GEORGE & FAY YEE
Centre for Healthcare Innovation

Dr. Rasheda Rabbani • March 24, 2022
rasheda.rabbani@umanitoba.ca

CONDUCTING SYSTEMATIC REVIEWS STATISTICS (i)

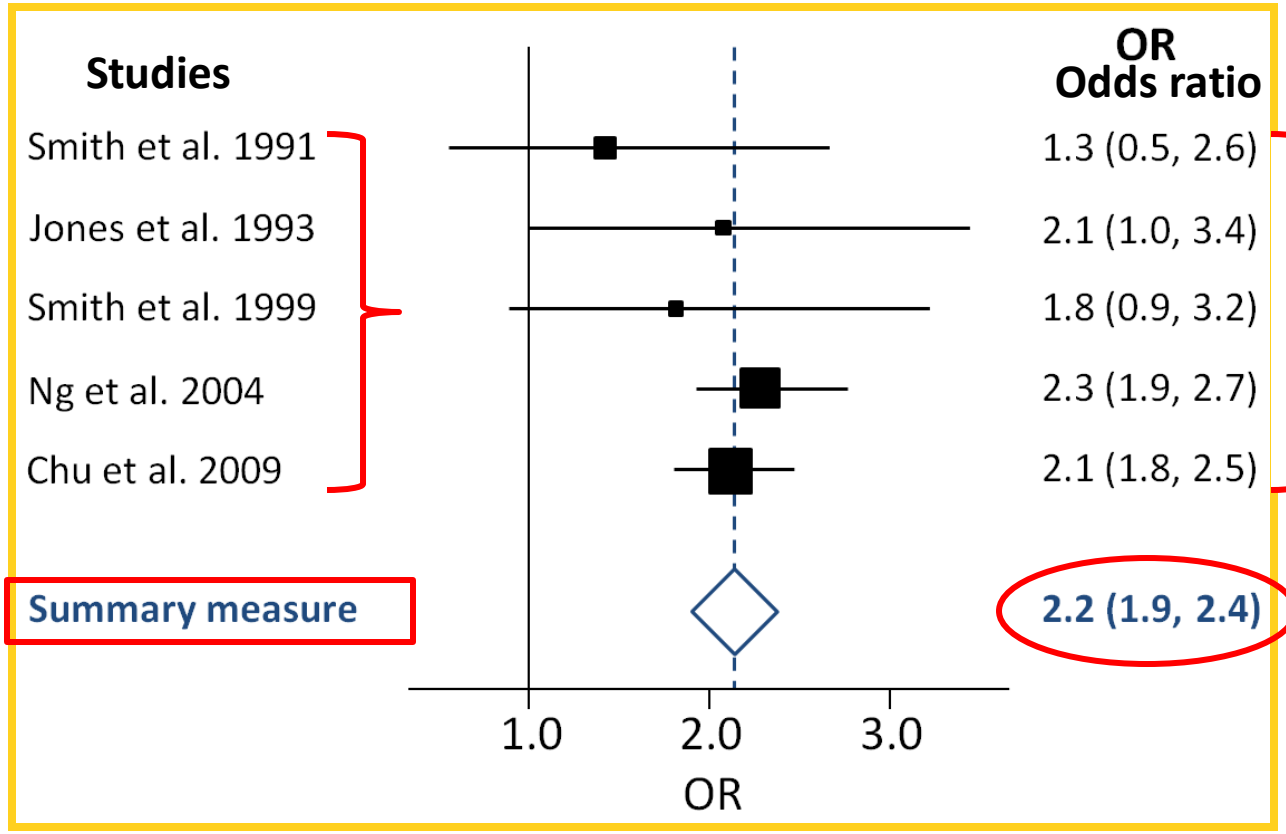
CONCEPTS OF META-ANALYSIS, AND HOW IT WORKS.

Today's objectives

- Discuss analytic issues pertaining to studies with dichotomous and continuous outcomes.

- Introduce the concepts of meta-analysis (weighting and pooling techniques).

What is meta-analysis



Meta-analysis is the statistical combination of results for same specific outcome from two or more separate studies.

Why perform a meta-analysis in a review?

- To increase power.
- To improve precision.
- To answer questions not posed by the individual studies.
- To settle controversies arising from apparently conflicting studies or to generate new hypotheses.

Study protocols

A clinical study protocol is a document that provides

- Details of the study plan and organization and is written prior to the start of subject, (study) recruitment and data collection.

- Protocols include information on study rationale, objectives, methodology (design and statistical approaches), types of participants (i.e., inclusion and exclusion criteria), treatments, outcome and the duration of the study.

Is it always right to perform a meta-analysis?

- NO; not all systematic reviews are (nor should be) including meta-analyses.
- A summary measure is only appropriate when no large discrepancies between studies in patient traits, study traits, etc.
  - i.e. "Clinically and Methodologically homogenous".
- Otherwise, the pooled estimate may be a biased representation of the true estimate of interest.
- Various approaches to assessing heterogeneity; need to consider these, and be clinically sensible.

we'll discuss details about heterogeneity later on today and in second day.

Basic principles in meta-analysis

Choice of

- Types of outcome data
- An effect measure
- Data extraction and conversions
- A model
- How to handle heterogeneity

Deeks JJ, Higgins JPT, Altman DG (editors). Chapter 10: Analysing data and undertaking meta-analyses. In: Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch VA (editors). *Cochrane Handbook for Systematic Reviews of Interventions* version 6.3 (updated February 2022). Cochrane, 2022. Available from **www.training.cochrane.org/handbook.**

Basic principles in meta-analysis

Choice of

- **Types of outcome data**
- An effect measure
- Data extraction and conversions
- A model
- How to handle heterogeneity

Types of outcome data?

Dichotomous (or binary) outcome: when the outcome for every participant is one of two possibilities, for example, dead/alive; smoking/non-smoking; cancer/not cancer.

Continuous outcome: where each individual's outcome is a measurement of a numerical quantity. For example, weight, area and volume, blood pressure.

Ordinal outcome: the classification of disease severity into 'mild', 'moderate' or 'severe', is of ordinal type. As the number of categories increases, ordinal outcomes acquire properties similar to continuous outcomes.

Types of outcome data?

Counts and rates: calculated from counting the number of events that each individual experiences. for example, fracture, an adverse reaction or a hospitalization.
Dichotomize – how many patients had at least one event.

Time-to-event (typically survival) outcome: the time until an event occurs, but where not all individuals in the study experience the event (censored data).

looking not only at *how many* patients died, but also at *how long* after treatment they died, gives a **more sensitive** assessment.

Time-to-event data can sometimes be analysed as dichotomous data. This requires the status of all patients in a study to be known at a fixed time-point.

Basic principles in meta-analysis

Choice of

- Types of outcome data
- **An effect measure**
- Data extraction and conversions
- A model
- How to handle heterogeneity

Choice of Effects measures: dichotomous outcome

Risk Ratio (RR) (also called the relative risk):

In intervention studies, it is the ratio of the risk in the intervention group to the risk in the control group. A risk ratio of one indicates no difference between comparison groups. Large RR's are impossible with common events.

Odds Ratio (OR):

OR describes the odds of an event in the treatment group relative to the odds of an event in the control group. When events are uncommon, chances are low odds ratios are very similar to risk ratios. When events are common, chances are high. OR>RR. Case control study calculate OR.

Choice of Effects measures: dichotomous outcome

Risk Difference (RD) *(also called the absolute risk reduction):* RD describes the risk of an event in the treatment group minus risk of an event in the control group.

Relative measures don't tell you the actual number of participants who benefited. E.g., RR 2.0….same for 80% vs 40% as for 10% vs 5%...but these are very different event rates!

Basic principles in meta-analysis

Choice of

- Types of outcome data
- An effect measure
- **Data extraction and conversions**
- A model
- How to handle heterogeneity

Choice of Data Extraction and converting to the desired format. (dichotomous outcome)

Dichotomous outcome: data extraction and conversion

Calculation of risk ratio (RR), odds ratio (OR) and risk difference (RD) from a 2 × 2 table:

The results of clinical trial can be displayed as a 2 × 2 table

Treatment Group	Event	No event	Total
Intervention	*a*	*b*	*a+b*
Control	*c*	*d*	*c+d*

Where a, b, c, d are the numbers of participants in each treatment group by Event and No event.

$$RR=\frac{risk\ of\ event\ in\ experimental\ group}{risk\ of\ event\ in\ control\ group}=\frac{a/a+b}{c/c+d}$$

$$OR=\frac{odds\ of\ event\ in\ experimental\ group}{odds\ of\ event\ in\ control\ group}=\frac{a/b}{c/d}=\frac{ad}{bc}$$

$$RD=risk\ of\ event\ in\ experimental\ group-risk\ of\ event\ in\ control\ group=\frac{a}{a+b}-\frac{c}{c+d}$$

Transformations for OR and RR

In a trial of two treatments for ulcer healing two groups were compared:

- elastic bandage: 31 healed out of 49 patients
- inelastic bandage: 26 healed out of 52 patients.

The risk ratio can be presented in two ways:

- RR = (31/49)/(26/52) = 1.27 (elastic over inelastic)
- RR = (26/52)/(31/49) = 0.79 (inelastic over elastic)

We want a scale where 1.27 and 0.79 are equivalent. They should be equally far from 1.0, the null hypothesis value.

We use the logarithm of the risk ratio:

- log10(1.273) = 0.102, log10(0.790) = –0.102
- log10(1) = 0 (null hypothesis value)

Effects measures: dichotomous outcome

Good vs. Bad Outcomes

	Good outcome (e.g. survival)	Bad outcome (e.g. infection)
RR<1	Reduced risk (not beneficial)	Reduced risk (beneficial)
RR>1	Increased risk (beneficial)	Increased risk (harmful)
OR=1, RR=1	No difference	No difference
OR<1	Reduced odds (not beneficial)	Reduced odds (beneficial)
OR>1	Increased odds (beneficial)	Increased odds (harmful)

Effects measures: dichotomous outcome

Cautions:

- Neither the risk ratio nor the odds ratio can be calculated for a study if there are no events in one group or arm. This is because, as can be seen from the formulae we would be trying to divide by zero.

$$RR=\frac{risk\ of\ event\ in\ experimental\ group}{risk\ of\ event\ in\ control\ group}=\frac{a/a+b}{c/c+d}$$

$$OR=\frac{odds\ of\ event\ in\ experimental\ group}{odds\ of\ event\ in\ control\ group}=\frac{a/b}{c/d}=\frac{ad}{bc}$$

$$RD=risk\ of\ event\ in\ experimental\ group-risk\ of\ event\ in\ control\ group=\frac{a}{a+b}-\frac{c}{c+d}$$

- In these situations, it is usual to add 0.5 to each cell of the 2 × 2 table (RevMan automatically makes this correction when necessary).

Effects measures: dichotomous outcome

Cautions:

- In the case where no events (or all events) are observed in both groups or arms the study provides no information about relative probability of the event and is automatically omitted from the meta-analysis.
- Zeros arise particularly when the event of interest is rare – such events are often unintended adverse outcomes.
- Counts of rare events should be treated differently, use rates (per persons-time-units of follow-up), or dichotomize – how many patients had at least one event.

Effects measures: dichotomous outcome

Number Needed to Treat for an additional Beneficial outcome (NNTB):

NNTB describes the number needed to be treated in order to prevent one Failure. The higher the NNTB, the less effective is the treatment. e.g., if you need to give a stroke prevention drug to 20 people before one stroke is prevented, then the number needed to treat to benefit for that stroke prevention drug is 20.

Number Needed to Treat for an additional Harmful outcome (NNTH):

associated with a harmful effect. It is an estimate of how many people need to receive a treatment before one more person would experience a harmful outcome.

Effects measures: dichotomous outcome

Relative risk reduction (RRR):

Is a convenient way of re-expressing a risk ratio as a percentage reduction.

For example: a risk ratio of 3 for a treatment implies that events with treatment are three times more likely than events without treatment. Alternatively we can say that treatment increases the risk of events by 100 × (RR – 1)% = 200%.
Similarly a risk ratio of 0.25 is interpreted as the probability of an event with treatment being one-quarter of that without treatment. This may be expressed alternatively by saying that treatment decreases the risk of events by 100 × (1 –RR)% = 75%.

The interpretation of the clinical importance of a given risk ratio cannot be made without knowledge of the typical risk of events without treatment: a risk ratio of 0.75 could correspond to a clinically important reduction in events from 80% to 60%, or a small, less clinically important reduction from 4% to 3%.

Choice of Effect Measures (continuous outcome)

Effects measures: continuous outcome

The mean difference (MD):

$$MD = Mean_{treatment} - Mean_{Control}$$

It can be used as a *summary* statistic in meta-analysis when outcome measurements in all studies are made on the same scale.

The standardized mean difference (SMD):

$$SMD = \frac{difference\ in\ mean\ outcome\ between\ groups}{standard\ deviation\ of\ outcome\ among\ particips ants}$$

The **SMD** is used as a summary statistic in meta-analysis when the studies all assess the same outcome but measure it in a variety of ways (for example, all studies measure depression but they use different psychometric scales).

Effects measures: continuous outcome

Caveats on Using Standard Mean Difference: (SMD)

- Sample variance heterogeneity.
- Adjusted Covariates effect measures.
- Directionality.
- Missing standard deviation.
- Multiplicity of data.
- Cannot pool reported mean change from baseline and final values.

In this situation what can we do?

- Ratio of Means.
- Convert in same scale and use mean difference (MD).
- Dichotomizing Continuous Outcomes in Meta-Analyses.

Choice of Data Extraction and converting to the desired format. (continuous outcome)

Continuous outcome: data extraction and conversion

1. Post-intervention (final values) versus change from baseline for both intervention and control groups (for SMD).

2. Obtaining standard deviations (SDs) from standard errors (SEs) and confidence intervals (CI) for group means.

3. Obtaining standard deviations from t values and P values for differences in means.

4. Transformations and skewed data.

5. No information on variability.

Continuous outcome: data extraction and conversion

6. Medians and interquartile ranges (IQR=1.35*SD, for Gaussian data).

7. Median Ranges (Hozo at al) (Walter et al, for Gaussian data).

8. Estimating the sample mean and standard deviation from the sample size, median, range and/or interquartile range (Wan et al, Bland et al).

9. Approximate Bayesian Computation (ABC) method (Mazer and Kwon D. et al).

10. Ratio of Mean method.

Weir et al. BMC Medical Research Methodology (2018) 18:25
https://doi.org/10.1186/s12874-018-0483-0

BMC Medical Research Methodology

RESEARCH ARTICLE Open Access

Dealing with missing standard deviation and mean values in meta-analysis of continuous outcomes: a systematic review

Christopher J. Weir[1*], Isabella Butcher[1], Valentina Assi[1], Stephanie C. Lewis[1], Gordon D. Murray[1], Peter Langhorne[2] and Marian C. Brady[3]

Results: For missing standard deviations (SDs), following screening of 503 articles, fifteen methods were identified in addition to those reported in a previous review. These included Bayesian hierarchical modelling at the meta-analysis level; summary statistic level imputation based on observed SD values from other trials in the meta-analysis; a practical approximation based on the range; and algebraic estimation of the SD based on other summary statistics. Following screening of 1124 articles for methods estimating the mean, one approximate Bayesian computation approach and three papers based on alternative summary statistics were identified. Illustrative meta-analyses showed that when replacing a missing SD the approximation using the range minimised loss of precision and generally performed better than omitting trials. When estimating missing means, a formula using the median, lower quartile and upper quartile performed best in preserving the precision of the meta-analysis findings, although in some scenarios, omitting trials gave superior results.

Conclusions: Methods based on summary statistics (minimum, maximum, lower quartile, upper quartile, median) reported in the literature facilitate more comprehensive inclusion of randomised controlled trials with missing mean or variability summary statistics within meta-analyses.

Continuous outcome: data extraction and conversion

Ratio of means for analyzing continuous outcomes in meta-analysis performed as well as mean difference methods

Jan O. Friedrich[a,b,c,d,e,*], Neill K.J. Adhikari[b,f], Joseph Beyene[g,h,1]

Accepted 4 September 2010

Results: Two hundred thirty-two of 5,053 reviews were included. Measures demonstrated similar treatment effects, with ≤6% discordant pairs and no asymmetry. A 0.5 SMD increase corresponded to 22 (95% confidence interval: 19, 24)% increase using RoM. There was less heterogeneity in RoM vs. MD ($n = 143$, $P = 0.007$), SMD vs. RoM ($n = 232$, $P = 0.005$), and SMD vs. MD ($n = 143$, $P = 0.004$). Comparing discordant pairs, fewer meta-analyses showed significant heterogeneity with SMD vs. RoM ($P = 0.04$), consistent with the known bias of SMD.

Conclusion: Empiric data from diverse meta-analyses demonstrate similar treatment effects and no large differences in heterogeneity of RoM compared with difference-based methods.

Continuous outcome: data extraction and conversion

If Studies provided post intervention Mean or change directly for each group,

- Is there a measure of variance (e.g. SD, SE, CI, P)?
- measure of variance is not SD (convert first).

Some basic formula for missing SD

SD Computation for each group means

- **If Mean, SE and N reported for each group mean**

$$SD = SE \times \sqrt{N}$$

- **If Confidence interval (CI) reported for each group mean:**

$$SD = \sqrt{N} \times (\text{upper limit} - \text{lower limit}) / 3.92$$

3.92 is For 95% confidence interval (CI)

Some basic formula for missing SD

SD Computations for group mean differences

From p-value:

1st step: calculate z value from standardized normal table or t-table or use excel, [**=abs(normsinv(p-value/2))**] [=ABS(T.INV.2T(0.001, 30-1))]

2nd step: SE = mean difference / Z. or t-value for small sample size

From 95% confidence interval:

Step 1: $SE = (\text{upper limit} - \text{lower limit})/3.92$

Step 2: **From standard error to standard deviation** $SD = \dfrac{SE}{\sqrt{\dfrac{1}{N_E} + \dfrac{1}{N_C}}}$

E= Intervention; C= Control

Some basic formula for data conversions

Computing risk ratio from an odds ratio (OR)

$$RR = \frac{OR}{1 - ACR \times (1 - OR)}$$

ACR = Assumed Control group risk/
Median control group risk from included studies in the meta-analysis

Re-expressing SMDs by transformation: to odds ratio

$$\ln OR = \frac{\pi}{\sqrt{3}} SMD$$

Effects measures: other outcome types

Count Data (counts of decayed, missing or filled teeth):

Rate ratio (RR): is used for count data and often estimated from a Poisson regression model. The estimate of rate takes into account both the number of new cases, and follow-up time of population. a woman may experience two strokes during a follow-up period of two years. Her rate of strokes is 1.0 per year of follow-up (or, equivalently 0.083 per month of follow-up).

Time to Events data:

the measure is hazard ratio (HR): and most commonly estimated from the Cox proportional hazards model.
Investigators can also calculate HR and its variance if observed (O) and expected events (E) can be extracted.

Effects measures: other outcome types

Calculation of rate ratio (RR) for Count data:

$$\text{rate ratio} = \frac{E_E/T_E}{E_C/T_C} = \frac{E_E T_C}{E_C T_E} \qquad \text{SE of ln rate ratio} = \sqrt{\frac{1}{E_E} + \frac{1}{E_C}}$$

E_E = events occurred during T_E participant-years of follow-up in the experimental intervention group

E_C = events during T_C participant-years in the control intervention group

Effects measures: other outcome types

Data extraction for time to event outcome:

Trials

BioMed Central

Methodology

Open Access

Practical methods for incorporating summary time-to-event data into meta-analysis

Jayne F Tierney*[1], Lesley A Stewart[2], Davina Ghersi[3], Sarah Burdett[1] and Matthew R Sydes[4]

Effects measures: other outcome types

Data needed for Time to event outcome:

Trial Reference	Treatment	Control
Randomization ratio (e.g 1:1)		
Total patient randomized		
Total patients analysed		
Observed events O		
logrank expected events E		
HR with 95% CI		
Logrank variance (V)		
Logrank observed minus Expected events (O-E)		
HR and Standard Error or variance from adjusted or unadjusted Cox test statistics		
2 sided p-value to 2 significant figures & what type of test used		
(logrank, Mantel hanzsel or unadjusted Cox)		
Advantage to treatment or control		
Actuarial or Kaplan Meier curves repoted		
Number at risk		
Follow-up details		

What to do if more than two intervention in a single trial

- Combine groups to create a single pair-wise comparison (recommended).
- Select one pair of interventions and exclude the others.
- Split the 'shared (control)' group into two or more groups with smaller sample size, and include two or more (reasonably independent) comparisons.
- Include two or more correlated comparisons and account for the correlation. ***Undertake a multiple-treatments meta-analysis.***

What to do if more than two intervention in a single trial

Scenario 1: to combine all relevant experimental intervention groups of the study into a single group, and to combine all relevant control intervention groups into a single control group.

As an example, suppose that a meta-analysis of 'acupuncture versus no acupuncture' would consider studies of either 'acupuncture versus sham acupuncture' or studies of 'acupuncture versus no intervention' to be eligible for inclusion. In the meta-analysis by combining the participants in the 'sham acupuncture' group with participants in the 'no intervention' group. This combined control group would be compared with the 'acupuncture' group in the usual way. For dichotomous outcomes, both the sample sizes and the numbers of people with events can be summed across groups.

Scenario 2: A further possibility is to include each pair-wise comparison separately, but with shared intervention groups divided out approximately evenly among the comparisons.

For example, if a trial compares 121 patients receiving acupuncture with 124 patients receiving sham acupuncture and 117 patients receiving no acupuncture, then two comparisons (of, say, 61 'acupuncture' against 124 'sham acupuncture', and of 60 'acupuncture' against 117 'no intervention

What to do if more than two intervention in a single trial

Continuous outcome:

Formulae for combining groups

	Group 1 (e.g. males)	Group 2 (e.g. females)	Combined groups
Sample size	N_1	N_2	$N_1 + N_2$
Mean	M_1	M_2	$\frac{N_1M_1 + N_2M_2}{N_1 + N_2}$
SD	SD_1	SD_2	$\sqrt{\frac{(N_1 - 1)SD_1^2 + (N_2 - 1)SD_2^2 + \frac{N_1N_2}{N_1 + N_2}(M_1^2 + M_2^2 - 2M_1M_2)}{N_1 + N_2 - 1}}$

1	Calculation of risk ratio (RR), odds ratio (OR) and risk difference
2	Risk Ratio (RR): from given Odds Ratio (OR)

Obtaining standard deviations (SD) from standard errors (SE) and confid	
1	Standardized Mean Differences (SMD): When both groups within mean, S
2	SD: when single group within SE and n are reported
3	Read Note
3a	SD: if given 95% Confidence interval around the mean and n
3b	SD: if given 90% Confidence interval around the mean and n
3c	SD: if given 99% Confidence interval around the mean and n
4	Calculation of t-value from given n
4a	Calculation of t-value for 95% CI: from given n
4b	Calculation of t-value for 90% CI: from given n
4c	Calculation of t-value for 99% CI: from given n

Obtaining standard deviations (SD) from standard errors (SE), confidence intervals (CI), t values and P values for Mean Differences (MD)	
Note	If given SE, p-value, confidence interval, z-stat, or t-stat from the differnce of means, then we can make simplifying assumption that treatment SD is equal to control SD.
1	Calculation of t-value from two sample t-test p-value
2	Calculation of (SE): From t-value and MD
3	Calculation of (SD): From SE for mean differences
4	calculation of z value when of p-value given

1	Formulae for combining groups

Missing correlation

Correlation need to be adjust in the calculation of SE/SD:

Crossover Trials:

SE for mean difference will be too large if parallel design pooled SE formula used. Because the positive correlation associated with the same patients in both the treatment and control groups lowers the variance of the mean difference.

Cluster Randomized Trials:

Ignoring this correlation in cluster randomized trials will produce an SE of the mean difference between intervention groups that is too small.
to account for this discrepancy is to compute a design effect (DE) using the average cluster size and ICC the intra-class correlation coefficient. The square root of the design effect can then be multiplied by the standard error of the regular mean difference (computed as if it were parallel) to produce the adjusted SE.

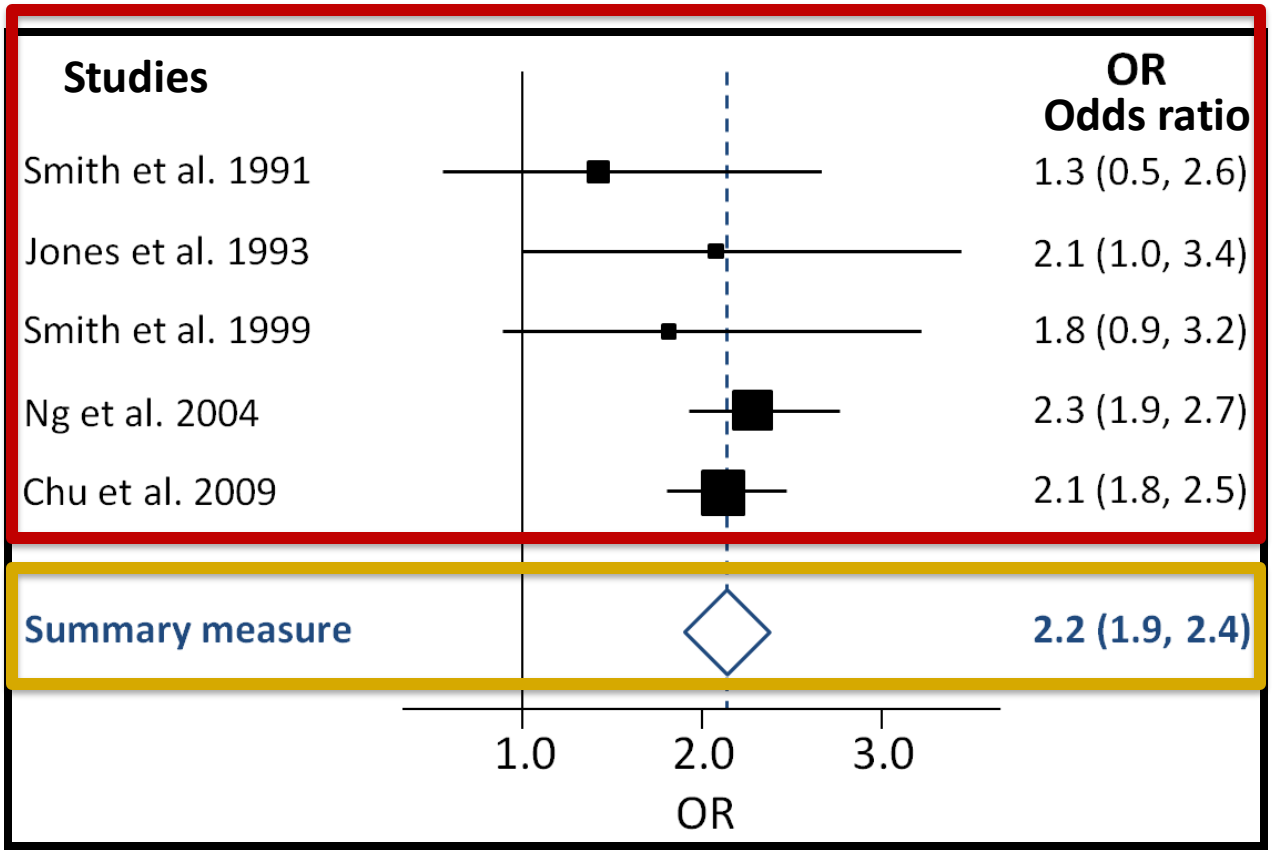
Basic principles in meta-analysis

Choice of

- Types of outcome data
- An effect measure
- Data extraction and conversions
- **A model**
- How to handle heterogeneity

Choice of Statistical Model for Combining Studies

Meta-analysis is typically a two-stage process



Meta-Analysis…the 'average' effect. How to get?

Thoughts about taking the OR/RR/RD/MD from all studies and just getting the crude average?

- All studies just the replications of 'same experiment' so why not?
- But are all studies likely to be in the same ballpark of giving a truthful answer to the research question?
- E.g. a study of 20 patients vs 1000 patients to assess the benefits of drug A versus B…would you have a preference about which to draw conclusions from? Why?

Meta-Analysis…the 'average' effect. how to get?

So if not a crude average, what makes sense?

- A weighted average, which "gives studies some credit" for a few things…Larger sample size? Lower variance?

- Makes intuitive sense a study of 5,000 patients gets more weight than a study based on 5. These features lead to better precision and may thus also give answers nearer the 'truth', so more weight seems rational.

Meta-Analysis…the 'average' effect. how to get?

So what steps go into pooling the RCT data?

- **STEP 1: calculate an OR/RR/RD/MD and 95% CI for individual trials**

  -These are 'basic' stats we already discussed.

- **STEP 2: calculate a pooled OR/RR/RD/MD and 95% CI for the included of trials**

  - Clinically are studies 'combinable' in our view?
  - If yes, calculate weights for each study (fixed, random effects…).
  - Look at measure of statistical heterogeneity (i.e. I^2 or Q, and how much variation in treatment effects is there across studies?).

Factors to choose between FE and RE???

Recall we said it's intuitive to give more weight to bigger, and more precise studies than smaller ones.

Another issue:

Choosing the right approach to modeling, i.e. 'fixed effects' versus 'random effects'.

What exactly does the choice between FE/RE mean?

- **Fixed effects:** assumes all studies estimate the SAME value; only source of variation in estimates is random error.

- **Random effects:** assumes the included studies are estimating related but different values. Two sources of variation…random error, and also between-study variation.

Factors to choose between FE and RE???

Item 1: Clinically and methodologically sensibility

- Do the studies enroll *similar patients*?

- *Careful review* of evidence tables…
  - If very similar, some will suggest fixed effects model

  - If clear differences, random effects model the sensible choice

Heterogeneous in several ways: age, severity, procedure, methods

Study	Mean age	NYHA Class	% repeat procedure	% combined procedure	Double blinded?
Smith	35.2	All class I	0%	0%	Y
Thompson	65.7	Some class 3, some class 4	50%	65%	N
Johnson	78.4	Near all class 4	80%	75%	N
Russell	46.7	Mostly class 1, some class 2	5%	10%	Y
Bryant	59.4	All class 3	20%	30%	Y

Factors to choose between FE and RE???

Item 2: Three Measures of Statistical Heterogeneity

1. **Cochrane Q : measures of weighted squared deviation.**

2. **I^2 (I-square) : ratio of true heterogeneity to total observed variance.**

3. **τ^2 (tau square): Between studies variance.**

Measuring inconsistency in Meta-Analysis
Julian PT Higgins et al

Guidance on interpreting cochrane Q, I^2

- **Cochrane Q:**
  - a chi-square statistic associated with a p-value.
  - Common practice is to "relax" our significance level to 0.1 (or even 0.2) so that we're more conservative on identifying heterogeneity.

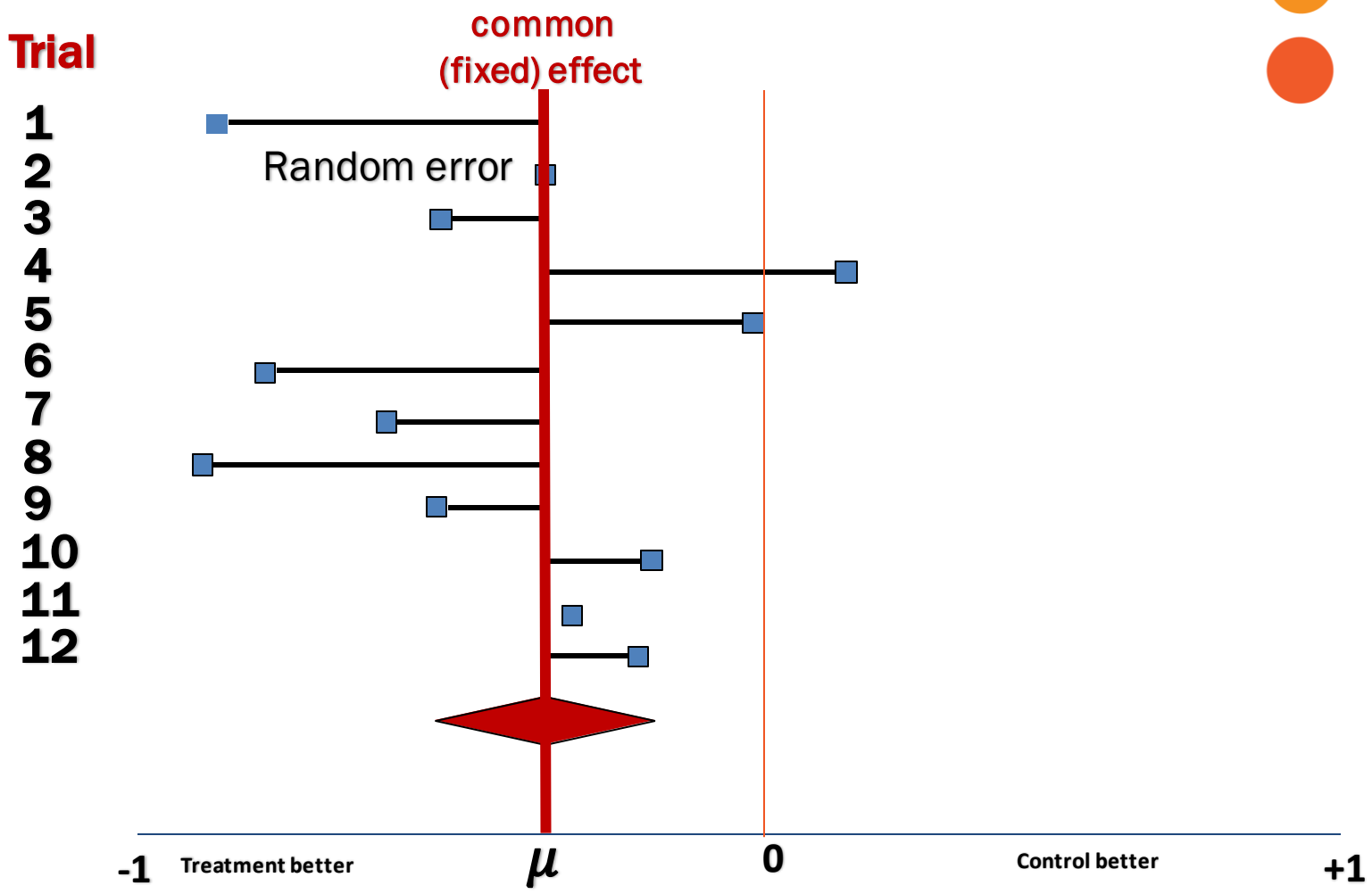
- **I^2 quantity**:
  - Note that I^2 is based on Q! $I^2 = [(Q - df) / Q] * 100\%$.
  - **Suggested categorizations (Higgins et al 2003):**
    - >25%= Low
    - >50% = moderate
    - >75% = High

Factors to choose between FE and RE???

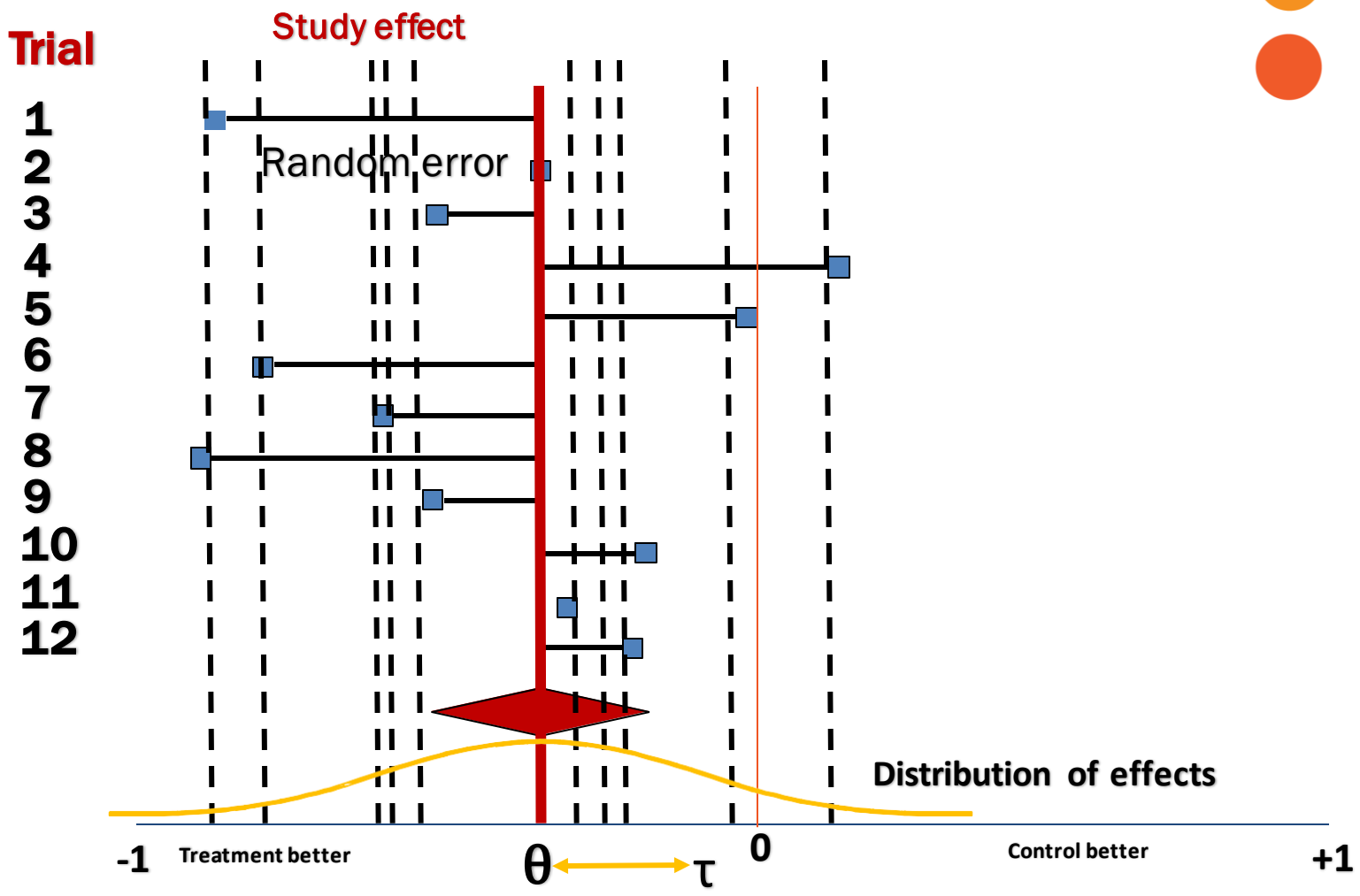
- In practice:
  - Random effects more conservative than fixed effects (i.e. wider 95% CIs); if no statistical heterogeneity, FE and RE return same pooled value.
  - Some choose FE/RE based on statistical heterogeneity measures Q and I^2 ; not ideal practice… better to think about clinical heterogeneity of studies and work with RE in general.
  - Easy to switch between models in software programs

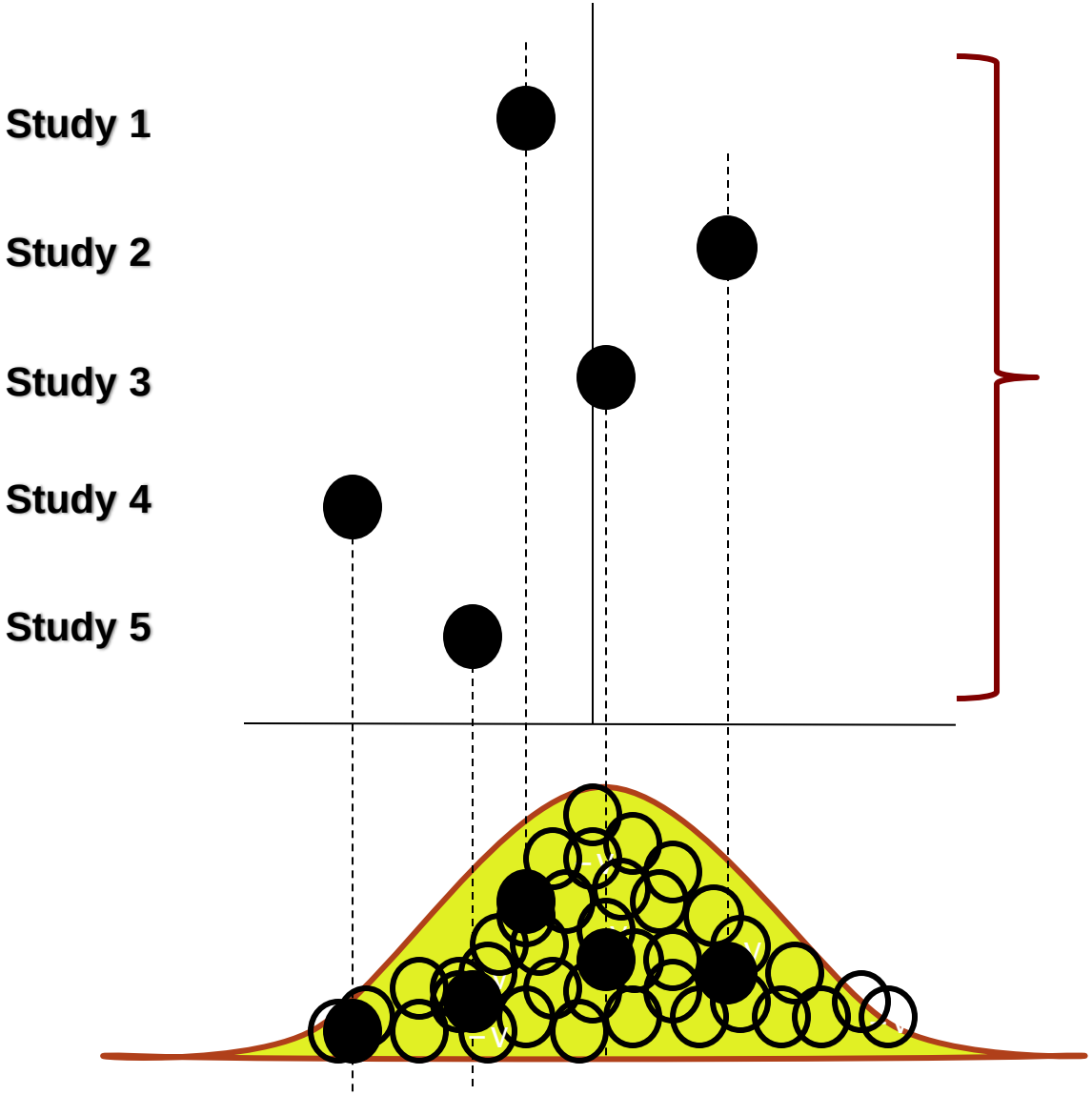
Bottom line: FE, RE are two ways to calculate the weights for our weighted average. Assumptions of RE generally more sensible.

Fixed effect meta-analysis



Random effect meta-analysis

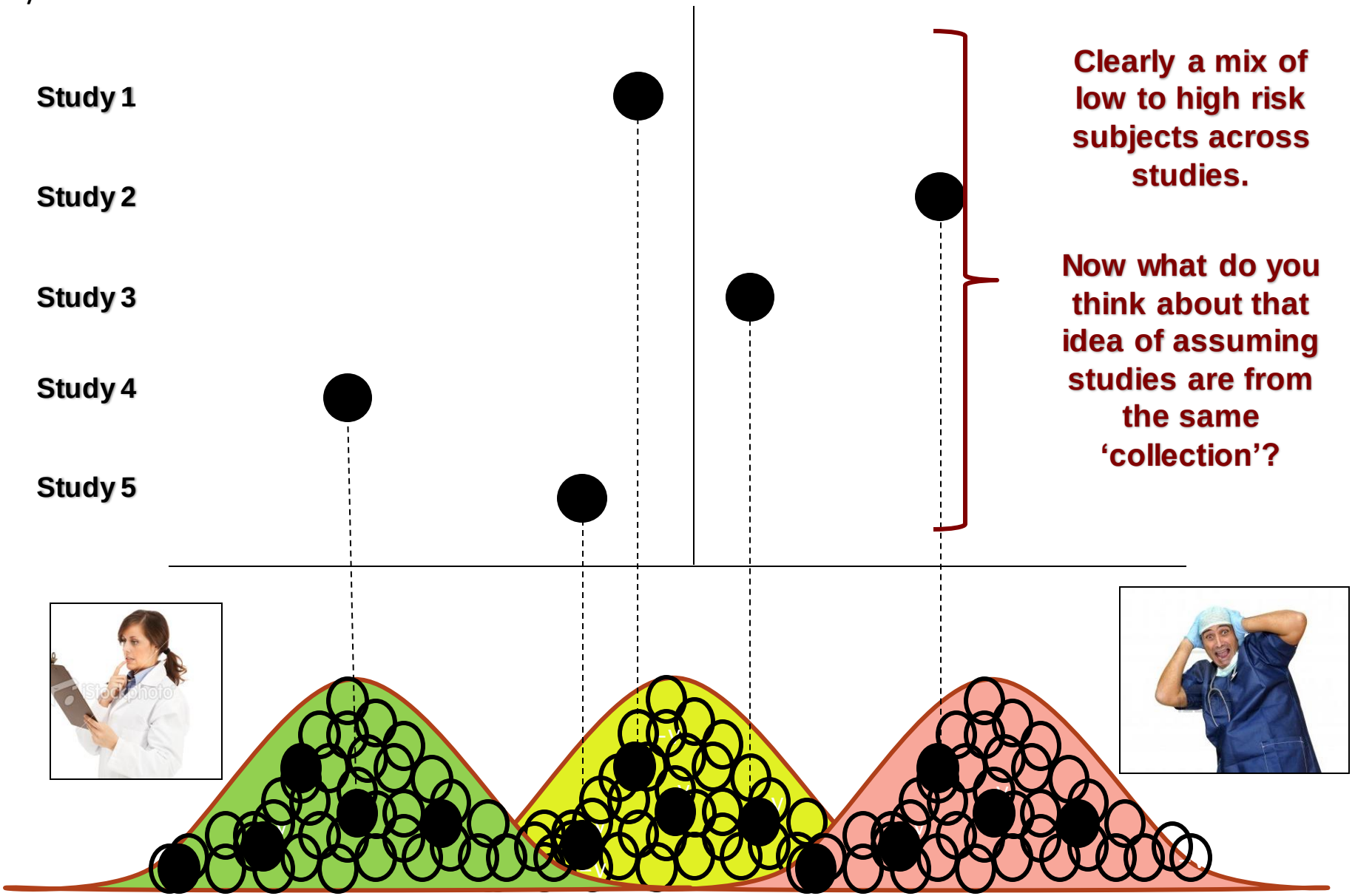




One source population, homogenous;
e.g. all first time, low risk surgery patients

We have 5 studies of seemingly similar ptnts, could argue they are from the same hypothetical collection of studies…

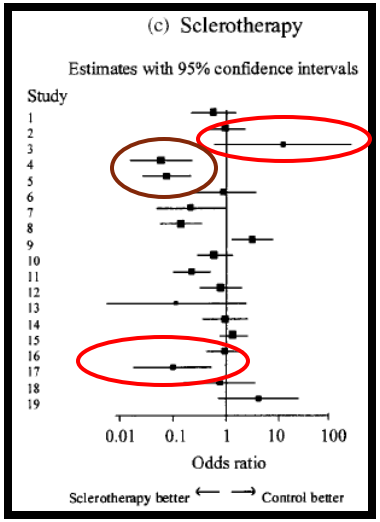
Fixed effects assumption perhaps okay???



Multiple source populations:
e.g. low risk, moderate risk, high risk surgery ptnts

Factors to choose between FE and RE???

- **Q, I^2 measure the degree of 'bouncing around of treatment effects' between studies; more than due to chance?**



- **$Q = \sum w_i*(logOR_i - logOR)^2$: i.e. in relation to the summary odds ratio, how much do the individual study ORs vary around it? Compare to a chi-square distribution…if 'large', need to consider RE model (as well as exploratory analyses and whether to pool)**

- **$I^2 = (Q - \#\ RCTs - 1) / Q$; range 0% - 100%: Q, I^2 a bit like a 'back-up plan'; if we miss important clinical or methods differences, 'significant statistical heterogeneity' tells us to go back and look again!**

Factors to choose between FE and RE???

What does all this matter for how I get weights?

- **Mathematically….**

Fixed effects: Weight (W) = 1 / Variance of each study

Random effects: Weight (W^*)= 1 / (Variance of each study)+(variance of between study)

$$\text{Pooled Effect (Y)} = \frac{Sum\ of (each\ study\ estimate\ x\ weight)}{sum\ of\ weight}$$

$$\text{with Variance} = \frac{1}{sum\ of\ Weights}$$

STEP 2: Formulae for fixed and random effects MA

"pooled estimate" $Y = \frac{\sum WiYi}{\sum Wi}$ 1

"study weights by FE" $W_i = 1/V_{Yi}$ 2

"Variance of pooled estimate" $V(Y) = \frac{1}{\sum Wi}$ 3

"Cochrane Q" $Q = \sum WiYi^2 - \frac{(\sum WiYi)^2}{\sum Wi}$ 4

"I^2" (Q-df/Q)*100% 5

"between study variance" (tau-square) $\tau^2 = Q - (k-1)/C$ 6

K is the number of study

"C" $= \sum Wi - \frac{\sum Wi^2}{\sum Wi}$ 7

"study weights by RE" $Wi = 1/(V_{Yi+T^2})$ 8

Statistical methods

- **Inverse-Variance (1930s)**
- Used with almost any measure with a standard error.

- **Generic Inverse Variance**
- Used for adjusted effects.

Review Manager 5 (RevMan 5)

- **Mantel-Haenszel (Mantel 1959, G and 1985)**
- Dichotomous measures only, Perfo be n inverse variance with sparse data.

- **Peto odds ratio method (Yusuf 1985)**
- Odds ratio only. Corrections for zero cell counts are not necessary. this method performs well when events are very rare.

Including analysis by subgroup

- **Advanced Methods**
- Maximum likelihood theory, Bayesian theory, Exact methods.

Generic Inverse Variance Method

Calculated effect estimate may be available.
(e.g. odds ratios from logistic regression or mean change from ANOVA or linear regression model).

only if they are accompanied by measures of uncertainty such as a standard error, 95% confidence interval or an exact P value.
The data should be entered as natural logarithms.

Software

- Review Manager (RevMan)
  - Note, set up for parallel designs.
  - Not for cross-over designs.
  - Cannot perform regression analysis.
- Comprehensive packages
  - Stata Intercooled, SAS, Comprehensive Meta-Analysis (not free).
  - Freeware: R.
- Your own coding
  - Whatever stats package you might use.
  - Excel.

  Generally, not advisable.

RevMan with dichotomous

Study level info

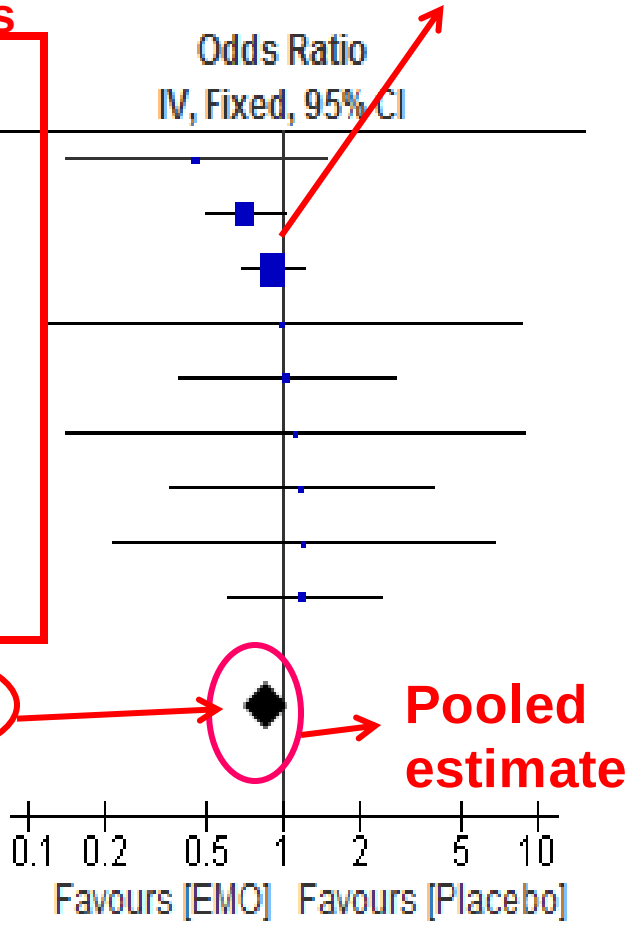
Study level estimates

No Effect line

Study or Subgroup	EMO Events	EMO Total	Placebo Events	Placebo Total	Weight	Odds Ratio M, Fixed, 95% CI
Silver et al	5	42	10	44	2.8%	0.46 [0.14, 1.48]
Corwin et al 3	62	733	83	727	32.0%	0.72 [0.51, 1.01]
Corwin et al 2	111	650	120	652	47.4%	0.91 [0.69, 1.21]
Van Iperen et al	2	12	2	12	0.8%	1.00 [0.12, 8.56]
Georgopolous et al	15	100	7	48	4.1%	1.03 [0.39, 2.73]
Still et al	2	19	2	21	0.9%	1.12 [0.14, 8.82]
Vincent et al	11	48	5	25	2.7%	1.19 [0.36, 3.90]
Ganriel et al	6	11	5	10	1.3%	1.20 [0.22, 6.68]
Corwin et al 1	24	80	21	80	8.0%	1.20 [0.60, 2.40]
Total (95% CI)		**1695**		**1619**	**100.0%**	**0.86 [0.71, 1.05]**
Total events	238		255			

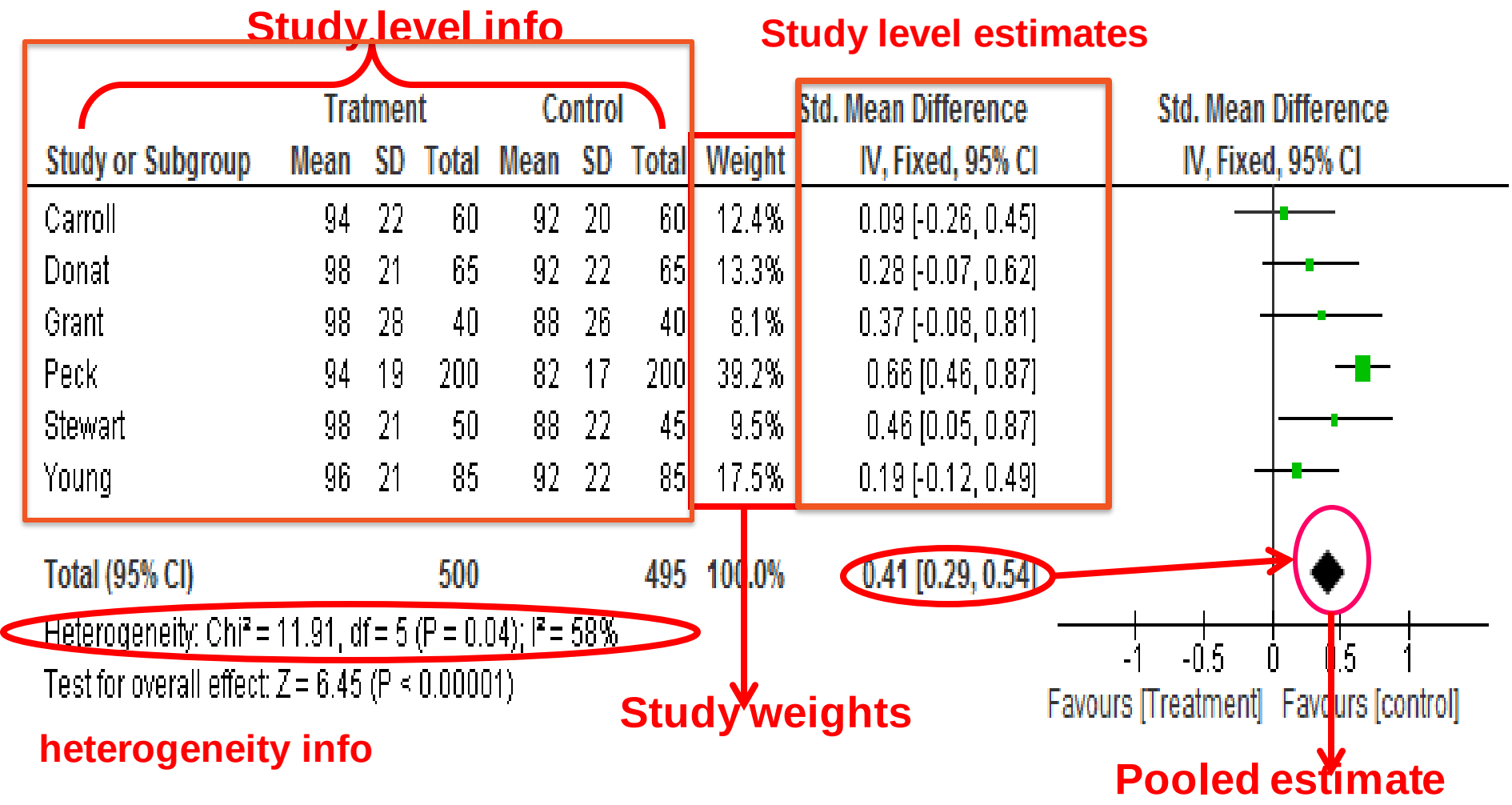
Heterogeneity: Chi² = 3.89, df = 8 (P = 0.87); I² = 0%

Test for overall effect: Z = 1.47 (P = 0.14)



Pooled estimate

RevMan with continuous outcome



Vote Counting

why “Vote counting” should be avoided:

1. Subjective decisions or statistical significance are used to define ‘positive’ and ‘negative’ studies (Cooper 1980, Antman 1992). In fact, small, moderate and even large effect sizes may yield a non significant p-value due to inadequate statistical power.

2. It takes no account of the differential weights given to each study.

Guideline in writing the analysis section of the protocol

1. Ensure that the analysis strategy firmly addresses the stated objectives of the review.
2. Consider which types of study design would be appropriate for the review. Parallel group trials are the norm, but other randomized designs may be appropriate to the topic (e.g. cross-over trials, cluster-randomized trials, factorial trials). Decide how such studies will be addressed in the analysis.
3. Decide whether a meta-analysis is intended and consider how the decision as to whether a meta-analysis is appropriate will be made.
4. Determine the likely nature of outcome data (e.g. dichotomous, continuous etc.).
5. Consider whether it is possible to specify in advance what intervention effect measures will be used (e.g. risk ratio, odds ratio or risk difference for dichotomous outcomes, mean difference or standardized mean difference for continuous outcomes).

Guideline in writing the analysis section of the protocol

6. Decide how statistical heterogeneity will be identified or quantified.

7. Decide whether random-effects meta-analyses, fixed-effect meta-analyses or both methods will be used for each planned meta-analysis.

8. Consider how clinical and methodological diversity (heterogeneity) will be assessed and whether (and how) these will be incorporated into the analysis strategy.

9. Decide how the risk of bias in included studies will be assessed and addressed in the analysis.

10. Pre-specify characteristics of the studies that may be examined as potential causes of heterogeneity.

11. Consider how missing data will be handled (e.g. imputing data for intention-to-treat analyses).

12. Decide whether (and how) evidence of possible publication and/or reporting biases will be sought.

Summary

- Meta-analysis is a way to combine data from RCTs of interest…weighted average of effects.

- To do: Extract relevant data from each study, and place into software program.

- Can determine weights by choosing between fixed effects and random effects MA.

- Can present study findings and pooled estimate using a forest plot. Can interpret relevance of pooled estimate by examining 95% CI, p-value.

Mathematically can always pool, but must be sensible and examine study characteristics closely.

GEORGE & FAY YEE

Centre for Healthcare Innovation

COLLABORATION INNOVATION TRANSFORMATION

University of Manitoba

Winnipeg Regional Health Authority
Caring for Health

Office régional de la santé de Winnipeg
À l'écoute de notre santé