An Introduction to Meta-analysis

Geoff Der
Unit Statistician
MRC/CSO Social and Public Health Sciences Unit,
University of Glasgow

1

In this session:

What is meta-analysis?

When is it appropriate to use?

Statistical methods

Software programmes

Publishing meta-analyses

2

Karl Pearson (1904) conducted the first meta-analysis commissioned by the British government on the effects of a typhoid vaccination

Gene Glass (1974) coined 'meta-analysis':
"*…the analysis of analyses. It connotes a rigorous alternative to the casual, narrative discussions of research studies which typify our attempts to make sense of the rapidly expanding research literature".*

3

What is meta-analysis?

"*Statistical combination of results from two or more separate studies*" to answer a common question

Why?

To provide a test with more power than separate studies

To summarise numerous and inconsistent findings

To investigate consistency of effect across different samples

http://www.cochrane-handbook.org/

4

What questions are addressed?

1. What is the direction of the effect?
2. What is the size of the effect?
3. Is the effect consistent across studies? (heterogeneity)
4. What is the strength of evidence for the effect? (quality assessment)

http://www.cochrane-handbook.org/

5

Some Background – Clinical Trials

Early trials show larger effects than later trials

Better designed trials show smaller effects

Larger trials show smaller effects

'Natural history' of novel interventions

Proliferation of small underpowered trials

Pocock, S J Clinical Trials A Practical Approach. Wiley 1983.

6

When is it appropriate?

Observational and Intervention studies

How many studies make it worth while?

Are there additional exclusion criteria for meta-analyses?
- Duplicate publications, e.g. in longitudinal studies
- Very small studies
- Poor quality
- Results not in suitable format? (But can approach authors)

Greenland, *Epidemiologic Reviews* 1987;9

7

Statistical Issues

Effect size measures
transformations; direction and magnitude of effect

Models: random vs fixed effects

Heterogeneity

Publication bias

Quality assessment and sensitivity analyses:
bias and confounding; subgroup analysis or meta-regression?

8

Effect Size Measures

Continuous outcome
Standardised Mean difference
- Cohen's d
- Hedges' g
- Glass's Δ

$$d = \frac{\bar{x}_1 - \bar{x}_2}{s},$$

Binary outcome
- Odds Ratio
- Relative Risk

Survival
- Hazard ratio

9

Effect Size Measures

Transform reported effect sizes to common measure
Eg measures of spread/variance: CI, SD, SE, IQR

Converting odds ratios to continuous outcome effect sizes, or vice versa (Chinn, *Statistics in Medicine*, 2000;19:3127)

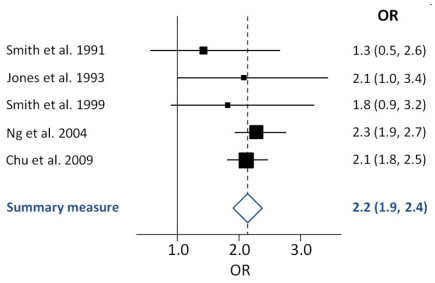
HR ~ OR ~ RR when the risk of an event is low: <20%
(Symons et al, *J Clin Epidemiol*, 2002;55:893-99).

Take care and check results!

Online effect size calculator:
http://www.campbellcollaboration.org/resources/effect_size_input.php

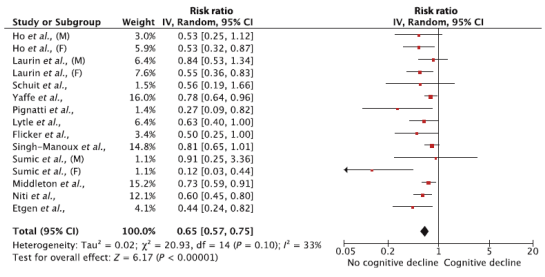
10

Forest Plots – the main output of MA



11

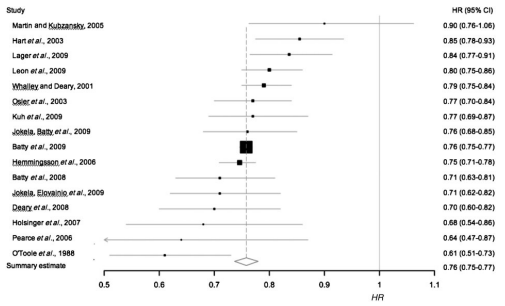
Example: RR with CIs



High physical activity & Cognitive decline (Sofi et al, *J Internal Med*, 2010;269:107-117)

12

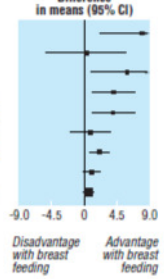
Example: HR with CIs



Childhood IQ and risk of mortality (Calvin et al., 2010)

Forest Plot– mean difference

Study name	Statistics for each study: Difference in means (SE)	P value	N	Age (years)	IQ test
Hay	7.92 (2.94)	0.007	130	11	WISC
Ghys	0.21 (2.82)	0.940	108	4	GOS
Morrow-Tlucak	5.83 (2.57)	0.023	219	2	MDI
Gomez-Sanchiz	3.98 (1.53)	0.009	164	2	MDI
Jacobson	4.00 (1.49)	0.007	279	11	WISC
Wigg	0.80 (1.37)	0.560	343	12	WISC
Fergusson	2.09 (0.65)	0.001	954	7	WISC
Richards	0.98 (0.61)	0.109	511	8	SC
NLSY	0.52 (0.36)	0.149	5475	10	PIAT



Example: standardised mean difference

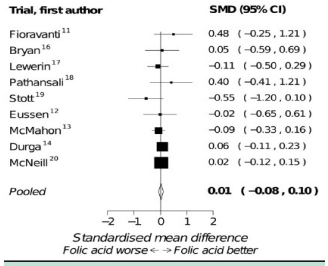


Figure 1 Standardized mean differences in cognitive function test scores (and 95% confidence intervals) between folic acid, with or without other B vitamins, and placebo groups for all tests within each randomized trial and pooled effect across all trials. CI = confidence interval; SMD = standardized mean difference.

Wald et al, *Am J Medicine*, 2010;123(6):522-7

Fixed vs Random effects

Fixed Effects
Each study is estimating the same quantity

Random Effects
Differences in study sample, design, measurement etc contribute to the effect size

DerSimonian and Laird method

Heterogeneity

Variability between studies caused by differences in:
- Study samples (e.g. healthy, clinical)
- Interventions or outcomes
- Methodology: design, measures, quality etc.

"*Statistical heterogeneity manifests itself in the… [study] effects being more different from each other than one would expect due to random error (chance) alone*" (Cochrane Handbook)

Assessing heterogeneity

- Visual inspection:
  - confidence intervals have poor overlap
- Formal test:
  - Chi-squared: are observed differences compatible with chance alone?
  (NB. low power with small number of studies; $p >$ 0.10 gives greater confidence of no heterogeneity)
- Additionally, look at the impact of heterogeneity on your aggregate estimate: inconsistency ($I^2 > 50\%$)

But, isn't there always clinical and methodological diversity?

Dealing with heterogeneity

- Check data!
- Choose random effect meta-analysis
- Explore the causes of heterogeneity:
  subgroup analysis or **meta-regression**
- Change the effect measure
- Exclude outlying studies
- Consider whether a meta-analysis is the right course

Must be dealt with sensitively and with a good rationale for the methods used

19

Subgroup Analyses

Dividing your studies by a design feature:
- Participant characteristic (sex, age, clinical diagnoses, geographical region)
- Study design characteristic (type of intervention, length of follow-up, type of measure used, e.g. cognitive function)

NB. More subgroup analyses increase the risk of false negatives and false positives (patients being denied an effective treatment, or given a harmful / ineffective one)

20

Meta-regression

Linear regression of the effect estimates on some study characteristic

Outcome: Study effect size

Explanatory variable: a characteristic of the studies that may influence the magnitude of the effect
(potential effect modifier or covariate)

Regression is weighted by study size/precision

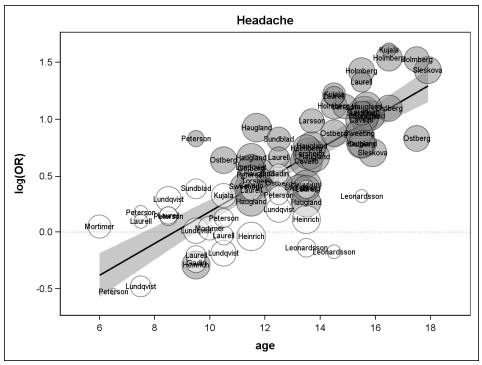
21

Subgroup analyses & meta-regression Considerations of both

- Are there enough studies that include the specified characteristics to justify these methods?
- Specify the characteristics in advance
- Keep numbers of characteristics to a minimum
- Is there adequate scientific rationale?
- Does one characteristic confound another?

22

Meta-regression example



23

Publication bias / small study bias: Addressing file drawer effects

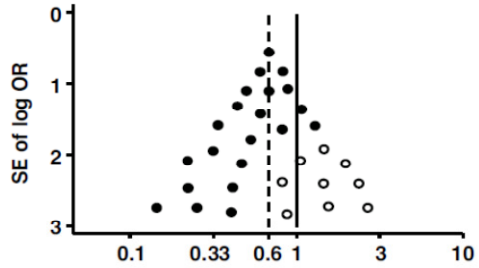
*"To control resulting overall effect sizes for publication bias, several tests were performed. These tests consisted of **visual inspection of funnel plots** (Light & Pillemer, 1984), **Rosenthal's Fail-safe N** (Rosenthal, 1979), a **weighted Failsafe N** (Rosenberg, 2005), **Orwin's Fail-safe N** (Orwin, 1983), **Begg and Mazumdar's rank correlation method** (Begg & Mazumdar, 1994), **Egger's regression test** (Egger, Smith, Schneider, & Minder, 1997; Sterne & Egger, 2005), **trim-and-fill analysis** (Duval & Tweedie, 2000) following the approach as suggested by Peters, Sutton, Jones, Abrams and Rushton (2007), a **sensitivity analysis for publications bias** as suggested by Vevea and Woods (2005), and **a method based on truncated normal distributions** (Formann, 2008).*

Application of this multitude of differential approaches originates in the increased awareness of problems of publication bias in general and the corresponding recent developments of enhanced methods to account for it."

Pietschnig et al, *Intelligence*, 2010;38:314-23.

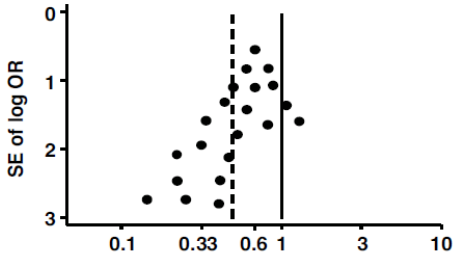
24

Funnel plots



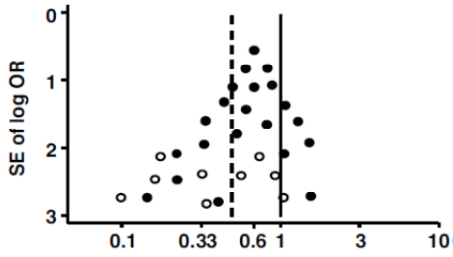
25

Funnel plots



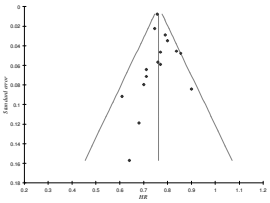
26

Funnel plots



27

Publication bias: Example 1 cognitive epidemiology



Childhood IQ and risk of mortality (Calvin et al., 2010)

28

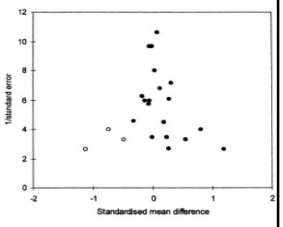
Trim-and-fill

1 trim off the asymmetric part of the funnel

2 use the symmetric remainder to estimate the true centre

3 replace the trimmed studies and their missing counterparts

4 estimate the true mean and its variance from the filled funnel plot



Duval & Tweedie, Biometrics, 2000;56(2):455-63.

29

Quality assessment

To control for bias, particularly in observational studies

Use a quality checklist/tool:
(eg Moher, 1995 for RCTs ; Sanderson, 2007 for observational studies)

Independent quality scoring (and blinded, if poss)

1. Forest Plot ordered by quality score.
   Is there an association?
2. Quality score as in meta-regression

3. Exclude low quality studies

30

Software: specially built programmes

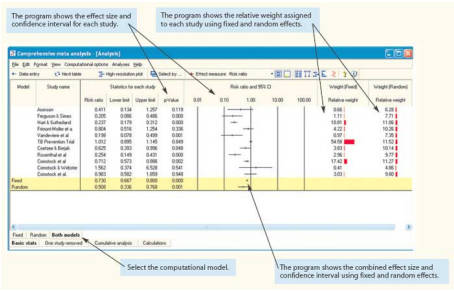
Comprehensive Meta-Analysis (CMA)
MetAnalysis
MetaWin
MIX – no longer *Free*
RevMan - *Free*
WEasyMA

Table 5: Meta-analysis software – usability ratings (best scoring software from left to right)

Items and subgroups	MIX	CMA	MetaWin	RevMan	WEasyMA
All researchers (36)					
Overall rating (min-max)	8.6 (6.7 to 10)	6.9 (3.7 to 9.7)	6.2 (4.3 to 8.7)	6.1 (4.3 to 8.3)	4.2 (1 to 7.3)
Getting started	8.6	7.4	6.8	7.6	4.5
Data preparation	8.3	6.3	6.3	4.5	2.6
Usability in analysis	8.8	7.1	5.6	6.3	3.9
Experienced (7)					
Overall rating (min-max)	8.1 (7.0 to 9.7)	6.8 (6.0 to 7.3)	5.9 (4.3 to 7.7)	5.4 (4.3 to 6.3)	3.3 (1 to 5.7)
Getting started	8.0	7.6	6.2	7.5	2.8
Data preparation	8.3	6.3	6.3	3.0	2.0
Usability in analysis	8.0	6.6	5.4	6.3	5.3
Inexperienced (19)					
Overall rating (min-max)	8.7 (6.7 to 10)	7 (3.7 to 9.7)	6.3 (4.3 to 8.7)	6.3 (4.7 to 8.3)	4.6 (1.3 to 7.3)
Getting started	8.8	7.3	6.9	7.7	5.0
Data preparation	8.3	6.3	6.3	5.0	2.8
Usability in analysis	9.1	7.3	5.6	6.3	4.1

Bax et al, *BMC Med Res Meth*, 2007;7:40.

Comprehensive Meta-Analysis



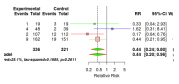
Standard statistical software

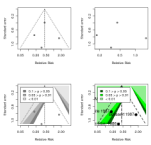
R
http://cran.r-project.org/web/packages/rmeta/rmeta.pdf
http://cran.r-project.org/web/packages/metafor/

STATA
http://www.medepi.net/meta/software/STATA_Metaanalysis_commands_V6_March2004.pdf

SAS
http://www.senns.demon.co.uk/SAS%20Macros/SASMacros.html

WinBUGS (Bayesian)
http://www.openbugs.info/w/





Publishing a meta-analysis

- Consider which journals have an interest in publishing meta-analyses - what are their instructions to authors?
- Does the quantitative reporting of results from meta-analysis reduce the need for qualitative discussion more typical of a systematic review?
- Are there standard protocols for writing up? Yes, MOOSE…

Stroup et al, JAMA, 2000;283(15):2008-12.

MOOSE Checklist

Reporting of background should include
- Problem definition
- Hypothesis statement
- Description of study outcome(s)
- Type of exposure or intervention used
- Type of study designs used
- Study population

Reporting of search strategy should include
- Qualifications of searchers (eg, librarians and investigators)
- Search strategy, including time period included in the synthesis and keywords
- Effort to include all available studies, including contact with authors
- Databases and registries searched
- Search software used, name and version, including special features used (eg, explosion)
- Use of hand searching (eg, reference lists of obtained articles)
- List of citations located and those excluded, including justification
- Method of addressing articles published in languages other than English
- Method of handling abstracts and unpublished studies
- Description of any contact with authors

MOOSE Checklist cont…

Reporting of methods should include
- Description of relevance or appropriateness of studies assembled for assessing the hypothesis to be tested
- Rationale for the selection and coding of data (eg, sound clinical principles or convenience)
- Documentation of how data were classified and coded (eg, multiple raters, blinding, and interrater reliability)
- Assessment of confounding (eg, comparability of cases and controls in studies where appropriate)
- Assessment of study quality, including blinding of quality assessors; stratification or regression on possible predictors of study results
- Assessment of heterogeneity
- Description of statistical methods (eg, complete description of fixed or random effects models, justification of whether the chosen models account for predictors of study results, dose-response models, or cumulative meta-analysis) in sufficient detail to be replicated
- Provision of appropriate tables and graphics

MOOSE Checklist cont…

Reporting of results should include
- Graphic summarizing individual study estimates and overall estimate
- Table giving descriptive information for each study included
- Results of sensitivity testing (eg, subgroup analysis)
- Indication of statistical uncertainty of findings

Reporting of discussion should include
- Quantitative assessment of bias (eg, publication bias)
- Justification for exclusion (eg, exclusion of non–English-language citations)
- Assessment of quality of included studies

Reporting of conclusions should include
- Consideration of alternative explanations for observed results
- Generalization of the conclusions (ie, appropriate for the data presented and within the domain of the literature review)
- Guidelines for future research
- Disclosure of funding source

37

Resources
Introduction to meta-analysis

Borenstein, M, Hedges LV, Higgins JPT, Rothstein HR. *Introduction to Meta-Analysis*, Wiley 2009
By the authors of Comprehensive Meta Analysis.

Cochrane Handbook for Systematic Reviews of Interventions 4.2.6 (Sept 2006) (PDF) **pages 97-166**, or, the latest version available to view online: Higgins JPT, Green S (editors). *Cochrane Handbook for Systematic Reviews of Interventions* Version 5.0.2 [updated September 2009]. The Cochrane Collaboration, 2009. Available from www.cochrane-handbook.org

Stangl DK, Berry DA. *Meta-analysis in Medicine and Health Policy*, New York, NY: Marcel Dekker, 2000. [Large focus on Bayesian approach].

Sutton AJ, Abrams KR, Jones DR, Sheldon TA, Song F. *Methods for Meta-analysis in Medical Research*. Chichester, UK: John Wiley & Sons, 2000. Including:
- Chapter 16 on Meta-analysis of Epidemiological and Observational Studies

Wolf FM. (1986). *Meta-analysis: quantitative methods for research synthesis*. Sage Publications.

38

Resources
Meta-analytic methods

Bax et al (2007) A systematic comparison of software dedicated to meta-analysis of causal studies. *BMC Medical Research Methodology* 2007, 7:40

Chinn, C. (2000). A simple method for converting an odds ratio to effect size for use in meta-analysis. *Statistics in Medicine*, 19:3127{3131)

Duval S, Tweedie S. (2000). Trim and Fill: A Simple Funnel-Plot-Based Method of Testing and Adjusting for Publication Bias in Meta-Analysis. *Biometrics*, 56(2), 455-463.

Greenland S. Interpretation and choice of effect measures in epidemiologic analyses. *Am J Epidemiol* 1987;125: 761–8.

Sterne JA, Egger M. (2001). Funnel plots for detecting bias in meta-analysis: Guidelines on choice of axis. *Journal of Clinical Epidemiology* 54 (2001) 1046–1055.

Sutton AJ, Abrams KR, Jones DR, Sheldon TA, Song F. *Methods for Meta-analysis in Medical Research*. Chichester, UK: John Wiley & Sons, 2000. Including:
- Chapters 3 to 9

39

Resources
Reporting a meta-analysis

Stroup DF, Berlin JA, Morton SC; et al. Meta-analysis of Observational Studies in Epidemiology: A Proposal for Reporting. *JAMA.* 2000;283(15):2008-2012

MOOSE (Meta-Analysis of Observational Studies in Epidemiology). This checklist for reporting observational studies was developed following a workshop convened to address the problem of increasing diversity and variability that exist in reporting meta-analyses of observational studies. (Stroup et al., 2000). Checklist: http://jama.ama-assn.org/cgi/content/full/283/15/2008/TABLEJST00003T1

Cochrane Handbook for Systematic Reviews of Interventions 4.2.6 (Sept 2006) (PDF) **pages 147-150: 8.9 Presenting, illustrating and tabulating results**. Available from www.cochrane-handbook.org

Sutton AJ, Abrams KR, Jones DR, Sheldon TA, Song F. *Methods for Meta-analysis in Medical Research*. Chichester, UK: John Wiley & Sons, 2000. Including:
- Chapter 10 Reporting the Results of Meta-analysis

40

Contact: Geoff.Der@glasgow.ac.uk

41