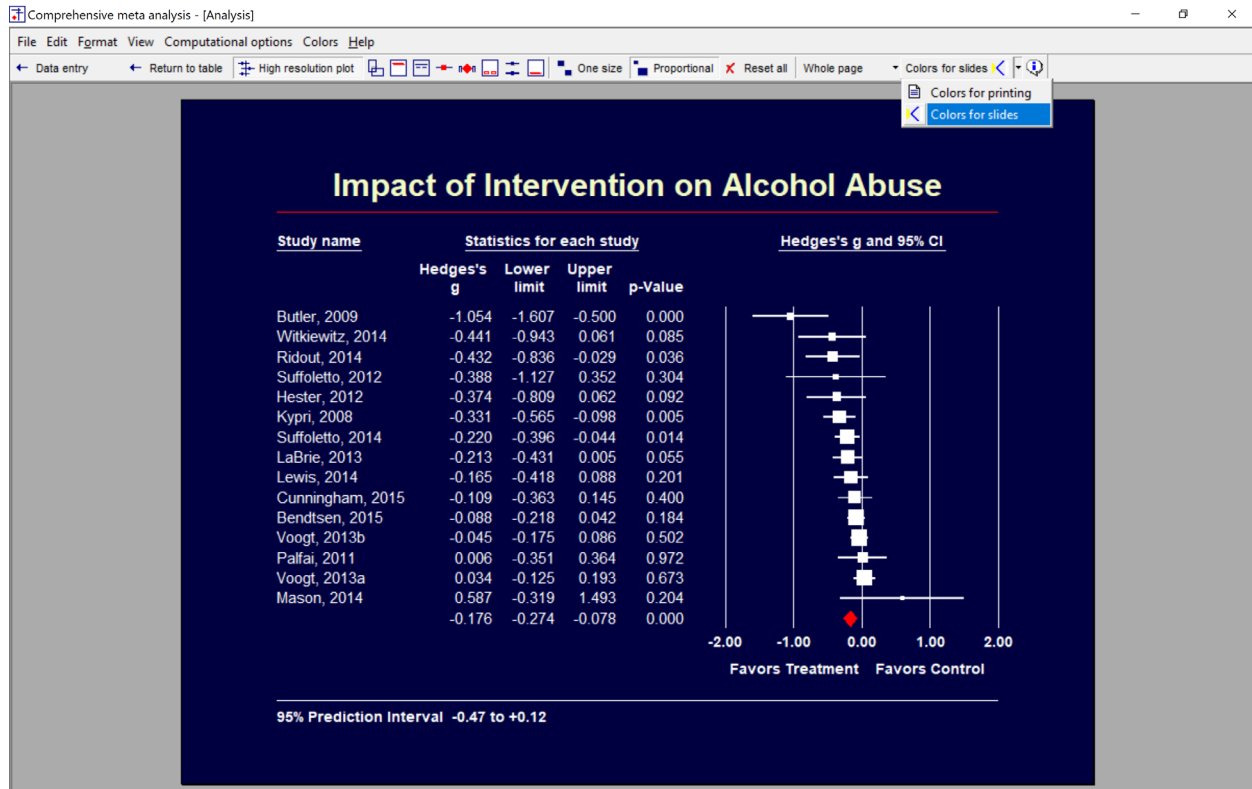
You can also set the background to any color. This can be useful if you want to match the colors to another set of slides

- To modify the colors for any element of the plot, click Colors on the menu
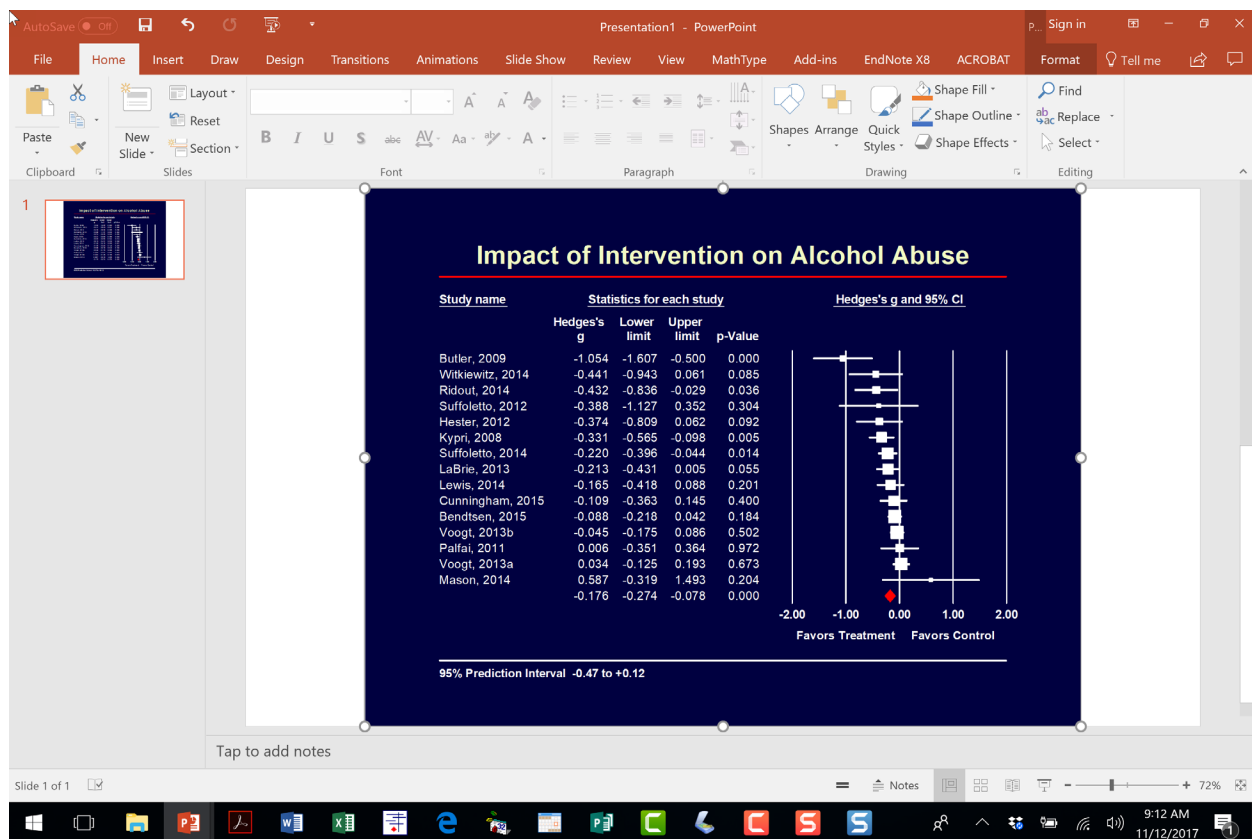
This screen shows the plot in PowerPoint

In some versions of PowerPoint, there will be space on either side of the plot

In that case, you may want to change the background to match the background on the plot.
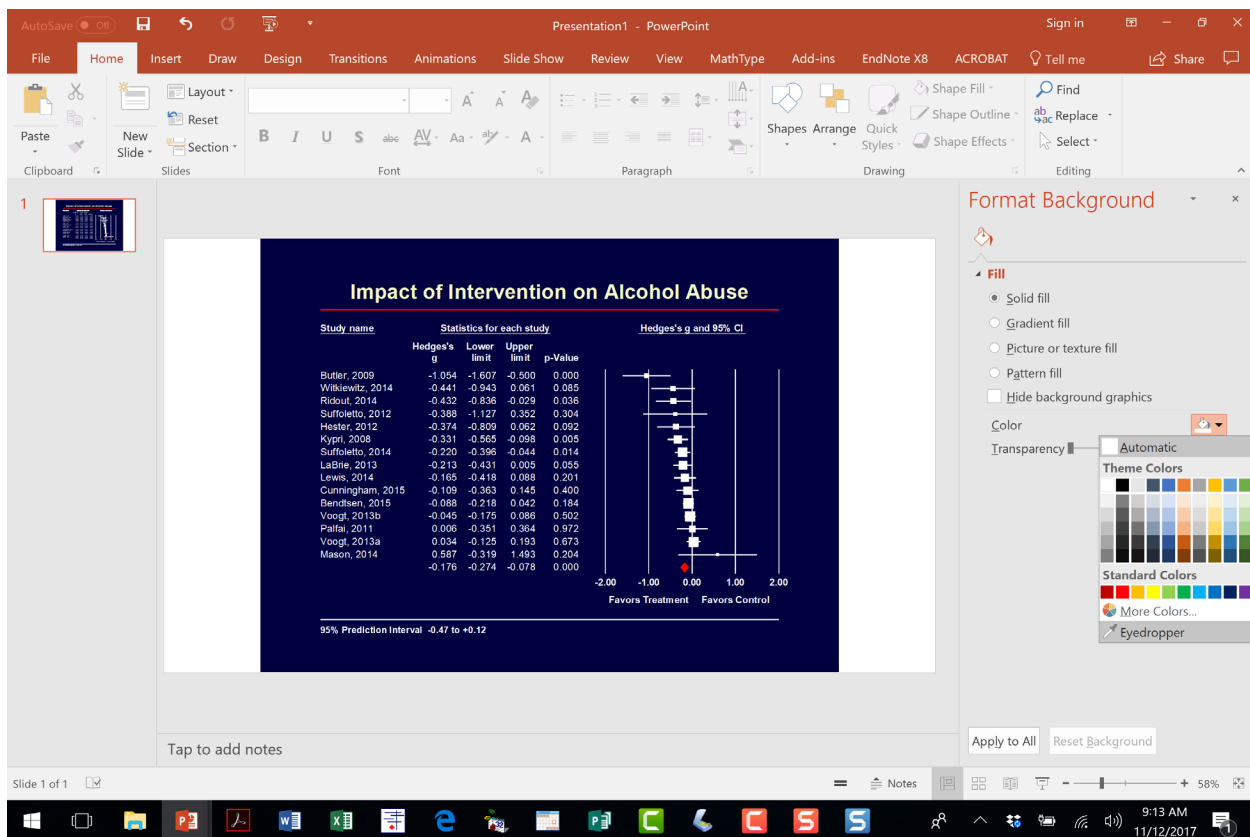
- Right-click on the background of the plot
- Choose Format background
- Under Fill, choose color
- Click on the Eyedropper
- Click on the blue section of the plot to select that color

The result is shown below

How to report the analysis

Overview

The analysis is based on fifteen studies that evaluated the impact of a computerized intervention on alcohol use. Each study compared alcohol use for young adults who had been randomized either to the intervention or to a control condition. The effect size is the standardized difference in means, Hedges' *g*. The results of this analysis will be generalized to comparable studies, and so the random-effects model was employed for the analysis.

Does the intervention decrease alcohol use?

The standardized difference in means is −0.176. On average, those who received the training scored 0.18 standard deviations lower (less drinking) than those who had not. On a scale that counts the number of drinks per unit of time, with a standard deviation of 10 drinks, this would correspond to a difference of 1.8 drinks. The confidence interval for the effect size is −0.274 to −0.078, which tells us that in the universe of comparable studies the effect size could fall anywhere in this range. This range does not include a difference of zero, which tells us that the true effect size is probably not zero. Similarly, the *Z*-value for testing the null hypothesis (that the effect size is 0.0) is −3.512, with a corresponding *p*-value of < 0.001. We can reject the null hypothesis, and conclude that (on average) the training does reduce alcohol use in the universe of populations which are comparable to those in the analysis. Given the dispersion in effects (as discussed below), it is important to recognize that the mean effect size applies to this particular mix of studies, and would be different for another mix of populations, dosages, and so on.

How much does the effect size vary across populations?

The *Q*-statistic provides a test of the null hypothesis that all studies in the analysis share a common effect size. If all studies shared the same effect size, the expected value of *Q* would be equal to the degrees of freedom (the number of studies minus 1). The *Q*-value is 29.664 with 14 degrees of freedom and a *p*-value of 0.008. We reject the null hypothesis that the true effect size is identical in all the studies. The I^2 statistic is 53%, which tells us that 53% of the variance in observed effects reflects variance in true effects rather than sampling error. The variance of true effects (T^2) is 0.016, and the standard deviation of true effects (*T*) is 0.127. The 95% prediction interval is −0.471 to +0.119. In the universe of populations represented by these studies, the true effect size in 95% of cases will fall somewhere in this range. In this context, we will assume that 0.20 is a small effect and 0.50 is a moderate effect. On that basis, the mean effect size is small. The prediction interval extends as far as −0.47, so the effect size may be called moderate in a small minority of populations.