10. Systematic Review and Meta-Analysis

10.1 Systematic Review

If a clinical trial has been properly conducted, it should provide information regarding the efficacy or effectiveness of a new therapy. Once a trial has been published, it might be thought that it would be unethical to undertake another trial making the same comparison. In practice, the decision is rarely so simple. Trial results generally need to be replicated before new treatment can be widely adopted. Clinical trials of new treatments are often repeated within different global regions (e.g. Europe, Americas, or Asia) to assess their generalisability. Modifications to the trial design may be made to remove perceived biases in the design of earlier studies or test the effect of treatment on other outcome measures.

Where several trials have been carried out to compare the same treatments, the traditional method for assessing the evidence involved selecting from the readily available trial reports, appraising each, before drawing conclusions in a narrative discussion. This type of review can be highly subjective and open to selection bias. An alternative is a *systematic review*, in which studies are identified systematically in attempt to find all before combining the results by an overall statistical summary. Systematic review is now an important component of the evaluation of new treatments and diagnostic test procedures, and is also used to combine evidence from epidemiological studies. *Meta-analysis* is the statistical methodology for combining data from several studies of the same question to produce an overall summary.

A systematic review followed by a meta-analysis can bring together the results from several inconclusive or conflicting studies to give a single conclusive result. It also gives greater power and precision to answer more refined research questions. For example, individual trials are usually designed to answer the question "does a treatment work on average". They rarely have sufficient power to investigate differences in the treatment effect for specific types of patient, but this may be possible by combining data from several studies in a meta-analysis. A systematic review may also enable one to investigate rarer outcomes, such as serious adverse events, that may not be possible in a single trial. For example by combining trials of treatments for depression it has been possible to show that some drug treatments increased the risk of suicidal behaviour, a result that could not be demonstrated in individual trials due to lack of power for this outcome measure.

Steps in a Systematic Review

A *systematic review* is similar to a clinical trial. It involves several steps.

1. Define precise objectives for the review.
2. Set inclusion and exclusion criteria for trials.
3. Search for trials satisfying the inclusion criteria.
4. Assess methodological quality of studies identified, possibly discarding methodologically poor studies.
5. Extract statistical summary data or obtain raw data for each study.
6. Estimate the overall treatment effect by a meta-analysis.

10.2 Meta-analysis

Meta-analysis of generally address the following questions

1. Are the effects in the studies homogeneous? This is needed to justify estimating an overall treatment effect.
2. What is the overall treatment effect?
3. Do study size, study characteristics or methodological quality correlate with the magnitude of the treatment effect?

The best way to carry out a meta-analysis is to combine the raw data from individual studies into a single large dataset, and then carry out an analysis of all the data to estimate the overall effect. This is method called *individual patient data* meta-analysis. Whilst this is similar to analysing a single large study, analysis should take account of data coming from several studies.

Individual patient data meta-analysis is often not possible, because the original raw data are no longer available for all studies, particularly where some may be many years old. For this reason, most meta-analyses use summary statistics extracted from published reports. This method is called *summary measures* meta-analysis and is a special set of methods.

Fixed or Random Effect Meta-analysis

Suppose there are *k* studies and the treatment effect estimate for the i^{th} study is $\hat{\theta}_i$. Suppose the overall treatment effect is θ. There are two main approaches to estimation of θ.

In the first, we assume that $\hat{\theta}_i$ each trial is estimating the a common effect of treatment θ. Any departure of $\hat{\theta}_i$ from θ is assumed to be simply due to sampling variation. This is called Fixed *Effects* estimation.

The second approach is called *Random Effects* estimation. This assumes that the studies are sampled from a larger population of studies. The treatment effect θ_i is then a random variable with mean equal to the overall effect θ and variance *v*.

If $\hat{\theta}$ is the overall estimate, $Var\left[\hat{\theta}\right]$ will be larger if estimated by random effect estimation than fixed effect estimation due to the additional variance term *v*.

In this module we will just describe methods of analysis for fixed effects estimation.

10.3 Summary Measures Estimation of the Overall Effect

Suppose there are *k* trials comparing two treatments. Let $\hat{\theta}_i$ be the estimate of the treatment effect for the i^{th} trial and let $v_i = Var\left[\hat{\theta}_i\right]$.

For a continuous outcome measure *y*, define $\bar{y}_{ij}$, s_{ij}^2 and n_{ij} $(i = 1,...,k; j = 1,2)$ to be sample mean and variance, and the sample size respectively of the j^{th} treatment in the i^{th} trial. The treatment effect of the i^{th} trial can be the mean difference, $\hat{\theta}_i = \bar{y}_{i2} - \bar{y}_{i1}$ with

$$\hat{v}_i = \hat{V}ar\left[\hat{\theta}_i\right] = \frac{s_{i1}^2}{n_{i1}} + \frac{s_{i2}^2}{n_{i2}}.$$

Formula differ from that for a t-test as we no longer assume that variance are equal used to justify the t-test

If *Y* is binary, one could use the rate difference (*RD*), as the summary statistic for each trial. If the observed number of events is r_{ij}, the observed proportions $p_{ij} = r_{ij}/n_{ij}$ $(i = 1,...,k; j = 1,2)$. One can define

$$\hat{\theta}_i = P_{i2} - P_{i1} \text{ with } \hat{V}ar\left[\hat{\theta}_i\right] = \frac{p_{i1}(1-p_{i1})}{n_{i1}} + \frac{p_{i2}(1-p_{i2})}{n_{i2}} = \hat{v}_i .$$

Note this is the non-null variance from page 37

Alternatively, one might want to estimate an overall odds ratio (*OR*) or rate ratio (*RR*). These are generally estimated by taking $\hat{\theta}_i$ equal to the $\log_e\left[\widehat{OR}\right]$ or $\log_e\left[\widehat{RR}\right]$.

For $\log_e\left[\widehat{OR}\right]$, $\hat{v}_i = \frac{1}{r_{i1}} + \frac{1}{n_{i1} - r_{i1}} + \frac{1}{r_{i2}} + \frac{1}{n_{i2} - r_{i2}}$ demonstrated in section 4. (page 45)

For $\log_e\left[\widehat{RR}\right]$, $\hat{v}_i = \frac{1}{r_{i1}} - \frac{1}{n_{i1}} + \frac{1}{r_{i2}} - \frac{1}{n_{i2}}$ See course work

The overall estimate for odd ratio and rate ratio are obtained by taking the exponent of the overall $\log_e[OR]$ and $\log_e[RR]$ estimates.

Summary Measures Estimate of the Overall Effect

Whichever type of summary measure is used (mean difference, *RD*, log [*OR*] or log[*RR*]), an overall estimate of θ can be estimated by a weighted mean defined as

$$\hat{\theta} = \sum_{i=1}^{k} w_i \hat{\theta}_i \Big/ \sum_{i=1}^{k} w_i .$$

Consider now the variance of the estimate for $\hat{\theta}$.

$$Var\left[\hat{\theta}\right] = \frac{1}{\left(\sum_{i=1}^{k} w_i\right)^2} Var\left[\sum_{i=1}^{k} w_i.\hat{\theta}_i\right]$$

Since the studies are independent, $Cov\left[\hat{\theta}_i, \hat{\theta}_j\right] = 0$ (trials)

Therefore, $Var\left[\hat{\theta}\right] = \dfrac{\sum_{i=1}^{k} w_i^2 Var\left[\hat{\theta}_i\right]}{\left(\sum_{i=1}^{k} w_i\right)^2}$.

Choice of Weights and the Minimum Variance Estimate

Different weights will give different estimates of θ and $Var\left[\hat{\theta}\right]$. We could weight studies equally by setting $w_i = 1, i = 1,...,k$, but this is rarely done as the size of studies generally varies greatly. It can be

shown that taking $w_i \propto \dfrac{1}{Var\left[\hat{\theta}_i\right]}$ for each *i* gives an estimator with minimum variance, that is with greater precision. For this reason, *inverse variance weights* are often used in meta-analysis.

The weighted variance $Var\left[\hat{\theta}\right] = \dfrac{\sum_{i=1}^{k} w_i^2 .Var\left[\hat{\theta}_i\right]}{\left(\sum_{i=1}^{k} w_i\right)^2}$ will have a minimum when if $w_i \propto 1/Var[\theta_i]$.

The proof uses the Lagrange Multiplier method for obtaining maxima or minima subject to a constraint.

Let $Var\left[\hat{\theta}\right] = F\left(w_1, w_2, ..., w_k\right) = \dfrac{\sum_{i=1}^{k} w_i^2 .Var\left[\hat{\theta}_i\right]}{\left(\sum_{i=1}^{k} w_i\right)^2}$

Without loss of generality one can apply the constraint $\sum_{i=1}^{k} w_i = 1$.

Define $G\left(w_1, w_2, ..., w_k\right) = \sum_{i=1}^{k} w_i - 1$.

Applying the Lagrange Multiplier Method one defines

$$H\left(w_1, w_2, ..., w_k, \lambda\right) = F\left(w_1, w_2, ..., w_k\right) + \lambda G\left(w_1, w_2, ..., w_k\right)$$

The minimum of *F* subject to the constraint *G* is found by equating the partial derivatives of $H\left(w_1, w_2, ..., w_k, \lambda\right)$ with respect to each w_i to zero. Considering the j^{th} study

$$\frac{\partial H}{\partial w_j}(w_1, w_2, ..., w_k, \lambda) = \frac{\partial}{\partial w_j}\sum_{i=1}^{k} w_i^2 . Var\left[\hat{\theta}_i\right] + \lambda \frac{\partial}{\partial w_j}\left(\sum_{i=1}^{k} w_i - 1\right)$$

Hence

$$\frac{\partial H}{\partial w_j}(w_1, w_2, ..., w_k) = 2w_j Var\left[\hat{\theta}_j\right] + \lambda = 0 \text{ giving } w_i = -\lambda \Big/ \left(2Var\left[\hat{\theta}_i\right]\right)$$

The second derivatives of *H* are positive so this must be a minimum.

Hence $w_i \propto 1 \Big/ Var\left[\hat{\theta}_i\right]$ gives the estimate with minimum variance●

If $\hat{\theta}_{MV}$ is the minimum variance estimate then

$$Var\left[\hat{\theta}_{MV}\right] = \frac{1}{\sum_{i=1}^{k} \frac{1}{Var\left[\hat{\theta}_i\right]}} = \frac{1}{\sum_{i=1}^{k} w_i}$$

Substitutes $\frac{1}{Var\left[\hat{\theta}_i\right]}$ for w_i, into

$$Var\left[\hat{\theta}\right] = \frac{\sum_{i=1}^{k} w_i^2 Var\left[\hat{\theta}_i\right]}{\left(\sum_{i=1}^{k} w_i\right)^2}$$

gives

$$Var\left[\hat{\theta}_{MV}\right] = \frac{\sum_{i=1}^{k} w_i^2 . Var\left[\hat{\theta}_i\right]}{\left(\sum_{i=1}^{k} w_i\right)^2} = \frac{\sum_{i=1}^{k} \frac{1}{Var\left[\hat{\theta}_i\right]}}{\left(\sum_{i=1}^{k} \frac{1}{Var\left[\hat{\theta}_i\right]}\right)^2} = \frac{1}{\sum_{i=1}^{k} \frac{1}{Var\left[\hat{\theta}_i\right]}} = \frac{1}{\sum_{i=1}^{k} w_i}$$

as required ●

Summary Measures Inference

Confidence Intervals

Even if the source data are not normally distributed, it is plausible that the individual study level estimates $\hat{\theta}_i$ are by the central limit theorem. Since $\hat{\theta}_{MV}$ is a linear function of $\hat{\theta}_i$s that are plausibly normally distributed, we can assume that $\hat{\theta}_{MV}$ is also normal. A $(1-\alpha)$ level confidence interval of $\hat{\theta}_{MV}$ can therefore be given by

$$\hat{\theta}_{MV} \pm z_{\alpha/2} \frac{1}{\sqrt{\sum_{i=1}^{k} w_i}}$$

Hypothesize Tests

To test the null hypothesis $H_0 : \theta = 0$, the following test statistic can be used

$$T = \frac{\hat{\theta}_{MV}}{SE\left[\hat{\theta}_{MV}\right]} = \hat{\theta}_{MV} \sqrt{\sum_{i=1}^{k} w_i},$$

which can be assumed to have a standardised normal distribution under the null hypothesis.

Ex 10.1 Systematic review of the effect of maternal steroid therapy on neonatal mortality.

The table over-page summarizes the results for 12 trial identified by a systematic review of trials testing maternal steroid therapy. The outcome measure is the number of neonatal deaths, which is death within the first 28 days of life. Note that in one study, there are no deaths in both arms and so this study cannot contribute to the meta-analysis and has to be excluded from the analysis.

(i) Estimate the minimum variance estimate and its 95% confidence interval.

(ii) Test the null hypothesis of no overall treatment effect.

Some of the computation is carried out on the table above summarizing the raw data.

(iii) Display the data graphically.

The standard method of graphical display of a meta-analysis is a forest plot illustrated below.

Fixed Effects Meta-Analysis of the Effect of Maternal Steroid Therapy on Neonatal Mortality (Crowley et al,1990)

Trial No.	Steroid Therapy			Control			$\hat{\theta}_i = P_S - P_C$	$v_i = Var[P_S - P_C]$	$w_i = 1/v_i$	$w_i\hat{\theta}_i$
	$Died\ (r_S)$	P_S	n_S	$Died\ (r_C)$	P_C	n_C				
Liggins	36	0.068	532	60	0.112	538	-0.044	0.000303	3302.88	-145.32
Block	1	0.014	69	5	0.082	61	-0.067	0.001441	694.17	-46.51
Scliutte	3	0.047	64	12	0.207	58	-0.160	0.003527	283.5	-45.37
Taeush	5	0.089	56	7	0.099	71	-0.009	0.002704	369.9	-3.44
Doran	2	0.025	81	10	0.159	63	-0.134	0.002417	413.8	-55.46
Teranin	0	0.000	38	0	0.000	42	0.000	-	-	-
Gamsu	14	0.107	131	20	0.146	137	-0.039	0.001639	610.3	-23.87
Collab. Grp.	36	0.097	371	37	0.099	372	-0.002	0.000477	2096.7	-5.09
Morales	7	0.058	121	13	0.105	124	-0.047	0.001207	828.3	-38.92
Papagecrgio	1	0.014	71	5	0.067	75	-0.053	0.001025	975.4	-51.29
Morrison	2	0.030	67	7	0.119	59	-0.089	0.002205	453.6	-40.28
Schmidt	5	0.147	34	5	0.161	31	-0.014	0.008053	124.2	-1.77
								Σ	10152.6	-457.2

(i) Estimate the minimum variance fixed effect estimate and its 95% confidence interval.

The fixed effect estimate

$$\hat{\theta}_{MV} = \frac{\sum_{i=1}^{k} w_i \hat{\theta}_i}{\sum_{i=1}^{k} w_i} = \frac{-457.2}{10152.6} = -0.0450$$

$$\hat{SE}[\theta_{MV}] = \sqrt{\frac{1}{\sum_{i=1}^{k} w_i}} = \sqrt{\frac{1}{10152.6}} = 9.924 \times 10^{-3}$$

95% C.I. is $\hat{\theta}_{MV} \pm z_{\alpha/2} \hat{SE}[\theta_{MV}]$, assuming normality which plausible due to the large sample size

95% C.I. is $-0.045 \pm 1.96 \times 9.924 \times 10^{-3}$ giving

-6.4% to -2.4%

(ii) Test the null hypothesis of no overall treatment effect

$$T = \frac{\hat{\theta}_{MV}}{SE[\hat{\theta}_{MV}]} = \frac{0.045}{9.924 \times 10^{-3}} = -4.54$$

T is assumed to be normally distributed

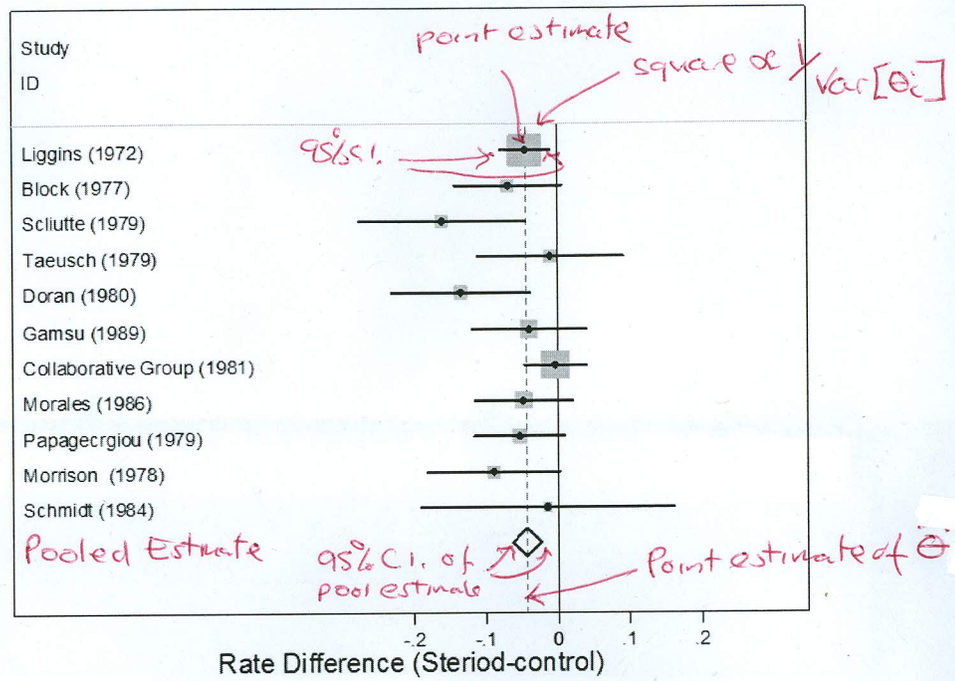
For a two-sided test (see page 4-1)

p-value $= 2(1 - \Phi(|T|)) = 2(1 - \Phi(4.54)) < 0.0004$ from tables

Hence we can reject $H_0: \theta = 0$ as the p-value is certainly less than 5% or 1%

(iii) **Graphical Display of Meta-Analyses**

Forest Plot of Data from Crowley et al.



The treatment effect in each study is represented by a square with bars represent the 95% confidence interval of the treatment effect. The combined treatment effect and its confidence interval are shown at the bottom of the figure as a diamond. The area of the block representing the point estimate for each study has been made inversely proportional to the variance. Since larger studies will have smaller variance, larger studies will be represented by a large block. This is added otherwise the eye would tend to be attracted towards the studies that have wider confidence intervals which are smaller.

10.4 Investigation of Biases

It is well known that studies that fail to find a statistically significant treatment effect are less likely to be published than those that do. This mean that a meta-analysis based only on published studies may be biased. The term used for this phenomenon is *Publication Bias*.

Possible Causes of Publication Bias

- Selective publications: Studies in which an intervention is found to be ineffective are sometimes never published. Sponsors of research, such as pharmaceutical companies or the innovator of the treatment, have been known to discourage or prevent publication of unfavourable results. If the results are negative, a clinical researcher may be less motivated to get a trial published as they are conscious that they may be considered less interesting to journal editors and so much more difficult to get accepted published.

- Identification: Studies in which results are statistically significant are likely to be published in more prestigious, and hence easily accessible, journals. As an illustration of this, it has been shown that trials carried out in non-English speaking countries are more likely to be published in English where the study result is statistically significant. Hence, a meta-analysis restricted to English language journals may overestimate the treatment effect as studies in other languages will tend to have a smaller effect.

- Selective reporting: Where studies have multiple outcomes measured, statistically significant results may be emphasized in reports whereas non-significant results may be given less prominence or even left out. Glaring examples of this are trial reports that fail to give the primary outcome measure previously specified in the trial protocol, but publish other measures that have been found to be statistically significant.

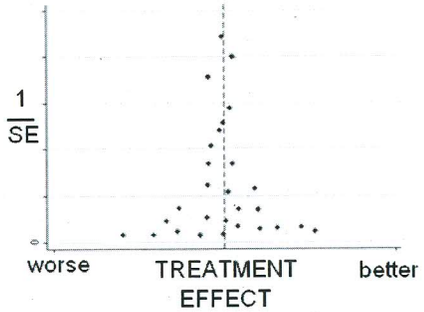
The Funnel Plots

One way to investigate publication bias is a funnel plot. This plots

$$\text{Precison}[\hat{\theta}] = \frac{1}{\text{SE}[\hat{\theta}]}$$

against the treatment effect for each trial $\hat{\theta}$. Assuming all studies in the meta-analysis are a random sample of all possible studies of the same treatment, the distribution of points should resemble an inverted funnel shape widening as the precision decreases. This is because studies with larger standard errors (i.e. less precision) will have wider confidence intervals and so estimates of the treatment effect will be more widely dispersed.

Figure 10.2 A Funnel Plots

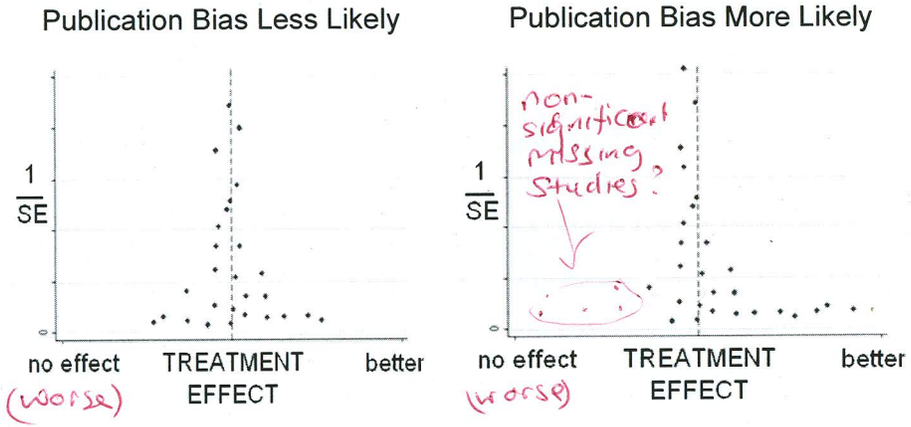


Funnel plots can also constructed by plotting total sample size against the treatment effect and give a similar shaped figure, as precision is related to the square root of the sample size.

Funnel Plots Asymmetry

Studies with greater precision have larger sample size and so tend to get published irrespective of statistically significance. In contrast, studies with less precision are less likely to be published, if they are not statistically significant. Hence, smaller studies showing a smaller treatment are more likely to be missed by a systematic review and so left out of a meta-analysis. This is illustrated in figure 10.3.

Figure 10.3 Illustration of publication bias

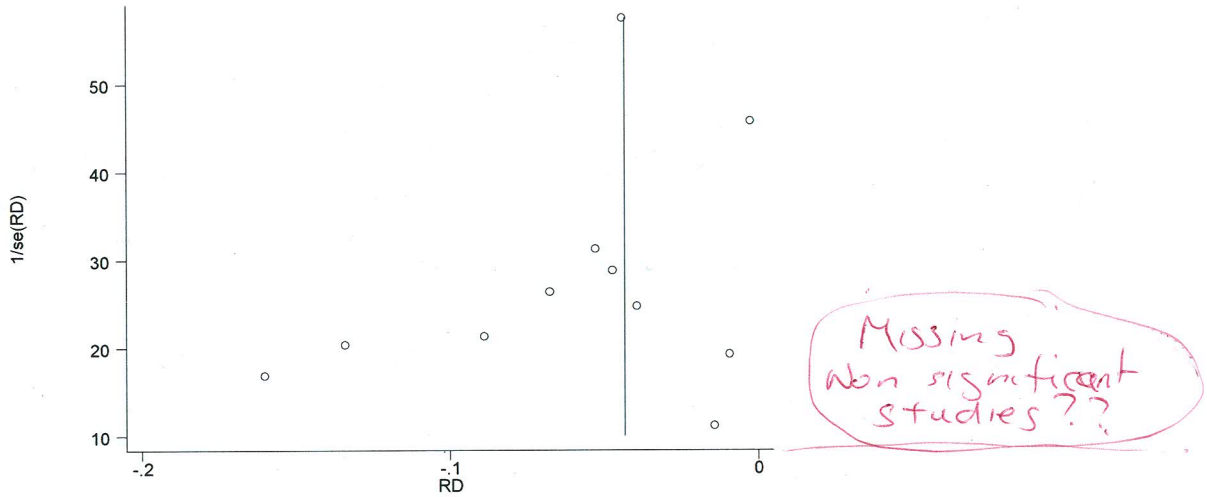


As well as publication bias, lack of symmetry in funnel plots may indicate:

- True Heterogeneity – the treatment in smaller studies may be more intensive than in larger studies or patients in smaller studies may differ systematically from those in larger studies.
- Outcome may be measured in different ways depending on trial size.
- Smaller trials may be more poorly conducted than larger studies and so more likely to be biased.

The possibility of publication bias means that it is particularly important that meta-analyses are based on all relevant studies and not just those that are conveniently available. Researchers carrying out systematic reviews are encouraged to identify trials that have not been published or are reported in more obscure journals. To aid this, an international directory of clinical trials (ISRCTN) has been established with which all new randomised trials should registers.

Figure 10.4 Funnel Plot of Crowley et al.



Missing non significant studies??

There is some evidence in the funnel plot above that smaller studies showed a larger effect. This could be due to publication bias.

If there is concern that there may be publication bias, one option would be to carry out a sensitivity analysis excluding smaller studies.

END OF MODULE