PBIO 504

Systematic Reviews and Meta-Analysis

Reviews

Literature Review

Critical Review

Conceptual Review

Mapping Review

Meta-Analysis

Methods Review

Overview

Qualitative Evidence Synthesis

Etc.

Grant MJ 2009, assigned reading

Why Reviews are Needed

- Number of publications overwhelming
- Enormous amount of information (in print and electronic) from different populations, countries and in many languages
- Results from studies on the same topic may disagree and differ in their conclusions
- Discuss different types of bias
- Address heterogeneity
- Summarize the current scientific evidence
- Implement interventions and treatments

Systematic Reviews

- Synthesis of scientific knowledge about a certain condition or disease, its diagnosis, treatment, prognosis, etc.
- Overview of different studies addressing the same research question, but with different populations, sample size and power, study design
- Systematic reviews constitute rigorous scientific research that is reproducible

Systematic Reviews Purpose

- Demonstrate feasibility of a treatment or intervention
- Implement new research findings to clinical practice
- Apply for research funding (grants)
- Suggesting a new direction for research
- Establish clinical performance
- Determine cost-effectiveness

Cochrane Database of Systematic Reviews (CDSR)

“Cochrane works collaboratively with contributors around the world to produce authoritative, relevant, and reliable evidence, in the form of Cochrane Reviews.

Cochrane Reviews are systematic reviews of primary research in human health care and health policy, and are internationally recognized as the highest standard in evidence-based health care resources. They investigate the effects of interventions for prevention, treatment, and rehabilitation. They also assess the accuracy of a diagnostic test for a given condition in a specific patient group and setting. They are published online in the Cochrane Library”.

Reference:

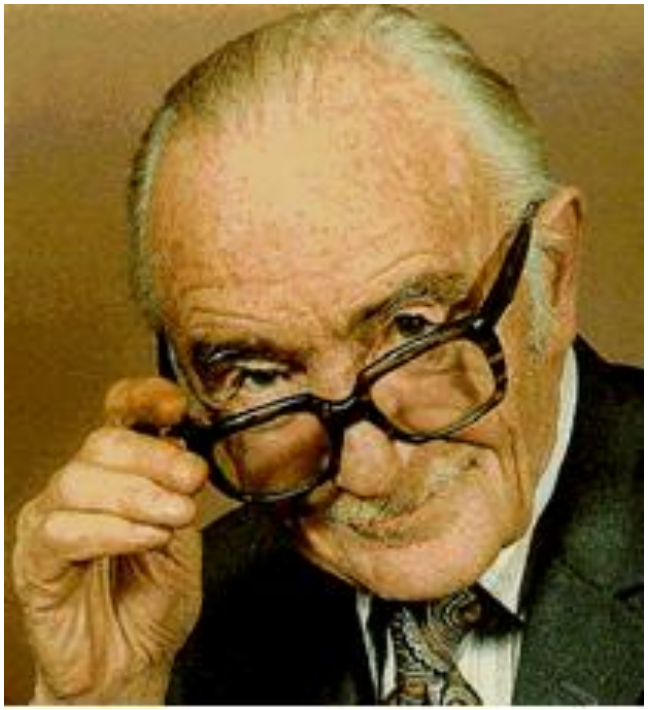
http://www.cochrane.org/what-is-cochrane-evidence

Professor Archibald Leman Cochrane, CBE FRCP FFCM, (1909-1988)

- The Cochrane Library which includes the Cochrane Archive is located at Cardiff University.

His work led to the opening in 1992 of the first Cochrane center in Oxford, UK

and the founding in1993 of the 'The Cochrane Collaboration‘



Picture thanks to Cardiff University Library, Cochrane Archive, University Hospital Llandough

Systematic Review Steps

- Decide on a focused review question
- Formulate a search strategy
- Set inclusion/exclusion criteria
- Protocol writing
- Conduct a thorough literature search
- Assess the studies to be included
- Extract the data
- Summarize the data from the selected studies
- Perform statistical analysis, if appropriate
- Report the findings in context and disseminate results

Note: Keeping accurate documentation is extremely important!

Systematic Review

- Include all relevant evidence to have a clear understanding of effects
- Try to minimize bias
- Describe in detail how the search was performed to find and include all relevant studies
- If studies have been excluded from a review, state the exclusion criteria
- Give an estimate of potential evidence not included and for what reason

Systematic Reviews

RCTs (randomised controlled trials)
of drug, immunization, or behaviour interventions

Observational Studies

Have potential to lead to misleading conclusions
Some reasons for concern include
bias and confounding
different study designs

Weaknesses of Systematic Reviews

- They vary in quality
- Need periodic updating
- Not all studies included (e.g. foreign language, early studies, studies with unclear results)
- Most positive results published in English
- Subjective assessment of included studies
- Publication bias
  - most research never published, only positive results are likely to be published
  - published trials are generally larger and may show an overall greater treatment effect than trials not published

Bias in Systematic Reviews

- Not all research can be easily retrieved by the specific search strategy
- Not all research published in journals is indexed on major databases or it is not always indexed in databases consistently
- Research some times reported in an unusual format
- Lead time to publication usually extensive
- Gaps in publication (from conference presentation)
- Indexing lag (from publication to record in database)
- Coverage of geographic areas may differ by journal

Examples

Authors describe some of their strategies to uncover as many relevant studies as possible:

- Hand searching
- Scanning reference lists
- Personal communication
- Searching web sites

- Trials found by searching beyond major databases

Detective Work

- Is there an existing Systematic Review? If yes, see their search strategy and improve on that
- Be creative with search words and phrases
- Search: Cochrane, Pubmed, Ovid, Embase
- Search Google and Google Scholar
- Search SCOPUS and TRIP
- Web of Science
- Theses and Dissertations
- Look at references and footnotes for cited articles

Find RCTs

- Cochrane Central Register of Controlled Trials (CENTRAL) includes records from Medline, Embase, conference proceedings

- clinicaltrials.gov
- clinicaltrialresults.com

- Critical Care, Trauma and Rehabilitation Trials Group (CCTR)
- Other databases such as Medline, Embase, Cinahl, Psychinfo

Additional Resorces

Campbell Collaboration Resource Centre (systematic reviews of the effects of social interventions)

EPPI-Centre

(systematic reviews of public policy)

PROSPERO

DARE

EQUATOR (Enhancing the QUAlity and Transparency Of health Research) aims to improve health research reporting, see next slide

EQUATOR NETWORK

Guidelines for reporting:

PRISMA systematic reviews

CONSORT Statement randomized controlled trials

STARD diagnostic accuracy studies

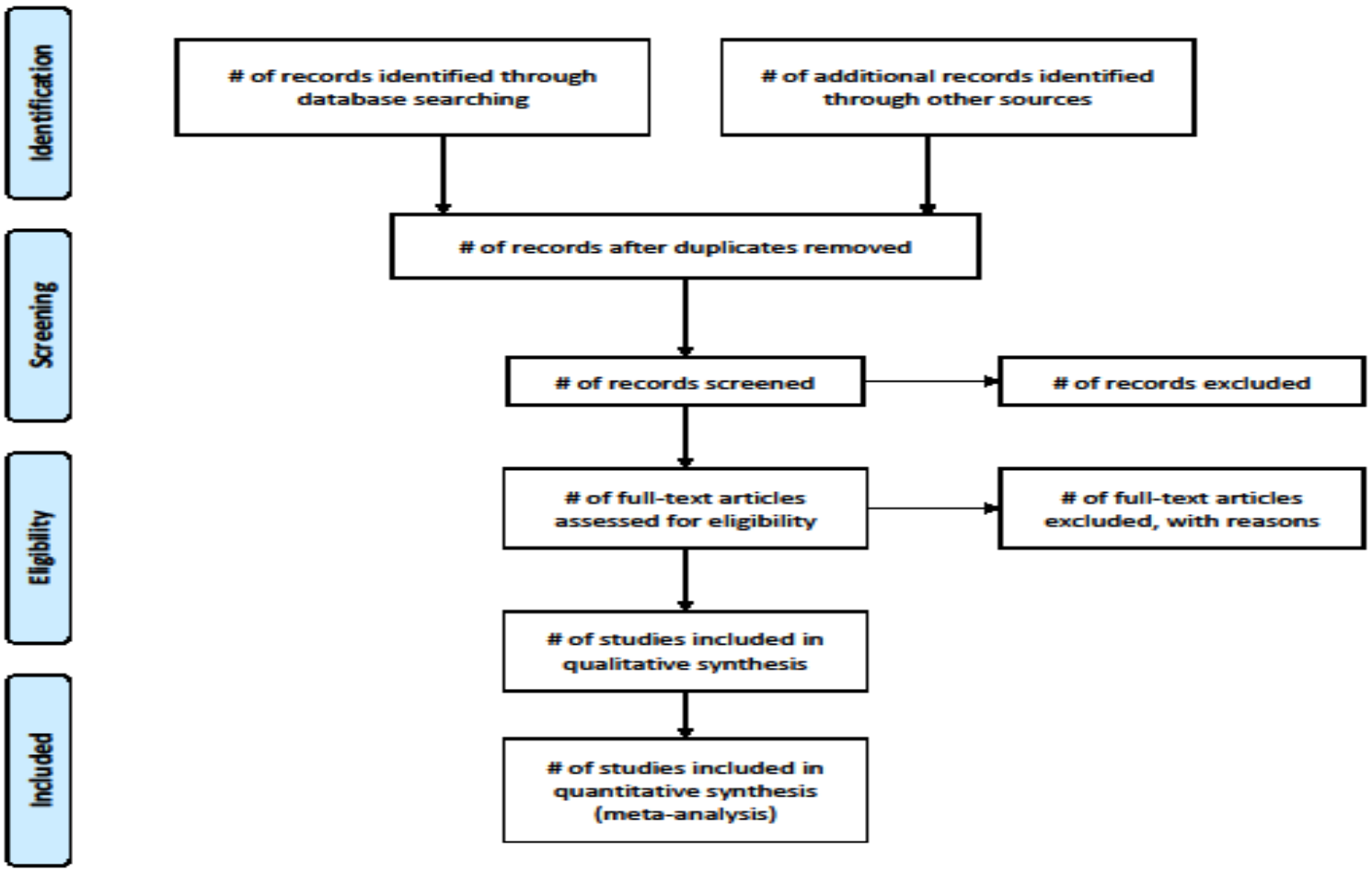
STROBE observational studies in epidemiology

MOOSE meta-analyses of observational studies

BioSharing

BioPortal and Minimum Information for Biological and Biomedical Investigation (MIBBI) portal bioscience reporting guidelines and tools

PRISMA 2009 Flow Diagram



From: Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group (2009). *Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement.* PLoS Med 6(6): e1000097. doi:10.1371/journal.pmed1000097

For more information, visit www.prisma-statement.org.

Meta-Analysis

- Meta-Analysis is the quantitative synthesis component of a systematic review. It is a statistical method usually performed when appropriate to combine study results.
- Focus is on methodology of combining statistical results and examine heterogeneity between studies.
- Examines the effect size and its direction and can ascertain if the effect is consistent across selected studies.
- Conflicting results can be formally assessed, and reasons for different results can be explored and quantified.
- If appropriate to pool the data, the power increases and precision improves, and the summary effect is calculated.

Meta-Analysis

- When combining the results of individual studies in a meta-analysis, random error is reduced, and there is an increased chance of detecting a statistically significant effect (based on increased power).

- Meta-analysis can also detect patterns across study results and analyze the observed differences.

- However, because studies included in a meta-analysis can be heterogeneous in their designs, quality, and populations, it may not always be valid to analyze with this method as it may lead to erroneous conclusions. This constitutes a major disadvantage and it must be extensively evaluated.

Meta-Analysis

- In a perfect situation we would analyze the raw data from each eligible study, but this is a rare event because the raw data may no longer be available, or those in charge of the data are not convinced easily to share.
- This is called pooled meta-analysis, when all the raw data is combined and analyzed.
- Main issue for feasibility is the proper use of the raw data, including:
  - proprietary rights, patient consent and confidentiality, authorship,
  - use of compatible technology and software

Combining Results of Individual Studies

- This is done in two steps
- First calculate summary statistics for each study.
- Second, combine results to calculate an overall summary estimate (weighted average of the individual study estimates). The weight assigned to a study influences the overall estimate.
- The assigned weight is usually in inverse proportion to the study variance.

Effect measures for continuous outcomes

- **Difference in means**, used when the outcome in all studies is measured on the same scale. The summary measure is called *weighted mean difference* (WMD) and it is the *weighted average of these differences in means.*
- **The standardized mean difference**, used when the outcome is not uniformly measured across studies. This involves standardization of results from each study before combining to calculate summary measure.

Standardization is the size of the effect (diff in means) in each study relative to the variability observed in that particular study.

Note: there are assumptions for this method and the overall effect is hard to interpret being expressed in units of stdev.

Summarizing effects across studies depends on the type of outcome var:

- continuous
- dichotomous (or binary)
- ordinal data (including measurement scales)
- counts and rates (counting number of events for each participant)
- time-to-event (the time until an event occurs, but not all participants experience the event, i.e. censoring)

The summary statistic may be a risk ratio for a dichotomous outcome variable or a difference between means if the outcome is continuous.

Summarizing effects across studies

- The summary (pooled) effect estimate is the weighted average of the individual estimated effects. This is calculated as the sum of all the individual estimates multiplied by their assigned weight divided by the sum of all the assigned weights.

- The standard error of the summary (pooled) effect is used to calculate its confidence interval and p-value

Statistical Models for Meta-Analysis

- **Fixed-effect models** weight the contribution of each study proportional to the amount of information observed in the study. Only variability in results within studies is considered and no variability btw studies
- **Random-effects models** consider between-study variability in results by weighting studies using a combination of their own variance and the between-study variance.
- If between-study variability is low, the within-study variance will dominate and the random-effects weighting will be close to the fixed-effect weighting.
- *It is common in practice to run both models and compare results.*

Meta-Analysis Methods for Dichotomous Outcomes

- **For Fixed-effect Models:**

  Mantel-Haenszel, Peto, and Inverse Variance

- **Random-effects Models:**

  DerSimonian and Laird Method

Note: Beside Peto Method, which can summarize only Odds Ratios, all of the methods mentioned here are able to summarize OR, RR, and Risk Differences

Sensitivity Analyses

- Assess robustness of the meta-analysis by performing again the meta-analysis and including only certain studies (exclude those with small sample size for example) and see if the results are similar.
- Researchers often prefer to perform a sensitivity analysis by applying the meta-analysis to subsets of studies based on high-quality versus low-quality studies, randomized versus non-randomized studies, early studies versus late studies, etc.
- Other methods include “leave one study out”

Assess Heterogeneity

- By visually examining forest plots to see if there is little or no overlap between confidence intervals.
- More formally a $\chi 2$ (chi-squared) test p-value provides evidence of heterogeneity (i.e. variation in effect estimates beyond chance).
- Interpret with caution as it has low power for small number of studies included and/or small sample size. A statistically significant result may indicate heterogeneity, but a non-significant result does not necessarily imply non-heterogeneity. To determine statistical significance, alpha of 0.10 is often used.
- If high powered, the test may detect small amounts of heterogeneity that may not be clinically relevant

Test for Heterogeneity

Heterogeneity statistic

- Q statistic

$$Q = \sum_{i=1}^{k} W_i \left(Y_i - \theta_{pooled} \right)^2$$

θ_{pooled} = mean RR
W_i = study-specific weight
Y_i = study-specific RR
k = # studies

Q follows a chi-square distribution with (k-1) degrees of freedom.
The null hypothesis: all the studies are homogeneous (similar OR or RR)
If p-value<0.05 (or 0.10) reject the null hypothesis and conclude that the studies included are heterogeneous (or at least one dissimilar).

Examining Heterogeneity

Sources of Between Study Heterogeneity:

- Different study designs
- Different length of follow-up
- Different distributions of covariates
- Different statistical methods/models used
- Different sources of bias

If Test for Heterogeneity Significant

Decision making:

- Do not include a meta-analysis in the systematic review
- Exclude studies
- Change the effect measure
- Separate studies in homogeneous groups and perform separate meta-analyses
- Meta-Regression

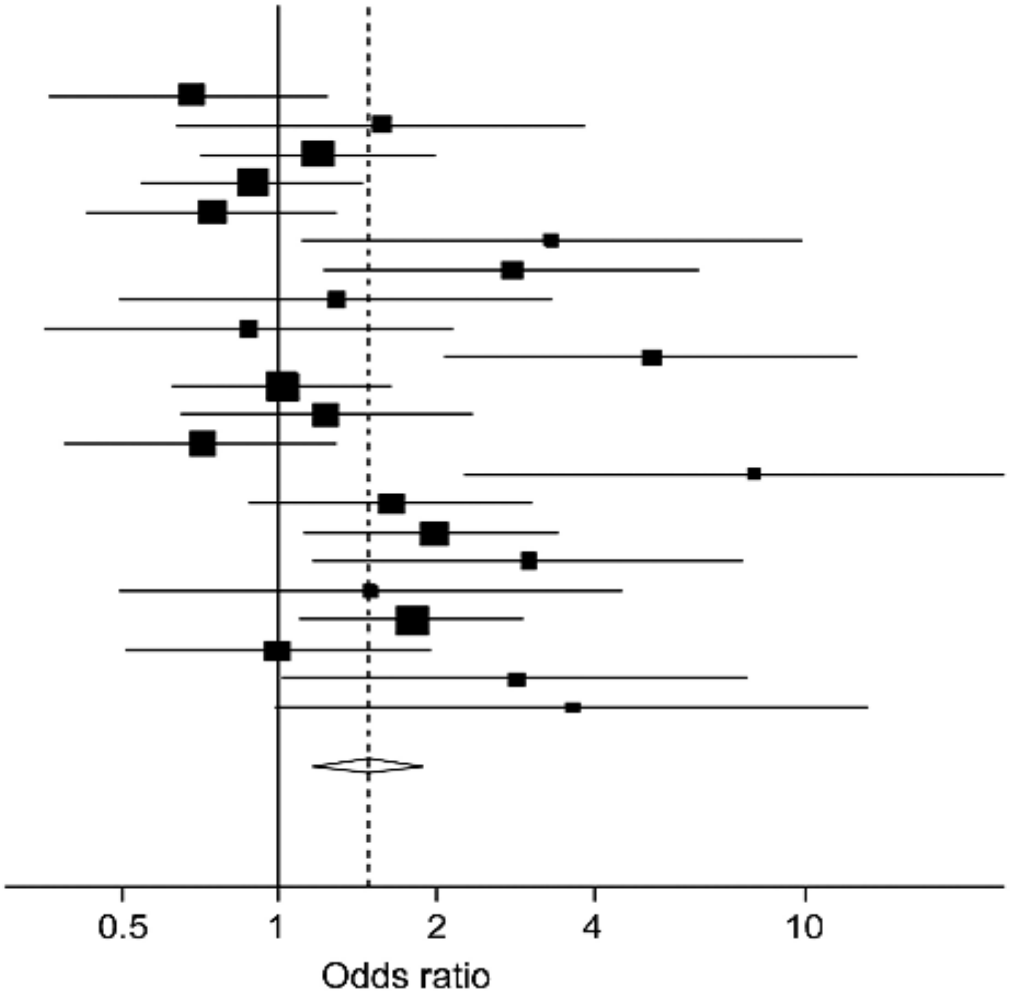
Creating a Forest Plot for the Meta-Analysis Results

- Arrange the studies you reviewed in chronological order, starting at the top with the oldest study, and list each study's estimated relative risk (**or Odds Ratio**) and its 95% CI.
- Draw a vertical line at RR=1.0
- Draw a dashed line to show the summary RR.
- Plot each individual RR and its 95% CI, as a dot or square and a horizontal line for the CI. If possible, make the size of the square proportional to the sample size of the study.
- Plot the calculated summary RR and 95% CI.

Forest Plot Example

STUDY	OR	95% CI
Rosenthal et al, 1998	0.68	0.4-1.3
Hayes et al, 1998	1.56	0.6-3.8
Josefsson et al, 1998	1.19	0.7-2.0
Hildesheim et al, 1998[a]	0.89	0.5-1.4
Sonoda et al, 1999	0.74	0.4-1.3
Zehbe et al, 1999[b]	3.32	1.1-9.9
Zehbe et al, 1999[c]	2.79	1.2-6.4
Giannoudis et al, 1999	1.29	0.5-3.3
Dybikowska et al, 2000	0.88	0.4-2.1
Dokianakis et al, 2000	5.07	2.1-12.0
Madeleine et al, 2000	1.02	0.6-1.6
van Duin et al, 2000	1.23	0.6-2.3
Tenti et al, 2000	0.71	0.4-1.3
Makni et al, 2000	8.00	2.2-28.5
Pegoraro et al, 2000	1.64	0.9-3.1
Zehbe et al, 2001[d]	1.97	1.1-3.4
Zehbe et al, 2001[e]	2.99	1.2-7.7
Gustafsson et al, 2001	1.50	0.5-4.5
Pegoraro et al, 2002	1.79	1.1-2.9
Bhattacharya et al, 2002	1.00	0.5-1.9
Arbel-Alon et al, 2002	2.83	1.0-7.8
Nagpal et al, 2002	3.64	1.0-13.3
ALL STUDIES (Random Effects)	1.48	1.2-1.9

Test for heterogeneity: $\chi^2(21df)=52$; p=0.0002



p53 codon 72 polymorphism and cervical cancer

Koushik et al, 2004

Interpretation

- For most studies observed odds ratios >1, and many were statistically significant based on CI.
- No obvious trend over time.
- The summary statistic shows OR=1.5 and 95% CI that excludes 1.0, suggesting that the gene mutation confers increased cancer risk.
- The Q test is highly significant (very small p-value) suggesting that the studies included are heterogeneous and it is not appropriate to summarize with an overall OR.

Meta-analysis: Strengths and Weaknesses

Strengths

- Quantitative summary of the evidence
- Greater power combining than relying on single studies
- Can assess heterogeneity
- Less subjective than a systematic review

Limitations

- Heterogeneity across studies can limit the conclusions
- Errors in original work cannot be checked
- Limited by quality of the studies included
- Influenced by **publication bias** (studies with non-significant results tend not to be published)