Attribution Key

for more information see: http://open.umich.edu/wiki/AttributionPolicy

Use + Share + Adapt

{ Content the copyright holder, author, or law permits you to use, share and adapt. }

PD-GOV **Public Domain – Government:** Works that are produced by the U.S. Government. (USC 17 § 105)

PD-EXP **Public Domain – Expired:** Works that are no longer protected due to an expired copyright term.

PD-SELF **Public Domain – Self Dedicated:** Works that a copyright holder has dedicated to the public domain.

ZERO **Creative Commons – Zero Waiver**

BY **Creative Commons – Attribution License**

BY-SA **Creative Commons – Attribution Share Alike License**

BY-NC **Creative Commons – Attribution Noncommercial License**

BY-NC-SA **Creative Commons – Attribution Noncommercial Share Alike License**

GNU-FDL **GNU – Free Documentation License**

Make Your Own Assessment

{ Content Open.Michigan believes can be used, shared, and adapted because it is ineligible for copyright. }

PD-INEL **Public Domain – Ineligible:** Works that are ineligible for copyright protection in the U.S. (USC 17 § 102(b)) *laws in your jurisdiction may differ

{ Content Open.Michigan has used under a Fair Use determination. }

FAIR USE **Fair Use:** Use of works that is determined to be Fair consistent with the U.S. Copyright Act. (USC 17 § 107) *laws in your jurisdiction may differ

Our determination **DOES NOT** mean that all uses of this 3rd-party content are Fair Uses and we **DO NOT** guarantee that your use of the content is Fair.

To use this content you should **do your own independent analysis** to determine whether or not your use will be Fair.

LECTURE 3: ANALYZING AND INTEGRATING THE OUTCOMES OF STUDIES

Joel J. Gagnier MSc, PhD

Overview

- What to do with your data
- Principles of Meta-analysis
- Meta-analysis models
- Effect Size Metrics
- Best evidence synthesis

Principles of Meta-analysis

- Summary statistic is calculated for each study
- A summary (pooled) intervention effect estimate is calculated as a weighted average of the intervention effects from individuals studies
  - Weights represent how much each study contributes to the overall estimate
- SE is used to
  - derive a confidence interval which communicates the precision of the summary estimate
  - Derive a p-value (strength of the evidence against the null hypothesis)

What do we need??

- Complete data from each study on our selected outcomes
- There are many types of data available and there may need to be some conversions done to use it

Analyze and present results

- Sometimes sufficient data are not provided for statistics to be done
  - Contact authors, ask for it
  - Impute data (e.g. SE; Cochrane handbook; various methods)
  - Do meta-analysis on the subset of studies with complete data

  OR
  - Do a qualitative review

Qualitative SR: Best evidence synthesis

- Consider quality of studies and results
- Give a summary table of study characteristics
- Write in the text your overall synthesis and related conclusions
  - Group by condition, intervention and outcomes

"The levels of evidence were defined as follows:

1. Strong – consistent findings among multiple high quality RCTs
2. Moderate – consistent findings among multiple low quality RCTs and/or CCTs and/or one high quality RCT
3. Limited – one low quality RCT and/or CCT
4. Conflicting – inconsistent findings among multiple trials (RCTs and/or CCTs)
5. No evidence from trials – no RCTs or CCTs"

Van Tulder 1997, 2003.

Meta-analysis

- The effect size (ES) makes meta-analysis possible
- The ES encodes the selected research findings on a numeric scale
- There are many different types of ES measures, each suited to different research situations
- Each ES type may also have multiple methods of computation

Types of Data

- Dichotomous (e.g., life and death)
  - 2X2 tables
  - RR (log), OR (log), RD (sensitive to baseline differences)
- Continuous (e.g., cholesteral levels)
  - Mean differences
  - Standardized mean differences (if different scales)
    - Mean difference for each study divided by the within sudy variance (SD) for that scale
  - Response ratios
- Correlational
  - Between two continuous variables

What Makes Something an Effect Size for Meta-analytic Purposes

- The type of ES must be comparable across the collection of studies of interest
  - May be the case for your research
  - May be accomplished through standardization
- Must be able to calculate a standard error for that type of ES
  - The standard error is needed to calculate the ES weights, called inverse variance weights (more on this latter)
- All meta-analytic analyses are weighted

The Mean Difference

$$\overline{ES} = \frac{\bar{X}_{G1} - \bar{X}_{G2}}{s_{pooled}} \qquad s_{pooled} = \sqrt{\frac{s_1^2(n_1 - 1) + s_2^2(n_2 - 1)}{n_1 + n_2 - 2}}$$

- Represents a group contrast on a *continuous* measure
- Uses the pooled standard deviation (some situations use control group standard deviation)
- Commonly called “d” or occasionally “g”

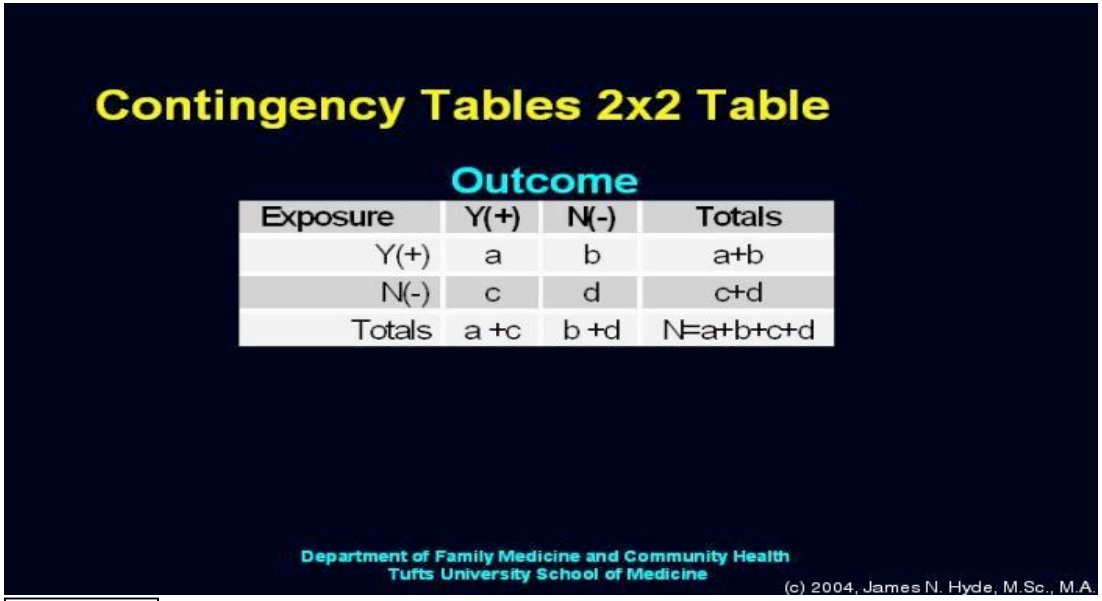
The Correlation Coefficient

$$\overline{ES} = r$$

- Represents the strength of association between two *inherently continuous* measures
- Generally reported directly as "r" (the Pearson product moment coefficient)

The Odds-Ratio

- The odds-ratio is based on a 2 by 2 contingency table, such as the one below

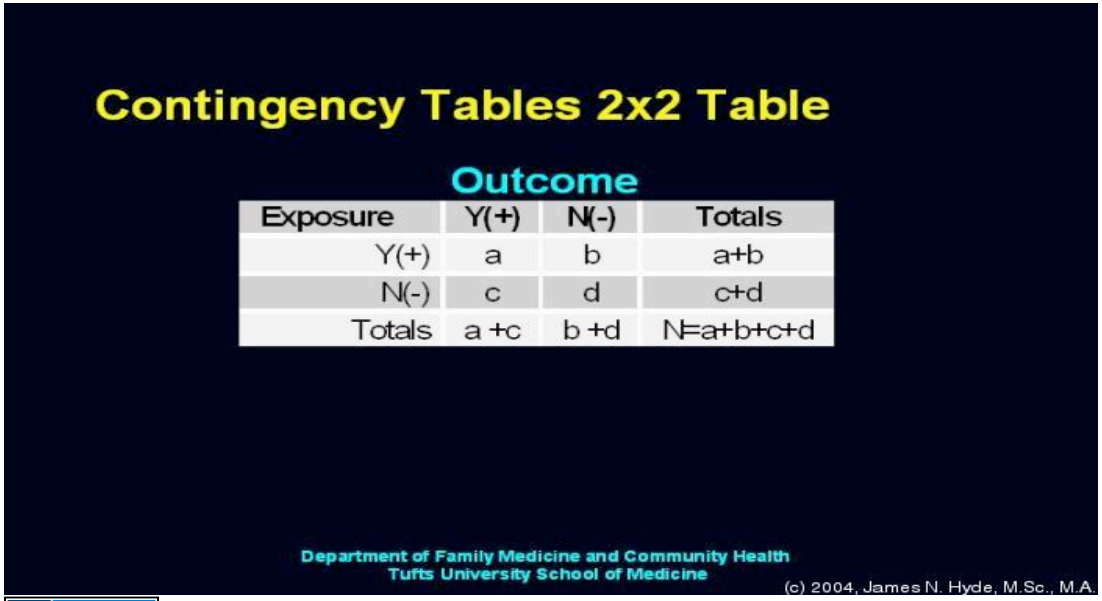


$$\overline{ES} = \frac{ad}{bc}$$

- The Odds-Ratio is the odds of success in the treatment group relative to the odds of success in the control group.

The Rate Ratio

- The rate ratio is based on a 2 by 2 contingency table, such as the one below



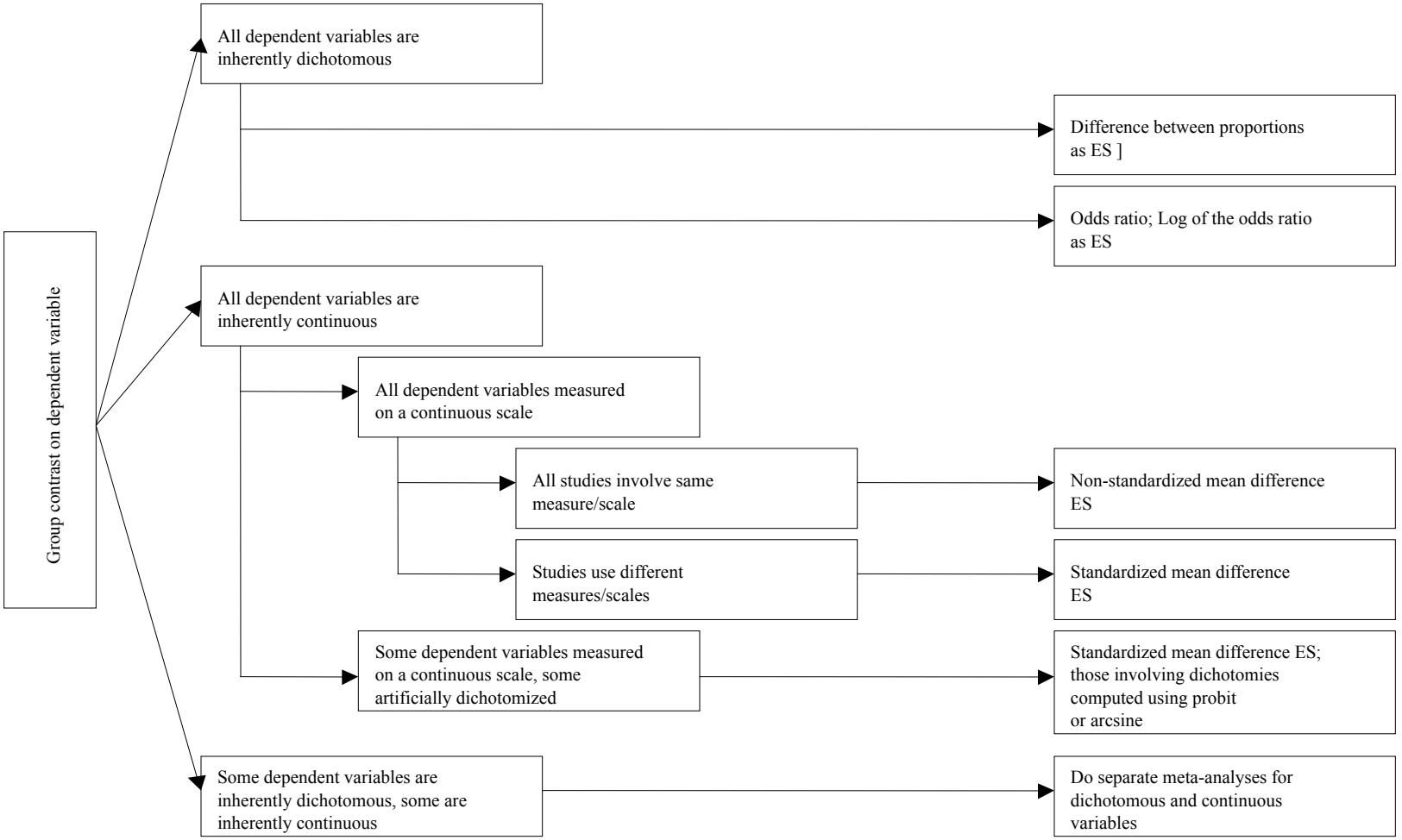
$$\overline{ES} = \frac{a/a+b}{c/c+d}$$

- The Rate Ratio is the success in the treatment group relative to the success in the control group.

Non-standardized Effect Size Metric

- Synthesizing a research domain that uses a common measure across studies
  - May wish to use an effect size that is non-standardized, such as a simple mean difference (e.g., LDL cholesterol)

Effect Size Decision Tree for Group Differences Research



Calculating the Standardized Mean Difference

- There are many methods
- Not within the scope of this course to cover them all
- See Chapter 4 of Borenstein

Methods of Calculating the Mean Difference

Direction Calculation Method

$$ES = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{s_1^2(n_1 - 1) + s_2^2(n_2 - 1)}{n_1 + n_2 - 2}}} = \frac{\bar{X}_1 - \bar{X}_2}{s_{pooled}}$$

Formulas for the Correlation Coefficient

- Results typically reported directly as a correlation
- Any data for which you can calculate a standardized mean difference effect size, you can also calculate a correlation type effect size
- See Chapter 6 of Borenstein

Formulas for the Odds Ratio

- Results typically reported in one of three forms:
  - Frequency of successes in each group
  - Proportion of successes in each group
  - 2 by 2 contingency table
- See Chapter 5 of Borenstein

Issues in calculating Effect Sizes

- Which formula to use when data are available for multiple formulas
- Multiple documents/publications reporting the same data (not always in agreement)
- Different time points reported
- How much guessing should be allowed
  - sample size is important but may not be presented for both groups
  - some numbers matter more than others

Overview of Meta-Analytic Data Analysis

- Transformations
- Converting among effect sizes (chapter 7 of Borenstein)
- The Inverse Variance Weight
- The Pooled Effect Estimate and Associated Statistics
  - Fixed OR Random Effects
- Test for Homogeneity
- Fixed Effects exploration of heterogeneity
  - Fixed Effects Analog to the one-way ANOVA
  - Fixed Effects meta-regression Analysis
- Random Effects exploration of heterogeneity
  - Random Effects Analog to the one-way ANOVA
  - Random Effects meta-regression Analysis

Transformations

- Odds-Ratio is asymmetric and has a complex standard error formula.
  - Negative relationships indicated by values between 0 and 1.
  - Positive relationships indicated by values between 1 and infinity.
- Solution: Natural log of the Odds-Ratio.
  - Negative relationship < 0.
  - No relationship = 0.
  - Positive relationship > 0.
- Finally results can be converted back into Odds-Ratios by the inverse natural log function.

Transformations (continued)

- Analyses performed on the natural log of the Odds-Ratio:

$$ES_{LOR} = \ln[OR]$$

- Finally results converted back via inverse natural log function:

$$OR = e^{ES_{LOR}}$$

Independent Set of Effect Sizes

- Must be dealing with an independent set of effect sizes before proceeding with the analysis.
  - One ES per study

  OR
  - One ES per subsample within a study

Meta-analysis model Assumptions

Fixed effects

- All studies are estimating the same "true" underlying effect
- Variability between studies is due to random variation (chance) only

Random effects

- There is a distribution of effects depending on the study methods
- Variability between studies is due to random variation (chance) and their methods

Characteristics of Fixed VS Random Effects models

	Fixed Effects	Random Effects
Pooled estimate weighted by	Within study variance	Within and between study variance
Weighting of small studies	Smaller weights	Larger weights
Weighting of large studies	Larger weights	Smaller weights
Confidence intervals (generally)	Narrow (within study variance only)	Wide (due to including variance between sudies)
In the presence of significant/ substantial statistical heterogeneity	Narrow (do not reasonably account for the variance in effect estimates between the studies)	Even wider
Effect size (pooled summary effect)	Estimate of some common effect size; use Z	Estimate of the mean of a normal distribution of effect sizes: use t distrib

Choice of models

- If you are looking for a single best effect estimate, and the studies appear to be homogeneous, then a fixed effects model is preferable
- If there is evidence of heterogeneity that cannot be explained
  - Random effects approach is preferrable
  - But … consider
- If studies are not likely functionally equivalent, and goal is to generalize to a range of scenerios
  - Random effects

Methods for combining study estimates

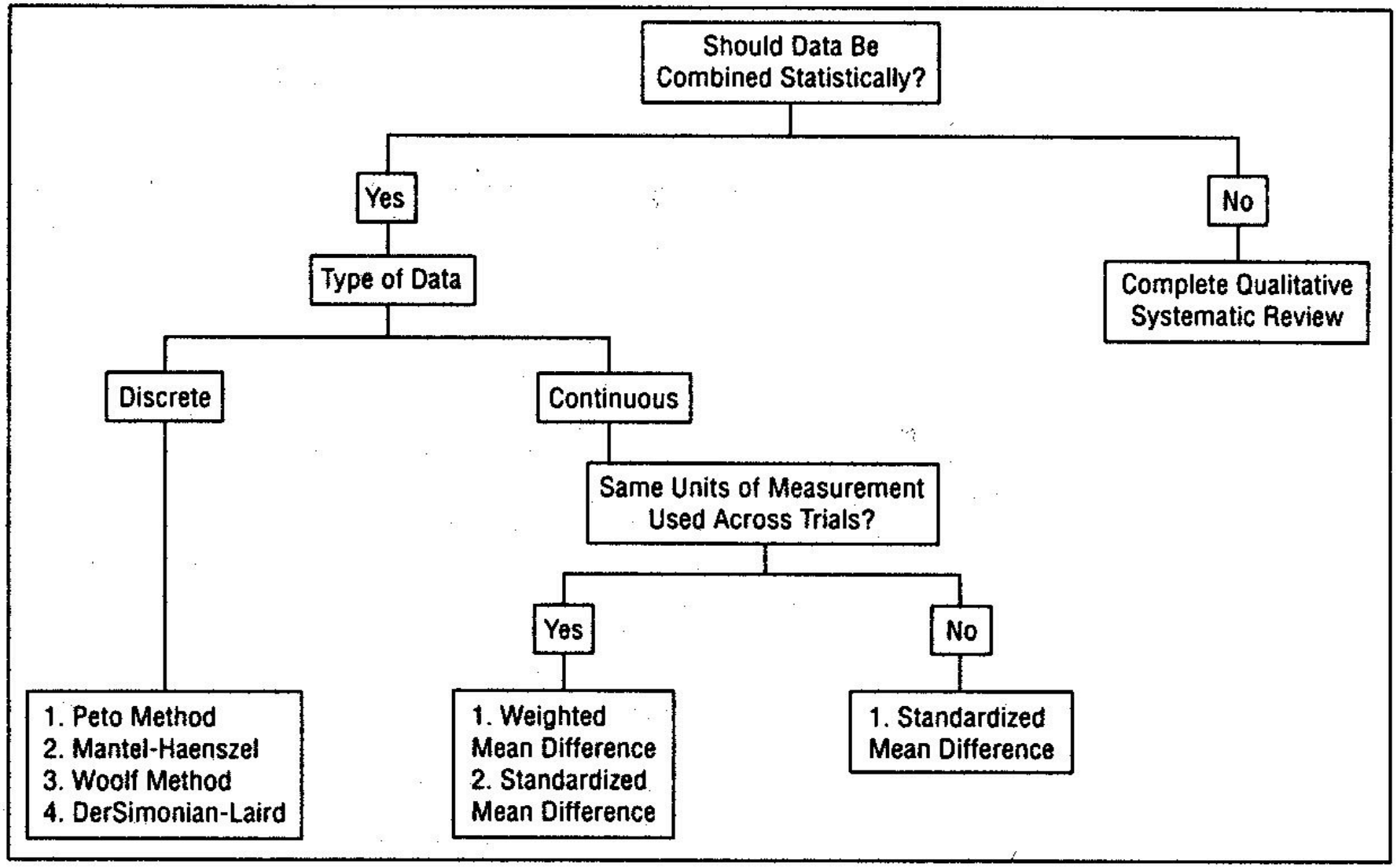
Fixed Effects Methods

- Methods:
  - Small number of studies, but sample sizes of the studies are large
    - Inverse variance
      - Preferred
  - If sparse data
    - Mantel & Haenszel method (RR, OR etc)
      - Preferred if many studies , each being small
      - Use continuity correction (add 0.5 to a cell) if cell is 0
    - Peto method for OR
      - Can be used if you have 0 in a cell of the 2x2 table
      - Produces serious underestimates when OR is large
      - Many criticisms
  - For large sample sizes
    - Maximum likelihood method

Methods for combining study estimates

Random Effects Methods

- Weighted least squares (WLS) regression
  - Inverse variance weighting
  - Most common method
  - Preferred
  - Dersimonian & Laird method
    - Method of moments estimate for tau-sq
- Maximum likelihood method (ML)
  - Assumes normality of the underlying effect distribution
    - Has another estimate for tau-sq
- Restricted maximum likelihood method (REML)
  - Assumes normality of the underlying effect distribution
- Exact method suggested by Van Houwelingen
  - Non-parametric approach if normality assumption is violated



Algorithm of statistical choices available to systematic reviewers.

 Moher D et al. Arch Pediatr Adolesc Med 1998;152:915-20

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.

What is 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 Effect Sizes

- Standardized Mean Difference:

$$se = \sqrt{\frac{n_1 + n_2}{n_1 n_2} + \frac{\overline{ES}_{sm}}{2(n_1 + n_2)}} \qquad w = \frac{1}{se^2}$$

Inverse Variance Weight for the 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.

Ready to Analyze

- We have an independent set of effect sizes (ES) that have been transformed and/or adjusted, if needed.
- For each effect size we have an inverse variance weight (w).

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 Weighted Mean Effect Size

Study	ES	w	w*ES
1	-0.33	11.91	-3.93
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.
- Next, multiply w by ES.

The Weighted Mean Effect Size

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

- Start with the effect size (ES) and inverse variance weight (w) for 10 studies.
- Next, multiply w by ES.
- Repeat for all effect sizes.

The Weighted Mean Effect Size

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

- Start with the effect size (ES) and inverse variance weight (w) for 10 studies.
- Next, multiply w by ES.
- Repeat for all effect sizes.
- Sum the columns, w and ES.
- Divide the sum of (w*ES) by the sum of (w).

$$\overline{ES} = \frac{\sum (w \times ES)}{\sum w} = \frac{41.82}{269.96} = 0.15$$

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.

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

Mean, Standard Error, Z-test and Confidence Intervals

Mean ES

$$\overline{ES} = \frac{\sum(w \times ES)}{\sum w} = \frac{41.82}{269.96} = 0.15$$

SE of the Mean ES

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

Z-test for the Mean ES

$$Z = \frac{\overline{ES}}{se_{\overline{ES}}} = \frac{0.15}{0.061} = 2.46$$

95% Confidence Interval

$$Lower = \overline{ES} - 1.96(se_{\overline{ES}}) = 0.15 - 1.96(.061) = 0.03$$

$$Upper = \overline{ES} + 1.96(se_{\overline{ES}}) = 0.15 + 1.96(.061) = 0.27$$

Random Effects Models

- Don’t panic!
- It sounds far worse than it is.
- Three reasons to use a random effects model
  - Cochran’s Q test (a test of statistical homogeneity) is significant (the studies are heterogeneous) and you assume that the excess variability across effect sizes derives from random differences across studies (sources you cannot identify or measure)
  - The Q within from an Analog to the ANOVA is significant
  - The Q residual from a Weighted Multiple Regression analysis is significant

The Logic of a Random Effects Model

- Fixed effects model assumes that all of the variability between effect sizes is due to sampling error
  - In other words, instability in an effect size is due simply to subject-level "noise"
- 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 (within study 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 (between study variance).

$$w_i = \frac{1}{se_i^2 + \hat{v}_\theta}$$

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 for the random effects variance component is:

$$\hat{v}_\theta = \frac{Q_T - k - 1}{\sum w - \left(\frac{\sum w^2}{\sum w} \right)}$$

Calculation of the Random Effects Variance Component: Q first

Study	ES	w	w*ES	w*ES^2	w^2
1	-0.33	11.91	-3.93	1.30	141.73
2	0.32	28.57	9.14	2.93	816.30
3	0.39	58.82	22.94	8.95	3460.26
4	0.31	29.41	9.12	2.83	865.07
5	0.17	13.89	2.36	0.40	192.90
6	0.64	8.55	5.47	3.50	73.05
7	-0.33	9.80	-3.24	1.07	96.12
8	0.15	10.75	1.61	0.24	115.63
9	-0.02	83.33	-1.67	0.03	6944.39
10	0.00	14.93	0.00	0.00	222.76
		269.96	41.82	21.24	12928.21

- Calculate a new variable that is the w squared.
- Sum new variable.

Calculating Q

We now have 3 sums:

$$\sum w = 269.96$$
$$\sum (w \times ES) = 41.82$$
$$\sum (w \times ES^2) = 21.24$$

Q is can be calculated using these 3 sums:

$$Q = \sum (w \times ES^2) - \frac{\left[\sum (w \times ES)\right]^2}{\sum w} = 21.24 - \frac{41.82^2}{269.96} = 21.24 - 6.48 = 14.76$$

The Q statistic

Allows us to check for **statistical heterogeneity**

- Are differences b/w trials > expected by chance?
- Cochran's Q = WSS= sum Wi $(Yi-M)^2$ (true variation and chance variation)
  - A test for the presence of statistical homogeneity (Ho= no difference between groups)
- Compared to the Chi-squared distribution
  - too little power with a collection of studies with small sample sizes
  - too much power with a collection of studies with large sample sizes
  - P usually set at 0.10 since has low power with small samples (as is mostly the case….SRs N=6-8 on average)

Calculation of the Random Effects Variance Component

- The total Q for this data was 14.76
- k is the number of effect sizes (10)
- The sum of w = 269.96
- The sum of w^2 = 12,928.21

$$\hat{v}_\theta = \frac{Q_T - k - 1}{\sum w - \left(\frac{\sum w^2}{\sum w}\right)} = \frac{14.76 - 10 - 1}{269.96 - \frac{12{,}928.21}{269.96}} = \frac{5.76}{269.96 - 47.89} = 0.026$$

Rerun Analysis with New Inverse Variance Weight

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

- Calculate a new weight.

$$w_i = \frac{1}{se_i^2 + \hat{v}_\theta}$$

- Rerun analysis.
- Congratulations! You have just performed a very complex statistical analysis.

The Weighted Mean Effect Size: Random Effects

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)
- Now have new random effects inverse variance weight (w) for 10 studies.

$$w_i = \frac{1}{se_i^2 + \hat{v}_\theta}$$

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

The Weighted Mean Effect Size

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

- Start with the effect size (ES) and inverse variance weight (w) for 10 studies.
- Next, multiply w by ES.
- Repeat for all effect sizes.
- Sum the columns, w and ES.
- Divide the sum of (w*ES) by the sum of (w).

$$\overline{ES} = \frac{\sum (w \times ES)}{\sum w} = \frac{41.82}{269.96} = 0.15$$

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.

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

Mean, Standard Error, Z-test and Confidence Intervals

Mean ES

$$\overline{ES} = \frac{\sum(w \times ES)}{\sum w} = \frac{41.82}{269.96} = 0.15$$

SE of the Mean ES

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

Z-test for the Mean ES

$$Z = \frac{\overline{ES}}{se_{\overline{ES}}} = \frac{0.15}{0.061} = 2.46$$

95% Confidence Interval

$$Lower = \overline{ES} - 1.96(se_{\overline{ES}}) = 0.15 - 1.96(.061) = 0.03$$

$$Upper = \overline{ES} + 1.96(se_{\overline{ES}}) = 0.15 + 1.96(.061) = 0.27$$

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.

Work on protocols!!!