Systematic Reviews and Meta-analyses

Introduction

- A systematic review (also called an overview) attempts to summarize the scientific evidence related to treatment, causation, diagnosis, or prognosis of a specific disease. A systematic review does not generate any new data – it reviews and summarizes already-existing studies.

- Although it appears that conducting a systematic review is easy, it requires a good deal of effort and care to do it well.

Introduction

- There are six basic steps to a systematic review:
  (1) Define a focused clinical question
  (2) Conduct a thorough literature search
  (3) Apply inclusion/exclusion criteria to the identified studies
  (4) Abstract/summarize the data from the eligible studies
  (5) Perform a statistical analysis (meta-analysis), if appropriate
  (6) Disseminate the results

1. Define a Focused Question

- If the question is too broad, it may not be useful when applied to a particular patient. For example, whether chemotherapy is effective in cancer is too broad a question (the number of studies addressing this question could exceed 10,000).

- If the question is too narrow, there may not be enough evidence (studies) that address it. For example, the following question is too narrow: Is a particular asthma therapy effective in Caucasian females over the age of 65 years in Central PA?

2. Conduct a Thorough Literature Search

- Many sources for studies (throughout the world) should be explored:
  - Bibliographic databases (Medline, Embase, etc.)
  - Publically available clinical trials databases (www.clinicaltrials.gov)
  - Conference proceedings
  - Theses/dissertations
  - Databanks of pharmaceutical firms
  - Personal contacts
  - Unpublished reports

- Beware of publication bias. Studies in which the intervention is not found to be effective sometimes are not published. In other words, a systematic review based only on published studies may be biased towards an overall positive effect.

2. Conduct a Thorough Literature Search

- Sometimes, “publication bias” is referred to as the “file-drawer problem” because studies with negative results tend to be filed in a drawer and not submitted for publication. Editors like to publish “positive” studies in their journals.

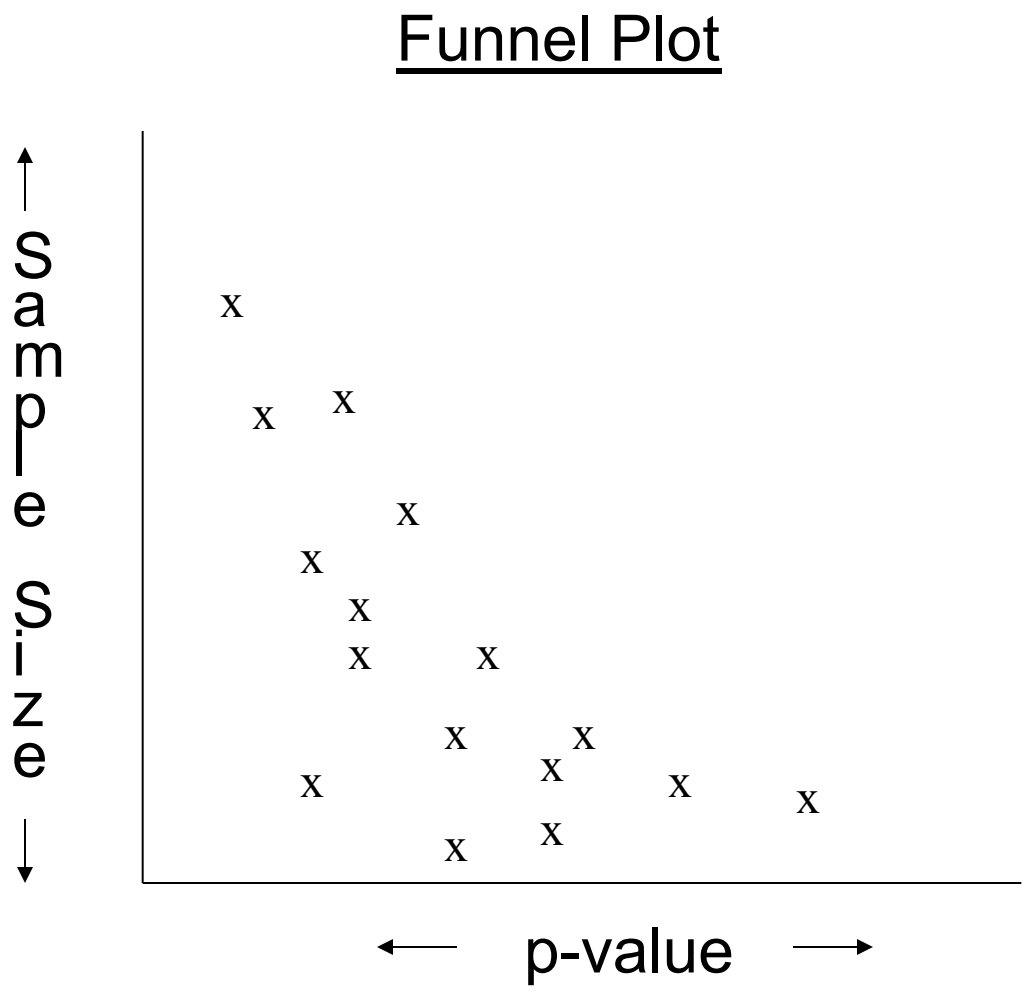
- Suppose there are some relevant studies with small sample sizes. If nearly all of them have a positive finding ($p < 0.05$), then this may provide evidence of a “publication bias” because of the following reason. It is more difficult to show positive results with small sample sizes. Thus, there should be some negative results ($p > 0.05$) among the small studies.

2. Conduct a Thorough Literature Search

- A “funnel plot” can be constructed to investigate the latter issue. Plot sample size (vertical axis) versus p-value or magnitude of effect (horizontal axis).

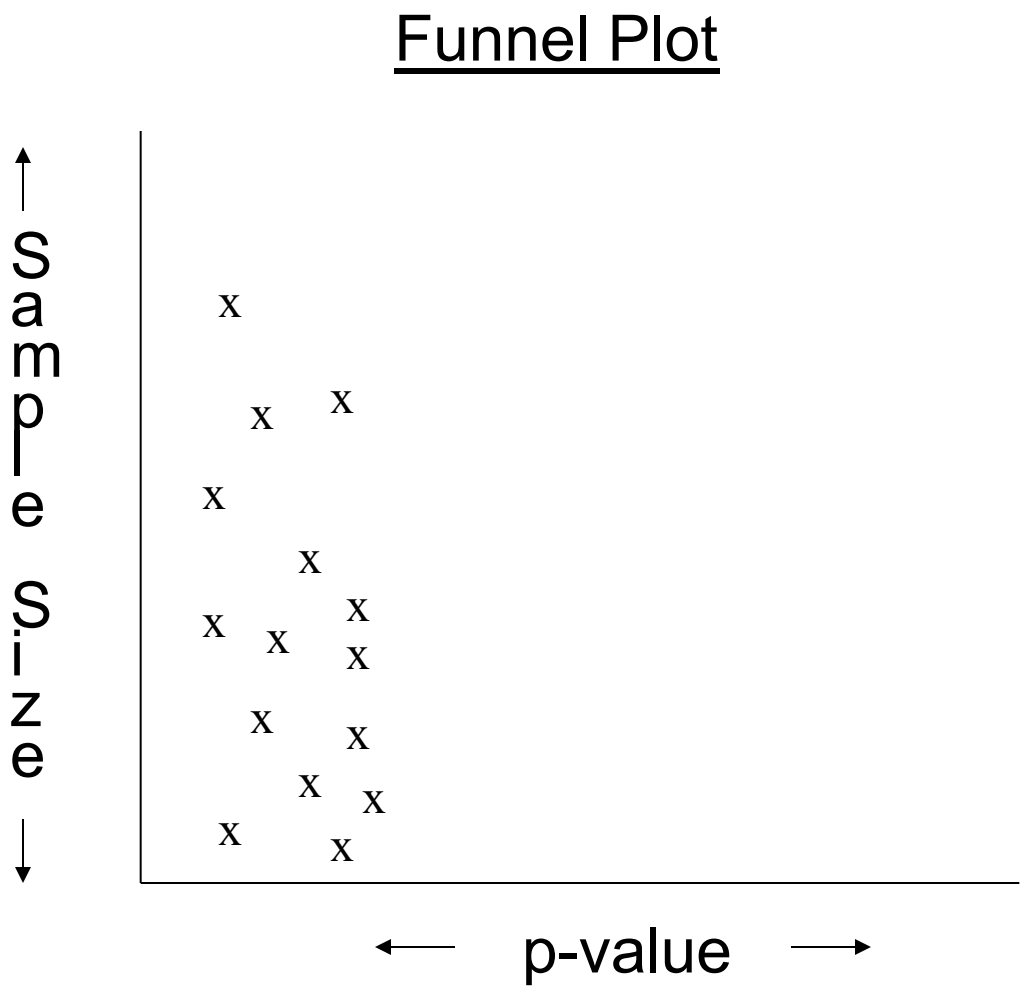
- Ideally, p-values for studies with small sample sizes should have a wide range, whereas studies with large sample sizes should have a narrow range. In such a situation, the funnel plot should take on the appearance of a funnel (or fan).

2. Conduct a Thorough Literature Search



The p-values for some of the small studies are relatively large, yielding a “funnel” shape for the scatterplot.

2. Conduct a Thorough Literature Search



None of the p-values for the small studies are large, yielding a "band" shape for the scatterplot and suspicion of publication bias.

3. Apply Inclusion/Exclusion Criteria

- Eligibility criteria for studies need to be established prior to the analysis. The researcher should base the inclusion/exclusion criteria on the design aspects of the trials, the patient populations, treatment modalities, etc. that are congruent with the objectives of the overview.

- Although subjective, some researchers grade the selected studies according to quality and may weight the studies accordingly in the analysis.

4. Abstract/Summarize the Data

- In most circumstances, the researcher easily can gather the relevant descriptive statistics (e.g., means, standard errors, sample sizes) from the reports on the eligible studies.

- Sometimes, older reports (say, prior to 1980) do not include variability estimates (e.g., standard errors). If possible, the researcher should attempt to contact the authors directly in such situations. This may not be successful, however, because the authors may no longer have the data.

- Ideally, the statistical analysis for a systematic review will be based on the raw data from each eligible study. This rarely occurs, however, because most authors are not willing to share their raw data or the raw data no longer are available.

5. Meta-Analysis

- The obvious advantage for performing a meta-analysis is that a large amount of data, pooled across multiple studies, can provide increased precision in addressing the research question.

- The disadvantage of a meta-analysis is that the studies can be very heterogeneous in their designs, quality, and patient populations and, therefore, it may not be valid to pool them.

- Researchers invoke two basic statistical models for meta-analysis, namely, fixed-effects models and random-effects models.

5. Meta-Analysis

- A fixed-effects model is more straightforward to apply, but its underlying assumptions are somewhat restrictive. It assumes that if all the involved studies had tremendously large sample sizes, then they all would yield the same result. In essence, a fixed-effects model assumes that there is no inter-study variability (study heterogeneity). The statistical model accounts only for intra-study variability.

- A random-effects model, however, assumes that the eligible studies actually represent a random sample from a population of studies that address the research question. It accounts for intra-study and inter-study variability. Thus, a random-effects model tends to yield a more conservative result, i.e., wider confidence intervals and less statistical significance than a fixed-effects model.

5. Meta-Analysis

- A random-effects model is more appealing from a theoretical perspective, but it may not be necessary if there is very low study heterogeneity. A formal test of study heterogeneity is available. Its results, however, should not determine whether to apply a fixed-effects model or random-effects model.

- The test for study heterogeneity is very powerful and sensitive when the number of studies is large. It is very weak and insensitive if the number of studies is small. Graphical displays provide much better information as to the nature of study heterogeneity.

- Some medical journals require that the authors provide the test of heterogeneity, along with a fixed-effects analysis and a random-effects analysis.

5. Meta-Analysis

- The basic step for a fixed-effects model involves the calculation of a weighted average of the treatment effect across all of the eligible studies.

- For a continuous outcome variable, the measured effect is expressed as the difference between sample treatment and control means. The weight is expressed as the inverse of the variance of the difference between the sample means.

- For a binary outcome variable, the measured effect usually is expressed as the logarithm of the estimated odds ratio from logistic regression. The weight is expressed as the inverse of the variance of the logarithm of the estimated odds ratio. For a time-to-event variable, a similar approach is taken for the logarithm of the estimated relative risk (hazard ratio).

5. Meta-Analysis

- Suppose that there are K studies.
- The estimated treatment effect (e.g., difference between the sample treatment and control means, logarithm of the odds ratio, etc.) in the k^{th} study, k = 1, 2, … , K, is Y_k
- The estimated variance of Y_k in the k^{th} study is S_k^2.
- The weight for the estimated treatment effect in the k^{th} study in the fixed-effects model is $w_k = 1/S_k^2$.
- The overall weighted treatment effect in the fixed-effects model is

$$Y = \left(\sum_{k=1}^{K} w_k Y_k \right) \Big/ \left(\sum_{k=1}^{K} w_k \right)$$

5. Meta-Analysis

- The estimated variance of Y in the fixed-effects model is

$$S^2 = 1 \Big/ \left(\sum_{k=1}^{K} w_k \right)$$

- Testing the null hypothesis of no treatment effect is performed from assuming that Y/S asymptotically follows a standard normal distribution. The approximate $100(1 - \alpha)\%$ confidence interval for the overall weighted treatment effect is

$$[Y - (z_{1-\alpha/2}S),\ Y + (z_{1-\alpha/2}S)]$$

- The statistic for testing H_0: {study homogeneity} is

$$Q = \sum_{k=1}^{K} w_k (Y_k - Y)^2$$

- Q has an asymptotic χ^2 distribution with K - 1 degrees of freedom.

5. Meta-Analysis

- The weighted analysis for the fixed-effects approach described previously corresponds to the following linear model for the k^{th} study, k = 1, 2, … , K:

  $$Y_k = \theta + e_k$$

  where Y_k is the observed effect in the k^{th} study, θ is the pooled population parameter of interest (difference in population treatment means, natural logarithm of the population odds ratio, etc.) and e_k is the random error term for the k^{th} study.

- It is assumed that $e_1, e_2, \ldots, e_K$ are independent random variables with e_k having a $N(0, \sigma_k^2)$ distribution. The variance term σ_k^2 reflects intra-study variability and its estimate is S_k^2. Usually, Y_k and S_k are provided as descriptive statistics in the k^{th} study report. The weighted average Y, described previously, is equal to $\hat{\theta}$.

5. Meta-Analysis

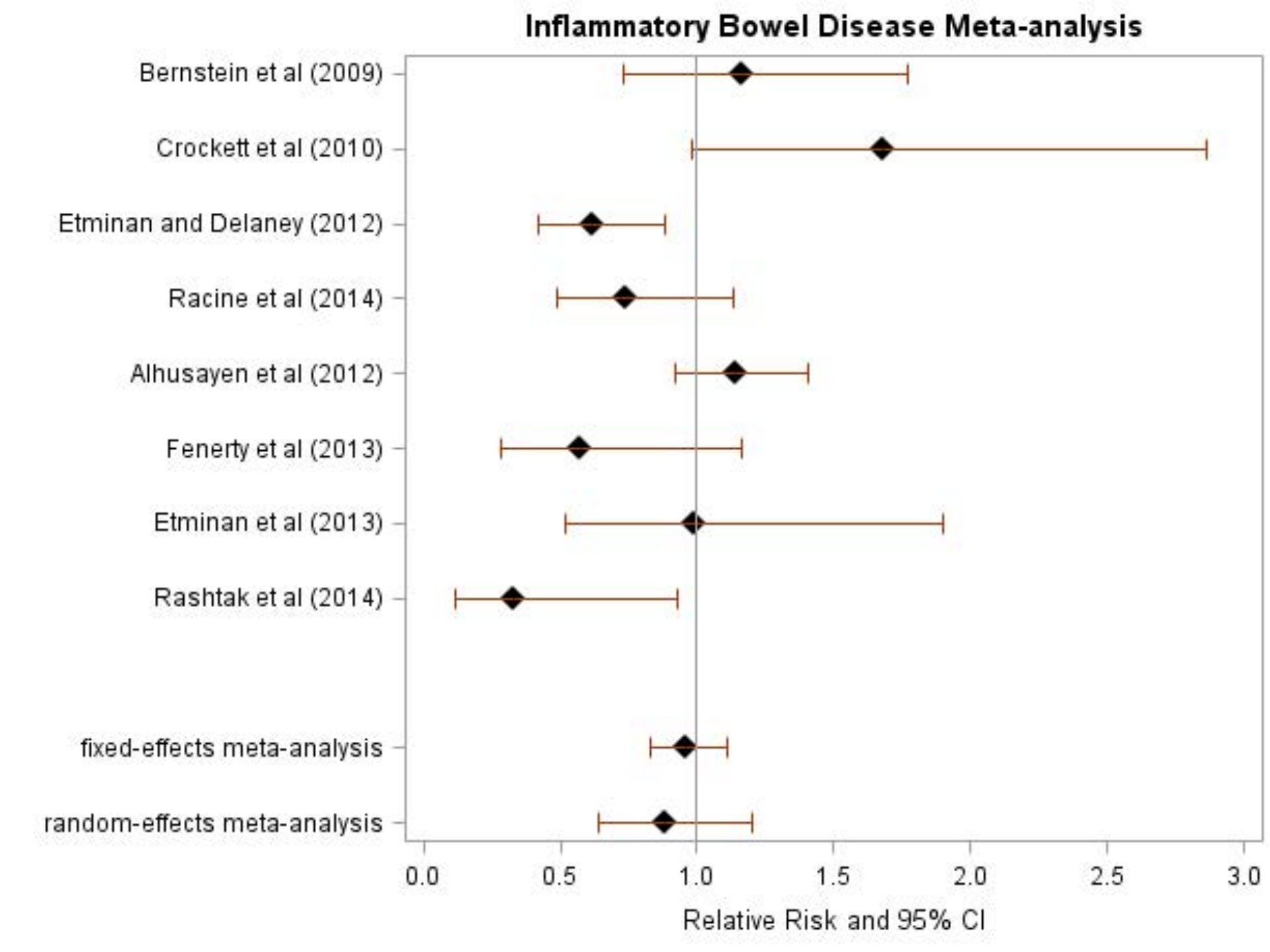
- A linear model for the random-effects approach is due to DerSimonian and Laird (*Controlled Clinical Trials* 1986):

$$Y_k = \theta + u_k + e_k$$

where Y_k, θ, and e_k are the same as described above and u_k is a random effect for the k^{th} study.

- It is assumed that $u_1, u_2, \ldots, u_K$ are independent and identically distributed as $N(0, \omega^2)$ random variables. The variance term ω^2 reflects inter-study variability.

- $Var(Y_k) = \sigma_k^2$ in the fixed-effects linear model, but $Var(Y_k) = \sigma_k^2 + \omega^2$ in the random-effects linear model.


Inflammatory Bowel Disease Meta-analysis
Bernstein et al (2009)
Crockett et al (2010)
Etminan and Delaney (2012)
Racine et al (2014)
Alhusayen et al (2012)
Fenerty et al (2013)
Etminan et al (2013)
Rashtak et al (2014)
fixed-effects meta-analysis
random-effects meta-analysis
0.0
0.5
1.0
1.5
2.0
2.5
3.0
Relative Risk and 95% CI