An Introduction to Meta-Analysis

Susan L. Norris, MD, MPH, MS
Associate Professor
Oregon Health & Science University
Portland, OR
norris@ohsu.edu

Biography, SL Norris

• MD, MS, University of Alberta
• MPH, University of Washington
• Board Certified: general surgery (Canada), family medicine (US)
• CDC: directed systematic review group with focus on diabetes, 1999-2004
• Oregon Health & Science University (current)
  – clinical practice guideline development
  – systematic review methodology (nonrandomized studies, sources of bias in reviews)
  – effects of physician-industry relationships on practice guidelines and systematic reviews

Disclosures

• Financial: none
• Intellectual
  – Member, GRADE Working Group
  – Methods work: nonrandomized studies, sources of bias in reviews
  – Impact of physician-industry relationships on primary research and evidence synthesis
• Professional
  – Investigator, Evidence-based Practice Center (AHRQ)
  – Funders: CDC, NIH, AHRQ, American College of Chest Physicians, American Urological Association
Learning Objectives

1. To understand the difference between a meta-analysis and a systematic review
2. To understand when a meta-analysis might be useful and when it may be misleading
3. To define heterogeneity and understand how it affects a meta-analysis
4. To understand the difference between a fixed effects and a random effects model
5. To be able to interpret a forest plot

Outline

1. Introduction
2. Terminology
3. Basic considerations
4. Models for combining data
5. Assessing and exploring heterogeneity
6. Publication bias
7. Reading and interpreting meta-analyses
8. Resources

Part 1. Introduction
Example of a Meta-analysis

Meta-analysis: Enoxaparin versus Unfractionated Heparin

Working Example of a Systematic Review and Meta-analysis
Proliferation of Meta-Analyses

• 1976 - 6 papers on meta-analysis
• 1991 - approximately 600 meta-analyses existed in the social and behavioral sciences
• 1993 - 678 meta-analyses were found in the medical literature alone
• 2005 - over 2000 reviews in the Cochrane library
• 2011 - over 4500 reviews in the Cochrane library

Why the Proliferation of Meta-analyses?

• We need more scholarly effort concentrated on the problem of finding knowledge that lies untapped in completed research studies
• We need methods for the orderly summarization of studies so knowledge can be extracted from the myriad individual researches
• This endeavor deserves higher priority now than adding a new experiment or survey to the pile.

(Glass, 1976)

Part 2.
Terminology
**Typology**

Integrative publications

- Systematic review
- Narrative review
- Nonsystematic review
- Review of comparative effectiveness
- Qualitative synthesis
- Quantitative synthesis
- Meta-analysis
- Practice guidelines
- Economic evaluation
- Decision analysis

**Systematic Review**

A concise summary of the best available evidence that addresses a sharply defined clinical question (Mulrow 1998)

- Qualitative synthesis = narrative summary
- Quantitative synthesis = meta-analysis
Meta-Analysis

IOM 2011, *Finding What Works in Health Care*
“the statistical combination of results from multiple individual studies”
“MA is a broad term that encompasses a wide variety of methodological approaches whose goal is to quantitatively synthesize and summarize data across a set of studies”

Other terms: quantitative synthesis, pooling, pooled analysis

Steps in a Systematic Review

- Develop the review question
- Develop inclusion/exclusion criteria
- Search for literature
- Quality assess individual studies
- Data abstraction and analysis
- Synthesis of findings
- Grading the strength of evidence

Synthesis in Systematic Reviews

Involves 4 questions:
1. What is the direction of effect?
2. What is the effect size?
3. Is the effect consistent across studies?
4. What is the strength of evidence for the effect?

Meta-analysis provides statistical methods for addressing Q 1-3
Narrative synthesis uses subjective methods to examine Q 1-4
Part 3.
Basic Considerations

Why Perform a Meta-analysis?

• To increase power
  – ↑ chance of detecting a real effect if it exists
• To improve precision
• To answer questions not posed by individual studies
  – Eg, examine effects across different populations, interventions, settings, outcomes
• To explore heterogeneity

Does it make sense to perform a meta-analysis?

“... given that the studies differ in various ways and the analysis amounts to combining apples and oranges?”

“... combining apples and oranges makes sense if your goal is to produce a fruit salad”

Does it make sense to perform a meta-analysis?

1. Does it make sense to combine studies?
2. What studies should be included?

Consider these questions in the context of the goal of the meta-analysis:
- Estimate a common effect across similar studies
- Assess dispersion (heterogeneity)
- Both

Does it make sense to combine data across studies?

- Conceptual considerations
  - Clinical diversity
  - Methodological diversity
- Statistical considerations
  - Statistical heterogeneity

When Not to Combine Studies in a Meta-analysis

- Studies are too diverse with respect to PICO to provide a meaningful answer
  - apples and oranges
- Included studies are biased
- In the presence of significant publication and/or reporting biases: summary measure is not valid
Steps in a Meta-analysis

1. Calculate a summary statistic for each study
2. Calculate a summary (pooled) intervention effect:
   - A weighted average of the intervention effects estimated in individual studies
3. Estimate heterogeneity
4. Assess potential for bias

Choice of Summary Statistic at the Individual Study Level

Depends on:

- Which one makes most sense
- Which is most consistently reported
- Which is mathematically appealing

Choice of Summary Statistic at the Individual Study Level

- Dichotomous (binary) data
  - Relative effect: risk ratio, odds ratio
  - Absolute effect: risk difference
  - NNT: not applicable to meta-analysis

- Continuous data
  - Difference in means: all outcomes are on the same scale
  - Standardized mean difference
    - When outcome is measured on a different scale
    - Can be difficult to interpret
  - Time-to-event (survival outcomes): hazard ratios
Part 4.
Models for Combining Data

Simple Average
\[
\frac{(-6.2) + (-7.7) + (-0.1)}{3} = -4.7 \text{ mmHg}
\]

<table>
<thead>
<tr>
<th>Study</th>
<th>Sample Size</th>
<th>$\Delta$ mmHg</th>
<th>95% Confidence Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>ANBP</td>
<td>554</td>
<td>-6.2</td>
<td>-6.9 to -5.5</td>
</tr>
<tr>
<td>EWPHE</td>
<td>304</td>
<td>-7.7</td>
<td>-10.2 to -5.2</td>
</tr>
<tr>
<td>Kuramoto</td>
<td>39</td>
<td>-0.1</td>
<td>-6.5 to 6.3</td>
</tr>
</tbody>
</table>

Weighted Average
\[
\frac{(554 \times -6.2) + (304 \times -7.7) + (39 \times -0.1)}{554 + 304 + 39} = -6.4 \text{ mmHg}
\]

<table>
<thead>
<tr>
<th>Study</th>
<th>Sample Size</th>
<th>$\Delta$ mmHg</th>
<th>95% Confidence Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>ANBP</td>
<td>554</td>
<td>-6.2</td>
<td>-6.9 to -5.5</td>
</tr>
<tr>
<td>EWPHE</td>
<td>304</td>
<td>-7.7</td>
<td>-10.2 to -5.2</td>
</tr>
<tr>
<td>Kuramoto</td>
<td>39</td>
<td>-0.1</td>
<td>-6.5 to 6.3</td>
</tr>
</tbody>
</table>
General Formula - Weighted Average Effect Size ($d_+$)

$$d_+ = \frac{\sum_{i=1}^{k} w_i d_i}{\sum_{i=1}^{k} w_i}$$

where:
- $d_i$ = effect size of the $i_{th}$ study
- $w_i$ = weight of the $i_{th}$ study
- $k$ = number of studies

Generic Inverse-Variance Weighted Average

Weighted average over $i$ studies

With weight defined as inverse of the SE squared

Two Meta-analytic Models

**Fixed effect meta-analysis**
- Assumes that the true effect size each study is trying to estimate is the same (i.e. fixed) across all the studies.
- There will be differences in the estimates each study arrives at, but these are just because of chance.

**Random effects meta-analysis**
- Incorporates an estimate of the between study variation (heterogeneity) by assuming that there is more than one true effect
FIXED EFFECTS MODEL

POOLED RESULT
SINGLE TRUE
TREATMENT EFFECT
RESULTS OF
MULTIPLE CLINICAL
TRIALS RANDOMLY
DISTRIBUTED
AROUND THE TRUE
TREATMENT EFFECT

TREATMENT EFFECTS (RD, OR, RR)

SINGLE
TRUE
TREATMENT EFFECT

POOLED RESULT
SINGLE ESTIMATED
TREATMENT EFFECT

RESULTS OF
MULTIPLE CLINICAL
TRIALS RANDOMLY
DISTRIBUTED
AROUND THE TRUE
TREATMENT EFFECT

TREATMENT EFFECTS (RD, OR, RR)

RANDOM EFFECTS MODEL

MULTIPLE TRUE
TREATMENT EFFECTS
Identification of treatment effects

RANDOM EFFECTS MODEL

POOLED RESULT
SINGLE ESTIMATED
TREATMENT EFFECT

RESULTS OF
MULTIPLE CLINICAL
TRIALS RANDOMLY
DISTRIBUTED
AROUND EACH OF THE
TRUE TREATMENT EFFECT

TREATMENT EFFECTS (RD, OR, RR)
Fixed Effect and Random Effects Models

**Fixed Effect Weight**

\[ W_i = \frac{1}{V_i} \]

**Random Effects Weight**

\[ W_i^* = \frac{1}{V_i + V^*} \]

where:  
\( V_i \) = within study variance  
\( V^* \) = between study variance

---

**Which Model is Appropriate?**

- **Fixed effect:**
  - All studies in the analysis are identical
  - Goal is to compute the common effect size ES for the identified population and not to generalize to other populations
  - relatively rare
- **Random effects:**
  - When expect variation in effects across studies are not all the same
  - Want to generalize to a range of scenarios
- **Choice between FE and RE should not be based on a statistical test, but rather an understanding of whether or not all studies share a common ES**
- **Generally the random effects model is more plausible**

---

**Fixed effect**

**Fixed effects**

**Random effects**

**Random effects**
Part 5.
Assessing and Exploring Heterogeneity

Why do Empirical Results Vary?

- Different study populations
- Different treatments or protocols
- Quality of technical design or execution
- Random variation
Heterogeneity

**Homogeneity**: degree to which studies are sufficiently similar that a single average effect would be meaningful

**Heterogeneity**: any kind of variability among studies in a systematic review

**Clinical diversity**: variability in participants, interventions, outcomes

**Methodological diversity**: variation in study design and risk of bias

**Statistical heterogeneity**: a consequence of clinical or methodological diversity

Only pool studies that will provide a meaningful summary

Presenting, Assessing, and Exploring Heterogeneity

- Present
  - Forest plot
- Assess
  - Cochran’s Q (Chi square test)
  - I^2 (Higgins)
- Explore
  - Subgroup analysis
  - Meta-regression
  - Sensitivity analysis

Example of a Forest Plot Showing Heterogeneity

Association between the metabolic syndrome and cardiovascular disease

*Ford ES. Diabetes Care, 2005. 28:1769-1778*
Heterogeneity has Two Components

- Variation in effect size includes both true heterogeneity and random error
- Want to isolate the true variance:
  - Compare the observed dispersion with the amount expected if all studies shared a common effect size
  - The excess is assumed to reflect real differences among studies
  - This portion of the variance is then used to create several measures of heterogeneity

Chi-Square Homogeneity Test (Mantel-Haenszel)

\[ Q = \chi^2_{(k-1)df} = \sum_{i=1}^{k} \left[ w_i (d_i - d^+)^2 \right] \]

NOTE: \( d = \ln(OR_i) \) \( d^+ = \ln(ORMH) \) \( w_i = 1/variance(OR_i) \)

Variance (OR_i) = 1/ai + 1/bi + 1/ci + 1/di

Heterogeneity: Higgins’ \( I^2 \)

\( I^2 \) quantifies the proportion of the observed dispersion that is real rather than spurious (due to random error)

\( Q \) is Cochran’s Q chi-square statistic

\( I^2 \) is expressed as a percentage [0-100%]

Negative values set to 0

25%, 50%, and 75% represent low, moderate, and high heterogeneity

\( I^2 \) can be directly compared between meta-analyses with different numbers of studies and different types of outcome data
Strategies for Dealing with Heterogeneity

- Check that the data are correct
- Do not do a meta-analysis
- Ignore heterogeneity: fixed effects model
- Use a random effects model
- Explore heterogeneity: subgroup analysis, meta-regression
- Change the effects measure
- Exclude studies

Dealing with Heterogeneity

<table>
<thead>
<tr>
<th>Ignore</th>
<th>Estimate (insensitive)</th>
<th>Incorporate</th>
<th>Explain</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fixed effects model</td>
<td>Random effects model</td>
<td>Meta-analyses</td>
<td>Meta-regression (control risk, covariates)</td>
</tr>
</tbody>
</table>

Subgroup Analysis

Explore heterogeneity with subgroup analyses

- Can be used to answer specific questions about patient groups
- Can be used to investigate heterogeneous results
- Data are often insufficient in the primary studies
- Subgroup analyses are observational: no randomized comparisons
  - Risk of confounding by bias from other study-level characteristics
- Risk of false (+) if excessive number of subgroup
Meta-Regression

- An observational study: confounding is a possibility
- Extends random effects meta-analysis to estimate the extent to which study-level covariates explain heterogeneity in the results
- Use even if the test for heterogeneity is non-significant
- Use study-level covariates
  - Avoid ecologic fallacy
- Specify the study covariates before you begin
  - Avoid data dredging

Sensitivity analyses

- Vary inclusion criteria
- Change your assumptions
- Fixed versus random effects model
- Different metrics
- Cumulative meta-analysis

Part 6.
Publication Bias
Publication Bias

- Definition: The publication of studies based on the magnitude and direction of the findings
- If missing studies are systematically different from identified studies, then bias is introduced into the systematic review
- Published studies are more likely to report positive results
- Literature-based meta-analyses tend to bias results towards statistically significance and appear convincing due to increased precision

Types of Publication Bias

- English language bias
- Availability bias
- Cost bias
- Familiarity bias
- Duplicate publication bias
- Citation bias
- Time lag bias

Impact of publication bias

Combination chemotherapy vs. monotherapy in ovarian cancer

Meta-analysis based on published and registered trials

Survival ratio (95% confidence interval)

Published trials (p=0.004)
Registered trials (p=0.17)
Publication Bias

1. How minimize PB?
2. How assess if PB has any impact on the observed effect?
3. How assess if PB might be entirely responsible for the observed effect?
4. How much effect does PB have on the observed effect size?

Assumptions of Models Examining Publication Bias

- Large studies are likely to be published regardless of statistical significance
- Small studies are at greatest risk for being lost: only the largest effects are likely to be published
- Effect of these assumptions:
  - Expect bias to increase in sample size does down
- Various methods for assessing PB are based on this model
Funnel Plots

- One mechanisms for displaying the relationship between sample size and effect size
- Plots effect size vs SS or a measure of variance: large studies cluster around mean effect; smaller studies tend to be spread across a broad range of values; creates a funnel
- Now usually plotted as effect size (X-axis) versus 1/SE (Y-axis)
- In the absence of publication bias, studies are distributed symmetrically about the mean effect size, as sampling error is random
- In the presence of publication bias, there will be missing studies with nonsignificant effect size and smaller sample size or large SE
- Interpretation of funnel plot is subjective, so various tests have been devised
Possible Causes of Funnel Plot Asymmetry

- Publication bias
  - Funnel plot assumes that if effect size is larger in smaller studies, it is due to publication bias
- True heterogeneity
  - There may be real, small-study effects
- Choice of precision measure
- Choice of effect measure
- Trial quality
- Chance

Part 7.
Reading and Interpreting Meta-Analyses
Working Example of a Systematic Review and Meta-analysis

Mortality with Tiotropium vs Placebo (Singh et al. BMJ 2011)

Mortality with Tiotropium vs Placebo: Subgroups (Singh et al. BMJ 2011)
Mortality with Tiotropium vs Placebo: Sensitivity Analysis (Singh et al. BMJ 2011)

Part 8.
Resources on Meta-Analyses

Resources for Performing Meta-Analysis

- Organizations
  - Cochrane Collaboration: www.cochrane.org
  - NICE, SIGN
Books on Meta-Analyses


Books on Meta-Analyses, Cont’d

Finding What Works in Health Care; Standards for Systematic Reviews. Institutes of Medicine of the National Academies, 2011.


Meta-analysis Software

• Customized software
  Commercial packages
  – Meta-Win [www.metawinssoft.com]
  Freely available
  – RevMan (www.cochrane.org)
  – Meta-Analyst
  – MIX (Meta-analysis with Interactive Explanations)

• Statistical packages with meta-analysis capability via custom code or macros
  – Stata, SAS, SPSS, R
Contemporary guidelines for reporting meta-analyses: PRISMA 2009

Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement

David Moher1,*, Alejandro Liberati2,*, Jo Ann Tetzlaff3,*, Jennifer Altman4, The PRISMA Group

1Ontario Health Technology Assessment Network, University of Ottawa, Ottawa, Ontario, Canada 2Prévention des Maladies Chroniques, Université de Sherbrooke, Sherbrooke, Quebec, Canada 3Clinical Epidemiology Unit, University of Oxford, Oxford, United Kingdom 4Department of Methodology, Institute of Education, University of London, London, United Kingdom

Adequate reporting of a systematic review or meta-analysis is essential for the reader to judge the validity and generalizability of the results. This statement was developed to provide guidance in this area. Development of the PRISMA statement was informed by previous work on reporting guidelines, including most recently the CONSORT statement. The PRISMA statement provides a list of items that should be considered when reporting a systematic review or meta-analysis. The aim is to ensure better reporting and hence a more accurate and confident interpretation of the results.

PRISMA 2009 Checklist

<table>
<thead>
<tr>
<th>Item</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Title page includes name(s) of authors, full address and contact information</td>
</tr>
<tr>
<td>1.2</td>
<td>Abstract starts with a statement of the purpose of the study and ends with a conclusion</td>
</tr>
<tr>
<td>1.3</td>
<td>Introduction includes background and the purpose of the review, a comprehensive literature search, inclusion and exclusion criteria, and the methods of data collection and analysis</td>
</tr>
<tr>
<td>1.4</td>
<td>Methods include a detailed description of the study design, data sources, and data collection and analysis methods</td>
</tr>
<tr>
<td>1.5</td>
<td>Results include a comprehensive, structured summary of the results, including their implications for future research and practice</td>
</tr>
<tr>
<td>1.6</td>
<td>Discussion includes a comprehensive, structured summary of the discussion, including implications for future research and practice</td>
</tr>
<tr>
<td>1.7</td>
<td>References include a comprehensive, structured summary of the references, including their implications for future research and practice</td>
</tr>
</tbody>
</table>

For more information, see the PRISMA Statement website.

*Corresponding author. E-mail address: dmoher@uottawa.ca
Suggested Readings


Conclusions

- Meta-analysis is the statistical combination of results from two or more studies
- Advantages: increase in power, improved precision, ability to answer questions not posed by individual studies, examine conflicting evidence
- Caveats
  - Relies on an underlying systematic review
  - Limited by available data
  - Beware of spurious precision
- Most MA involve a weighted average of effects estimates from different studies
- Heterogeneity across studies must be considered and explained
- Use sensitivity analyses to examine robustness of findings

References & Resources

References & Resources


