“Combining heterogeneous studies using the random-effects model is a mistake and leads to inconclusive meta-analyses”

Suhail A. R. Doi
Clinical Epidemiology Unit
SPH, UQ

What is a meta-analysis?

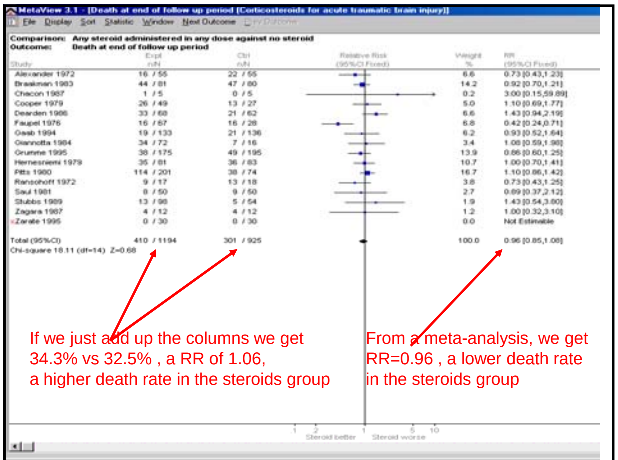
- A weighted average of the effects in the trials under consideration
- This means that instead of each of the studies contributing equally to the final average, some studies contribute more than others
- Normalized weights are used (weights summing to 1)
- In a fixed effects approach this assumes of course that trials are exchangeable

The Inverse Variance Weight

- Studies generally vary in size.
- An ES based on 100 subjects is assumed to be a more “precise” estimate of the population ES than is an ES based on 10 subjects.
- Therefore, larger studies should carry more “weight” in our analyses than smaller studies.
- Simple approach: weight each ES by its sample size.
- Better approach: weight by the inverse variance.

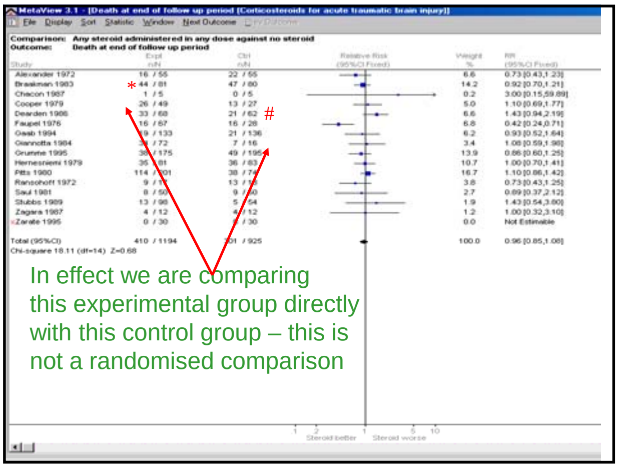
Could we just add the data from all the trials together?

- One approach to combining trials would be to add all the treatment groups together, add all the control groups together, and compare the totals
- This is wrong for several reasons, and it can give the wrong answer
- This is because imbalances between trials may produce bias



Study	Expt n/N	Ctrl n/N	Weight %	RR (95%CI Fixed)
Alexander 1972	16 / 55	22 / 55	6.6	0.73 [0.43,1.23]
Braakman 1983	44 / 81	47 / 80	14.2	0.92 [0.70,1.21]
Chacon 1987	1 / 5	0 / 5	0.2	3.00 [0.15,59.89]
Cooper 1979	26 / 49	13 / 27	5.0	1.10 [0.69,1.77]
Dearden 1986	33 / 68	21 / 62	6.6	1.43 [0.94,2.19]
Faupel 1976	16 / 67	16 / 28	6.8	0.42 [0.24,0.71]
Gaab 1994	19 / 133	21 / 136	6.2	0.93 [0.52,1.64]
Giannotta 1984	34 / 72	7 / 16	3.4	1.08 [0.59,1.98]
Grumme 1995	38 / 175	49 / 195	13.9	0.86 [0.60,1.25]
Hernesniemi 1979	35 / 81	36 / 83	10.7	1.00 [0.70,1.41]
Pitts 1980	114 / 201	38 / 74	16.7	1.10 [0.86,1.42]
Ransohoff 1972	9 / 17	13 / 18	3.8	0.73 [0.43,1.25]
Saul 1981	8 / 50	9 / 50	2.7	0.89 [0.37,2.12]
Stubbs 1989	13 / 98	5 / 54	1.9	1.43 [0.54,3.80]
Zagara 1987	4 / 12	4 / 12	1.2	1.00 [0.32,3.10]
Zarate 1995	0 / 30	0 / 30	0.0	Not Estimable
Total (95%CI)	410 / 1194	301 / 925	100.0	0.96 [0.85,1.08]

Chi-square 18.11 (df=14) Z=0.68



Closer look at the Inverse Variance Weight

The standard error (SE) is a direct index of ES precision.

SE is used to create confidence intervals.

The smaller the SE, the more precise the ES.

Hedges' showed that the optimal weights for meta-analysis are:

$$w = \frac{1}{SE^2}$$

Inverse Variance Weight for the Major Effect Sizes

Logged Odds-Ratio:

$$se = \sqrt{\frac{1}{a} + \frac{1}{b} + \frac{1}{c} + \frac{1}{d}} \qquad w = \frac{1}{se^2}$$

Where a, b, c, and d are the cell frequencies of a 2 by 2 contingency table.

Logged RR:

$$se = \sqrt{\frac{(1-P_T)}{P_T n_T} + \frac{(1-P_C)}{P_C n_C}} \qquad w = \frac{1}{se^2}$$

The Weighted Mean Effect Size

Study	ES	w
1	-0.33	11.91
2	0.32	28.57
3	0.39	58.82
4	0.31	29.41
5	0.17	13.89
6	0.64	8.55
7	-0.33	9.80
8	0.15	10.75
9	-0.02	83.33
10	0.00	14.93

Start with the effect size (ES) and inverse variance weight (w) for 10 studies.

$$\overline{ES} = \frac{\sum (w \times ES)}{\sum w}$$

The Standard Error of the Mean ES

Study	ES	w	w*ES
1	-0.33	11.91	-3.93
2	0.32	28.57	9.14
3	0.39	58.82	22.94
4	0.31	29.41	9.12
5	0.17	13.89	2.36
6	0.64	8.55	5.47
7	-0.33	9.80	-3.24
8	0.15	10.75	1.61
9	-0.02	83.33	-1.67
10	0.00	14.93	0.00
		269.96	41.82

The standard error of the mean is the square root of 1 divided by the sum of the weights.

$$se_{\overline{ES}} = \sqrt{\frac{1}{\sum w}} = \sqrt{\frac{1}{269.96}} = 0.061$$

Homogeneity Analysis

Homogeneity analysis may test

- H0_a: treatment effect is zero AND identical in every trial.
- H0_b: treatment effect is identical to the weighted average in every trial

H0_a (according to Senn) is tested by the fixed effects analysis and requires the principle of exchangeability of trials

H0_b is tested using Cochran's Q statistic

H0_b can be rejected regardless of the FE results

After such rejection:

- if FE is null this means that a large study was null and
- if not it means that the smaller studies were null.

Two options after H0_b is rejected:

- model between study differences
- fit a random effects model

Why fit a random model

We have rejected H0_b therefore we need to estimate the mean of the distribution of true effects that will tell us what is the likely result of a similar future trial of the type represented by the studies in the meta-analysis

H0_b is tested using

$$Q = \sum w_i (T_i - \overline{T})^2,$$

Where T is the treatment effect in each trial and $\overline{T}$ is the fixed effect inverse variance weighted estimate.

Random Effects Models

Total Q is significant and you assume that the excess variability across effect sizes derives from random differences across studies (sources you **have not** identified or measured – this is different from sources you **cannot** identify or measure)

The Logic of a Random Effects Model

Random effects model assumes that the variability between effect sizes is due to sampling error ***plus*** variability in the population of effects (unique differences in the set of true population effect sizes)

In other words, instability in an effect size is due to subject-level "noise" and true unmeasured differences across studies (that is, each study is estimating a slightly different population effect size)

The Basic Procedure of a Random Effects Model

Fixed effects model weights each study by the inverse of the sampling variance.

$$w_i = \frac{1}{se_i^2}$$

Random effects model weights each study by the inverse of the sampling variance ***plus*** a constant that represents the variability across the population effects.

$$w_i = \frac{1}{se_i^2 + \tau^2}$$

This is the random effects variance component.

How To Estimate the Random Effects Variance Component

The random effects variance component is based on Q.

The formula is:

$$\tau^2 = \frac{Q_T - (k-1)}{c}$$

Rerun Analysis with New Inverse Variance Weight

Add the random effects variance component to the variance associated with each ES.

$$w_i = \frac{1}{se_i^2 + \tau^2}$$

Calculate a new weight.

Rerun analysis.

A very complex statistical analysis is done.

Comparison of Random Effect with Fixed Effect Results

The biggest difference you will notice is in the significance levels and confidence intervals.

- Confidence intervals will get bigger.
- Effects that were significant under a fixed effect model may no longer be significant.

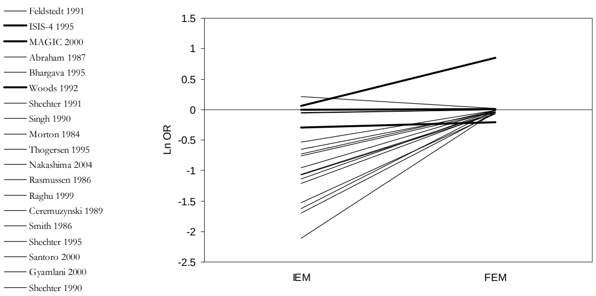
Random effects models are therefore more conservative **sometimes**.

Unfortunately, because of the random nature of the random model, this may not necessarily be true

Example 1

Start off with a heterogenous group of studies and run a fixed effect analysis

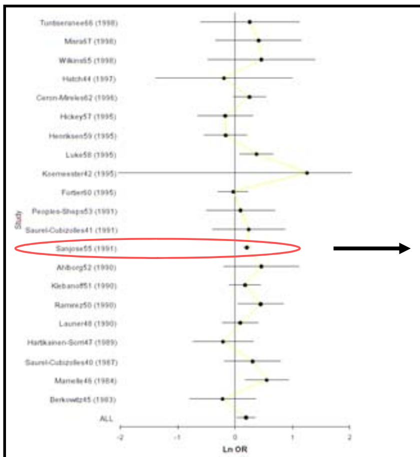
OR 1.01 (95% CI 0.96-1.07)



ISIS-4 = 78%

Example 2

Start off with a homogenous group of studies and create heterogeneity



Study size in rest of studies is 83 to 7100 subjects

104262 subjects

Study 9

Figure 1

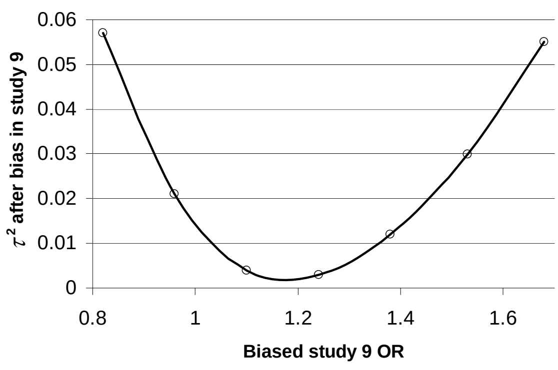


Figure 2

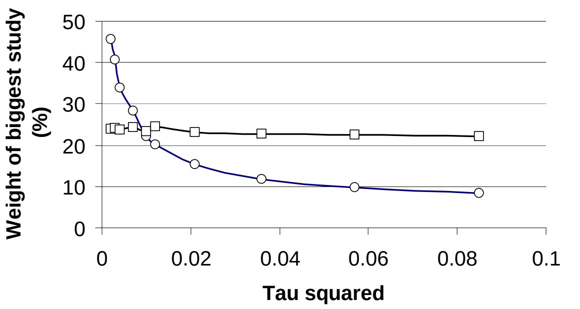


Figure 3

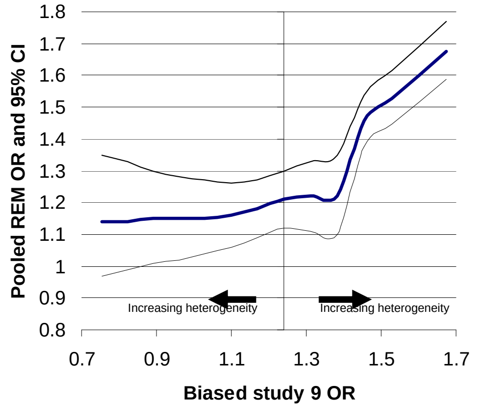


Figure 7

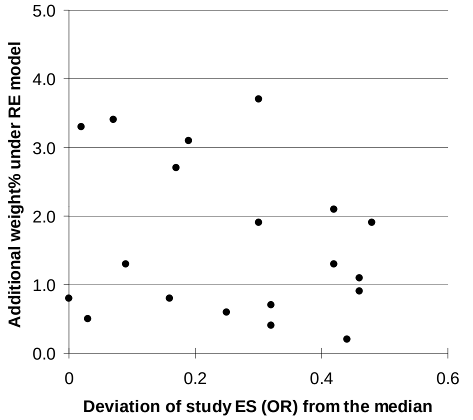
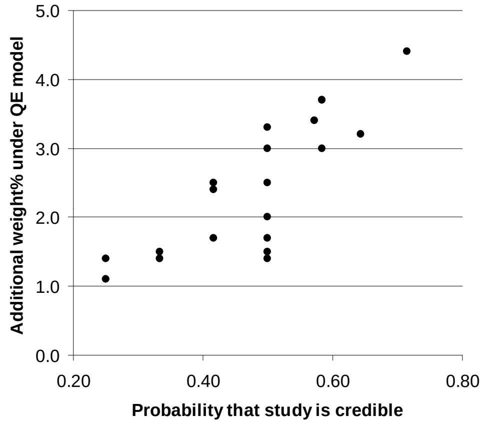


Figure 8



Comparison of Random Effect with Fixed Effect Results

Even the random effects method, when used in a meta-analysis of badly designed studies, will still result in bad statistics even though there is adjustment for heterogeneity.

Furthermore, such adjustments, based on an artificially inflated variance, lead to a widened confidence interval, supposedly to reflect heterogeneity, but do not have much clinical relevance.

Comparison of Random Effect with Fixed Effect Results

One of the reasons for inter-study differences is the quality of individual studies

If the quality of the primary material is inadequate, this may falsify the conclusions of the review, regardless of the random-effects model.

Clinical studies may become corrupted accidentally or deliberately,

- in the randomization process,
- in the masking to the allocated treatment,
- in the random generation of number sequences,
- in the analysis.

The need for analysis of the quality of these studies has become obvious and becomes more important than just inserting a random term based on heterogeneity.

Conclusion

Weighting studies: Random Effect

Choice of model entails a judgement of how much weight should be accorded to a trial by virtue of its numbers and effect size alone.

- **Random effects** meta-analysis will award relatively more weight to smaller studies than such studies would receive in a fixed effect meta-analysis.
- if results of smaller studies are systematically different from results of larger ones, (publication bias or low study quality bias) then a random effects meta-analysis will exacerbate the effects of the bias.
- Alternatively if the large study is biased, then the random effects meta-analysis will exacerbate the effects of the bias.
- If neither is biased, the random and fixed models will give the same results