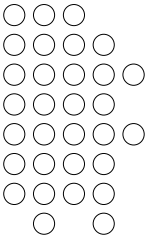
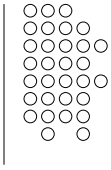
Meta Analysis

LJ Wei
Harvard University



Outline



- **Introduction to meta-analysis**
- **Principles of meta-analysis**
- **Formulating hypothesis and effect measures**
- **Methods for combining results across studies**
  - **Fixed effects pooling of primary study results**
  - **Between study heterogeneity**
  - **Random effects analysis**
  - **Meta Regression**
- **Meta analysis of rare events**
  - **Exact inference for combining 2x2 tables**

Introduction

What is Meta Analysis

- What is meant by the word "***Meta-analysis***"
  - ***Meta*** is Greek for "later in time"
  - ***Meta*** is now used to denote something that goes to a higher level or is more comprehensive.
  - **How is an analysis made more comprehensive?**
- In empirical research, there are often multiple studies addressing the same research question
  - A ***standard analysis*** attempts to reach a conclusion based on **a single study** without reference to any other studies.
  - A ***meta-analysis*** attempts to reach a conclusion based on **a set of studies** that address the same hypothesis.

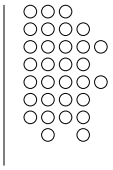
Introduction

What is Meta Analysis

- The **National Library** of **Medicine** (1989, pp. 1-40) defines **meta-analysis** as
  - "A quantitative method of combining the results of independent studies (usually drawn from the published literature) and synthesizing summaries and conclusions which may be used to evaluate therapeutic effectiveness, plan new studies, etc., with application chiefly in the areas of research and medicine"
- **Meta-analysis** began to be used as an index term that year.
- However, **Gene V. Glass** had begun using the term in 1976 (p. 3).
  - " Meta-analysis refers to the analysis of analyses. I use it to refer to the statistical analysis of a large collection of results from individual studies for the purpose of integrating the findings. 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. "

Introduction

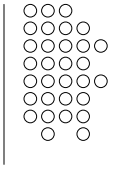
History of Meta Analysis



- In 1805, **Legendre** developed least squares to combine data on the orbits of comets from different observatories.
- In 1930's, statisticians working in agricultural research developed methods for combining the results of studies. Most notable are Fisher and Cochrane.
- In 1960's, Cohen popularized the notion of **effect size** for use in sample size determination in the social and behavioral sciences
  - Effect size measures the differences between null hypothesis and the truth
  - Effect size + sample size determines the power.
- In 1976, Grass published an article "Primary, secondary and meta-analysis of research". This is when the term "meta-analysis" was first used.

Introduction

History of Meta Analysis

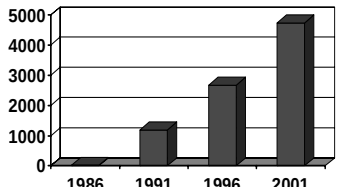


- In 1980's, meta-analysis became a widely used tool in social sciences and also received increasing attention from **clinical research**:
  - "*Statistical methods for meta-analysis*" by Hedges & Olkin published in 1980
  - Peto's method for binary data published in 1985; DerSimonian and Laird's random effects models for meta-analysis published in 1986.
- By 1990's, Meta-analysis starts being done in large scale in medicine within a new emphasis on **evidence based medicine**:
  - The English Cochrane center for systematic review was established in 1992
  - The International Cochrane Collaboration was formed in 1993

Introduction

Why Meta Analysis

- There are several reasons for conducting a meta-analysis of the results of previous studies:
  - **The increasingly large # of research studies**
    - 40,000 journals for the sciences
    - 1 article every 30 seconds



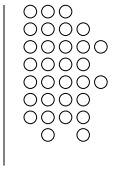
    - Unsystematic expert reviews of an area of research are often biased or years behind the current research.
    - Systematic and quantitative reviews are needed to summarize findings in a timely manner without bias.

Introduction

Why Meta Analysis

- **Scientific research is supposed to be replicable, and it is common to have several studies addressing the same hypothesis in different settings.**
  - Meta-analysis provides a way to consider replication and consistency of results across a set of studies without requiring that each study necessarily have large enough power to reach significance.
  - **Meta analysis can be used to increase power.**
  - **Meta analysis can be used to examine whether studies do not replicate each other and reach different conclusions.**
    - This may lead to scientific advances in several ways:
      - Determine variations of the treatment effect that produce a greater effect
      - Identify sub-populations that respond to treatment better

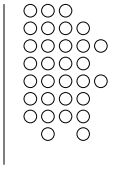
Principles of Meta-analysis



- **Meta-analysis is typically a two stage process**
  - **1) Summary statistic calculated from each study**.
    - For controlled trials, these values describe the treatment effects observed in each trial.
  - **2) *Pooled effect* is calculated by combining treatment effect estimates from individual studies.**
    - Typically a weighted average of individual effects

- **The combination of treatment effect estimates across studies may assume**
  - **Fixed effects**: treatment effects the same across studies.
  - **Random effects**: treatment effects ~ a distribution across studies.

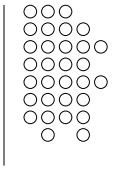
Principles of Meta-analysis



- **The standard error or sampling distribution of the *pooled* effect can be used to derive**
  - A confidence interval
  - A p-value for testing whether there is a treatment difference.

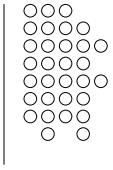
- **In addition to providing a measure of overall average treatment effect, meta-analysis methods can provide an assessment of**
  - Whether the variation among the results of the separate studies is compatible with random variation
  - Whether it is large enough to indicate inconsistency of treatment effects across studies → heterogeneity

Formulating Hypothesis and Effect Measures



- Before conducting a meta analysis, it is important to decide **the hypothesis or aim of the analysis**. When formulating a hypothesis for meta-analysis, it is important to determine
  - the precise question the meta-analysis aims to address
  - whether the meta analysis is exploratory or hypothesis testing
    - Hypothesis testing: is the intent of the study to provide a definitive test (usually a test of average effect = 0)
    - Exploratory: are there variations in the treatment or characteristics of the studies that lead to better outcomes?
- One also needs to select an appropriate **effect measure**

Type of Data and Effect Measures



- **Dichotomous or binary outcome**

	Event	No Event	Total	Event Rates
Intervention	m_1	$n_1 - m_1$	n_1	$p_1=m_1/n_1$
Control	m_0	$n_0 - m_0$	n_0	$p_0=m_0/n_0$

- Relative Risk (RR) $= \dfrac{m_1 / n_1}{m_0 / n_0} = \dfrac{p_1}{p_0}$
- Odds Ratio (OR) $= \dfrac{m_1 /(n_1 - m_1)}{m_0 /(n_0 - m_0)} = \dfrac{p_1 /(1 - p_1)}{p_0 /(1 - p_0)}$
- Risk Difference (RD) $= p_1 - p_0$

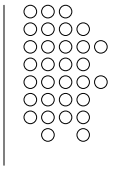
Type of Data and Effect Measures

- **Dichotomous or binary outcome**
  - when events of interest are rare
    - $1 - p_1 \cong 1$ and $1 - p_0 \cong 1$ ➔ RR ≅ OR
    - essential to assess the absolute risk in addition to relative risk
  - RR and OR = ∞ if there are no events in the control group
  - RR and OR **not defined** if both groups have zero events
    - Standard procedures either exclude studies with 0/0 events or add 0.5 to empty cells.

Type of Data and Effect Measures

- **Continuous outcome**
  - Mean difference (MD)
  - Standardized mean difference (SMD) $= \frac{\text{MD}}{\text{Std Deviation of Outcome}}$
    - All trials assess the same outcome, but measure it in a variety of ways (e.g depression measured with different psychometric scales.)
    - Inherently assumes that the differences in the standard deviation among trials reflect differences in the measurement scales and not real differences in variability among trial populations.
    - Overall treatment effect difficult to interpret (units in the standard deviation), but related to the term "effect size" which is frequently used in the social sciences.

Type of Data and Effect Measures

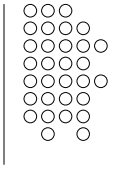


- **Ordinal outcome**
  - Takes values $y_1 < y_2 < \ldots < y_K$
    - Example: "mild", "moderate" and "severe".
  - When the number of categories is large, such data are often analyzed as continuous data.
  - One may transform ordinal data into binary data by combining adjacent categories.
  - Proportional odds ratio under a proportional odds model

$$\text{logit}\, P(Y \le y_k \mid \text{Trt}) = \alpha_k + \beta \cdot \text{Trt}$$

    - β, the log-odds ratio, summarizes the treatment effect

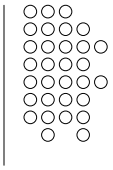
Type of Data and Effect Measures



- **Time to event (survival) outcome**
  - Examples: time to death, time to the onset of disease
  - Event time outcomes are subject to censoring due to loss to follow up or end of study
  - Time to event data sometimes are analyzed as dichotomous data by considering the probability of t-year survival.
  - The most common approach is to express the treatment effect as a hazard ratio under a proportional hazard assumption

$$\text{logit}\, P(Y \le y_k \mid \text{Trt}) = \alpha_k + \beta \cdot \text{Trt}$$

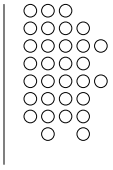
Statistical Methods for Combining Results Across Studies



- What is the ***true effect***? Depends on the underlying assumption about the study specific effect
  - **Fixed Effects Assumption**
    - Assumes that all studies have the same true effect
    - Variability only within each study
    - Precision depends mainly on study size
  - **Random Effects Assumption**
    - Studies allowed to have different underlying or true effects
    - Allows variation between studies as well as within studies
- **Basic assumption**: study results are independent.

Fixed Effects

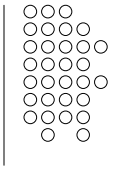
Pooling of Primary Study Results



- **Under the fixed effects framework, various procedures have been proposed**
  - **Vote count**
  - **p-value methods**
    - **Minimum p-value**
    - **Sum of -2log p-value**
    - **Probit of the p-value**
  - **Pooling using effect sizes**
    - **Inverse variance**
    - **Binary outcomes**
      - **Cochrane-Mantel-Hansel**
      - **Peto method**

Fixed Effects

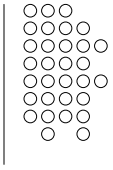
Vote Count



- Consider # of studies in favor of the conclusion (say, reach the 0.05 level of significance) and examine if they are the majority
- This approach has been used a lot due to its simplicity, but has several drawbacks
  - Significance depends on study sample size and effect size.
  - Even if the null hypothesis is wrong and studies are not small, the percentage of trials reaching significance could still be less than 50% ➔ low power of detecting a treatment effect
  - Vote counts do not provide an estimate of effect size.

Fixed Effects

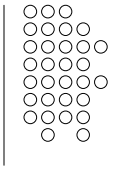
P-value Methods



- Consider S studies with the same null hypothesis H_0. Each study has a test statistic T_i and p-value p_i. We are interested in testing using all studies at α level.
  - Under H_0, $p_i \sim$ Uniform(0,1) and $\Phi^{-1}(p_i) \sim N(0, 1)$.
  - **Minimum p-value (Tippett)**:
    - $\min\{p_1, \ldots, p_S\} < 1 - (1-\alpha)^{1/S}$
  - **Sum of -2log p-value (Fisher)**
    - $-\sum_{i=1}^{S} 2\log(p_i) > \chi^2_{2S}(1-\alpha)$
  - **Probit of the p-value**
    - $S^{-1/2}\sum_{i=1}^{S} \Phi^{-1}(p_i) < \Phi^{-1}(1-\alpha)$

Fixed Effects

Pooling Using Effect Sizes



- Since Glass's work in 1976, combining effect sizes has become the main form of meta-analysis.
- Suppose the estimated effect sizes are

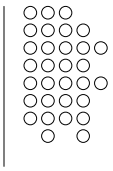
$$\{\hat{\beta}_i,\ i = 1,\ldots,S\}$$

- To ascertain the true underlying effect, a common approach is to consider a weighted average of the effect estimates from individual studies:

$$\hat{\beta} = \sum_{i=1}^{S} w_i \hat{\beta}_i \Big/ \sum_{i=1}^{S} w_i$$

Fixed Effects

Pooling Using Effect Sizes



- **Inverse Variance Method:**
  - **Under fixed effects framework:**

$$\hat{\beta}_i = \beta_0 + \varepsilon_i,\ i = 1,\ldots,S$$

    - β_0 is the true value of the common effect
    - ε_i represents the sampling variability of $\hat{\beta}_i$
      - $E(\varepsilon_i) \cong 0,\ \mathrm{var}(\varepsilon_i) = \sigma^2_{\varepsilon_i} = \mathrm{var}(\hat{\beta}_i)$
  - **Pooled estimate of β_0:** $\hat{\beta}_w = \sum_{i=1}^{S} w_i \hat{\beta}_i \Big/ \sum_{i=1}^{S} w_i$
    - Weights chosen to minimize the variance
    - **Optimal minimum variance weights:** $\hat{w}_i^{\min \mathrm{var}} = \hat{\sigma}_{\varepsilon_i}^{-2}$

Fixed Effects

Pooling Using Effect Sizes

- **Binary Outcomes**
  - When the event rates are low or trial sizes are small, the standard error estimates used in the inverse variance method may be poor.
  - **Cochrane-Mantel-Hansel** uses a different weighting scheme that depends upon the effect measure (eg RR, OR, RD).
    - **Cochrane-Mantel-Hansel pooled OR:**

$$\mathrm{OR}^{\mathrm{MH}} = \sum_{i=1}^{S} \hat{w}_i^{\mathrm{MH}} \mathrm{OR}_i \Big/ \sum_{i=1}^{S} \hat{w}_i^{\mathrm{MH}} \qquad \hat{w}_i^{\mathrm{MH}} = \frac{m_{0i}(n_{1i} - m_{1i})}{n_{1i} + n_{0i}}$$

i^{th} Study	Event	No Event	Total
Intervention	$\mathbf{m_{1i}}$	$\mathbf{n_{1i} - m_{1i}}$	$\mathbf{n_{1i}}$
Control	$\mathbf{m_{0i}}$	$\mathbf{n_{0i} - m_{0i}}$	$\mathbf{n_{0i}}$

$$\mathrm{OR}_i = \frac{m_{1i}(n_{0i} - m_{0i})}{m_{0i}(n_{1i} - m_{1i})}$$

Fixed Effects

Pooling Using Effect Sizes

- **Binary Outcomes**
  - **Peto's Method:**
    - based on an approximation to the data likelihood, expressed as a difference between observed and expected counts to estimate the pooled log OR and to test heterogeneity.

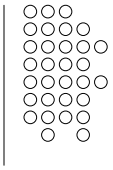
i^{th} Study	Event	No Event	Total
Intervention	$\mathbf{m_{1i}}$	$\mathbf{n_{1i} - m_{1i}}$	$\mathbf{n_{1i}}$
Control	$\mathbf{m_{0i}}$	$\mathbf{n_{0i} - m_{0i}}$	$\mathbf{n_{0i}}$

$$\mathrm{O}_i = m_{1i}\ ,\ \mathrm{E}_i = \frac{n_{1i}}{n_{1i}+n_{0i}}(m_{1i}+m_{0i})$$

$$\mathrm{V}_i = \mathrm{E}_i \frac{n_{1i}(n_{1i}-m_{1i}+n_{0i}-m_{0i})}{(n_{1i}+n_{0i})(n_{1i}+n_{0i}-1)}$$

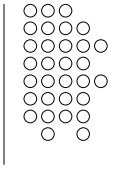
  - $\log \mathrm{OR}^{\mathrm{Peto}} = \sum_{i=1}^{S} (\mathrm{O}_i - \mathrm{E}_i) \Big/ \sum_{i=1}^{S} \mathrm{V}_i \qquad \mathrm{var}(\log \mathrm{OR}^{\mathrm{Peto}}) = 1 \Big/ \sum_{i=1}^{S} \mathrm{V}_i$

Between Study Heterogeneity



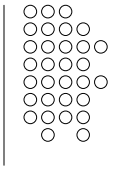
- The key assumption of fixed effects meta analysis methods is that all primary studies are estimating the same underlying ***true effect***
- The underlying effects across studies may be heterogeneous
  - Each study effect size $\hat{\beta}_i$ is estimating an individual population effect β_i.
  - As study sample size $N_i \rightarrow \infty$, $\hat{\beta}_i \rightarrow \beta_i$
  - Some of the β_i may be the same, but not all of them.

Between Study Heterogeneity



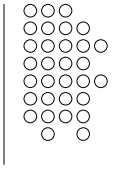
- Sources of heterogeneity
  - Patient selection
    - inclusion/exclusion criteria, disease severity/ type, patient characteristics, geographic differences
  - Treatment administration
    - duration of treatment, dose, blinding of treatment, compliance
  - Study type
    - clinical trial, case control study, cohort study
  - Types of controls
    - hospital controls, population controls, different disease controls
  - Analysis performed
    - intent to treat vs completer analysis, outcome measure used

Between Study Heterogeneity



- Testing for heterogeneity
  - Cochran's Q-test: $Q = \sum_{i=1}^{S} \hat{w}_i (\hat{\beta}_{\text{pooled}} - \hat{\beta}_i)^2 \sim \chi^2_{S-1}$ under H_0
    - Provides a measure of between study variation.
  - Other descriptive measures of heterogeneity
    - H statistic: $H = \sqrt{Q/(S-1)}$ has mean 1 under H_0.
      - Hoggins and Thompson (2002) suggested: H > 1.5 ➔caution regarding heterogeneity; H < 1.2 ➔ little heterogeneity
    - I^2 statistic = $(H^2 - 1)/H^2$
      - % of total variability in effect size due to between study variation
      - $I^2 \sim 0$ ➔ little heterogeneity; $I^2 \sim 1$ ➔ high heterogeneity
      - termed the "inconsistency" of the trials included in meta-analysis and has become a preferred measure of heterogeneity.

Random Effects



- Under random effects model, $\hat{\beta}_i$ estimates an underlying study specific effect size β_i.

$$\hat{\beta}_i = \beta_i + \varepsilon_i, \quad \varepsilon_i \sim N(0, \sigma_i^2), \ \beta_i \sim N(\beta_0, \tau^2), \quad i = 1,...,S$$

  - Bayesian model with a Gaussian prior distribution for β_i
  - The variance τ^2 represents between study variation.
  - The study specific variance σ_i^2 represents within study variation which goes to 0 as study sample size ➔ ∞
  - Pooled estimate: $\hat{\beta}^{\text{RE}} = \sum_{i=1}^{S} \hat{w}_i^{\text{RE}} \hat{\beta}_i \Big/ \sum_{i=1}^{S} \hat{w}_i^{\text{RE}} \quad \hat{w}_i^{\text{RE}} = \frac{1}{\hat{\tau}^2 + \hat{\sigma}^2_{\varepsilon_i}}$

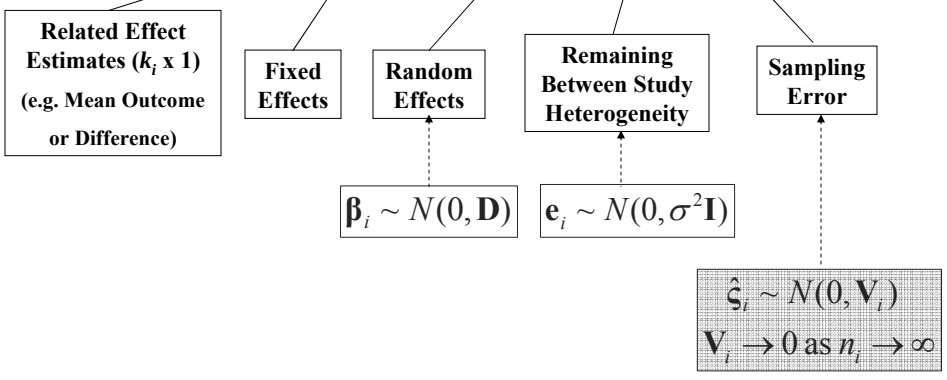
Meta Regression

- The pooled effect size estimates the ***average effect*** across all studies
- In the presence of heterogeneity
  - the validity of such an average measure?
  - not a single population effect size that applies to all studies
  - random effects pooling addresses heterogeneity to some extent
- Meta regression
  - provide an alternative approach that allows exploration of **why** studies have varied effect sizes.
  - one uses characteristics of the studies to explain the excess variation in effect sizes
  - Thompson and Higgins (2002) reviewed several meta-regression methods

Meta Regression

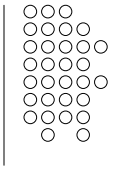
- Stram (1996) proposed a ***linear mixed effects*** model

$$\hat{\mathbf{Y}}_i = \mathbf{X}_i\boldsymbol{\alpha} + \mathbf{Z}_i\boldsymbol{\beta}_i + \mathbf{e}_i + \hat{\boldsymbol{\zeta}}_i$$



Meta Regression

Special Cases



- The DerSimonian and Laird (1986) Model:

$$\hat{Y}_i = \alpha + \beta_i + \hat{\zeta}_i$$

- The Begg and Pilote (1991) Model:

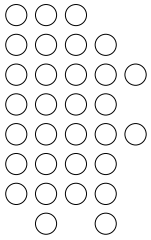
$$\hat{\mathbf{Y}}_i = \mathbf{X}_i\boldsymbol{\alpha} + \beta_i + \hat{\zeta}_i$$

  - This model allows the inclusion of single treatment historical controls as well as comparative trials in a treatment effect assessment.

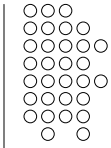
$$\hat{\mathbf{Y}}_i = \mathbf{X}_i \begin{pmatrix} \alpha_{\text{chemo}} \\ \delta \end{pmatrix} + \beta_i + \hat{\zeta}_i$$

- **20 estimates of the probability of 2-year disease free survival**
  - **4 trials with both BMT and Chemotherapy, → Y_i is a 2x1 vector, $\mathbf{X}_i = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$**
  - **12 single arm studies:**
    - **4 BMT only → Y_i is a scalar, $X_i = (1, 1)$**
    - **8 Chemotherapy only → Y_i is a scalar, $X_i = (1, 0)$**

Exact Procedure for Combining 2x2 Tables for Rare Events



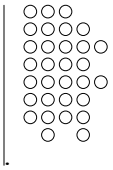
Exact Procedure for Combining 2x2 Tables for Rare Events



- Standard inference procedures for meta-analysis rely on large sample approximations to the distributions of the combined point estimators. However, such approximation may be inaccurate when
  - the individual study sample sizes are small, or
  - total number of studies is not large, or
  - event rates are low
- When the events of interest are very rare, many studies may have 0 events in one or both groups. Standard procedures
  - either excluding studies with 0 events,
  - or add 0.5 to empty cells

Example

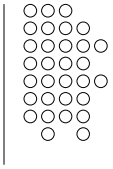
Effect of Rosiglitazone on MI or CVD Deaths



- Nissen and Wolski (2007) performed a meta analysis to examine whether Rosiglitazone (Avandia, GSK), a drug for treating type 2 diabetes mellitus, significantly increases the risk of MI or CVD related death.
- Avandia was introduced in 1999 and is widely used as monotherapy or in fixed-dose combinations with either Avandamet or Avandaryl. The original approval of Avandia was based on the ability of the drug to reduce blood glucose and glycated hemoglobin levels.
- Initial studies were not adequately powered to determine the effects of this agent on microvascular or macrovascular complications of diabetes, including cardiovascular morbidity and mortality.

Example

Effect of Rosiglitazone on MI or CVD Deaths



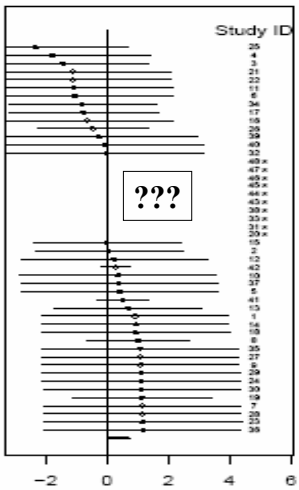
- However, the effect of any anti-diabetic therapy on cardiovascular outcomes is particularly important because more than 65% of deaths in patients with diabetes are from cardiovascular causes.
- Of 116 screened studies, 48 satisfied the inclusion criteria for the analysis proposed in Nissen and Wolski (2007).
  - 42 studies were reported in Nissen and Wolski (2007), the remaining 6 studies have zero MI or CVD death
  - 10 studies with zero MI events
  - 25 studies with zero CVD related deaths

ID	Rosiglitazone Group			Control Group		
	Patient	MI	Death	Patient	MI	Deat
1	357	2	1	176	0	0
2	391	2	0	207	1	0
3	774	1	0	185	1	0
4	213	0	0	109	1	0
5	232	1	1	116	0	0
6	43	0	0	47	1	0
7	121	1	0	124	0	0
8	110	5	3	114	2	2
9	382	1	0	384	0	0
10	284	1	0	135	0	0
11	294	0	2	302	1	1
12	563	2	0	142	0	0
13	278	2	0	279	1	1
14	418	2	0	212	0	0
15	395	2	2	198	1	0
16	203	1	1	106	1	1
17	104	1	0	99	2	0
18	212	2	1	107	0	0
19	138	3	1	139	1	0
20	196	0	1	96	0	0
21	122	0	0	120	1	0
22	175	0	0	173	1	0
23	56	1	0	58	0	0
24	39	1	0	38	0	0

ID	Rosiglitazone Group			Control Group		
	Patient	MI	Death	Patient	MI	Deatl
25	561	0	1	276	2	0
26	116	2	2	111	3	1
27	148	1	2	143	0	0
28	231	1	1	242	0	0
29	89	1	0	88	0	0
30	168	1	1	172	0	0
31	116	0	0	61	0	0
32	1172	1	1	377	0	0
33	706	0	1	325	0	0
34	204	1	0	185	2	1
35	288	1	1	280	0	0
36	254	1	0	272	0	0
37	314	1	0	154	0	0
38	162	0	0	160	0	0
39	442	1	1	112	0	0
40	394	1	1	124	0	0
41	2635	15	12	2634	9	10
42	1456	27	2	2895	41	5
43*	101	0	0	51	0	0
44*	232	0	0	115	0	0
45*	70	0	0	75	0	0
46*	25	0	0	24	0	0
47*	196	0	0	195	0	0
48*	676	0	0	225	0	0

- **Event Rates from 0% to 2.70% for MI**
- **Event Rates from 0% to 1.75% for CVD Death**

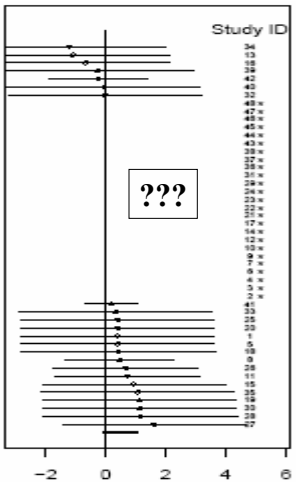
MI



Log Odds Ratio

95% CI: (1.03, 1.98); p-value = 0.03 (in favor of the control)

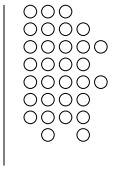
CVD Death



Log Odds Ratio

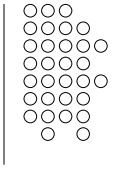
95% CI: (0.98, 2.74); p-value = 0.06

Exact Meta-Analysis Procedure
Combining Exact Intervals



- **Questions**:
  - *Could we combine information across studies without*
    - *excluding studies with 0 events or*
    - *artificial imputation?*
  - *Could we make **exact inference** without relying on possibly inaccurate large sample approximations when*
    - *the total number of studies is small, or*
    - *the sample sizes of individual studies are small, or*
    - *when the event rates are low.*

Exact Meta-Analysis Procedure
Combining Exact Intervals

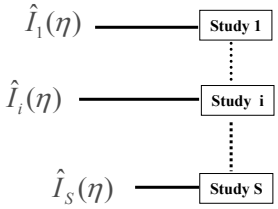


- Suppose we are interested in making inferences about a effect measure Δ whose true value is Δ_0
  - For example, the risk difference between two groups for the diabetes studies with respect to MI incidences.
- Specifically, suppose we are interested in constructing a $100(1 - \alpha)\%$ one-sided confidence interval (a, ∞) for Δ_0 based on all data from S independent studies.

Exact Meta-Analysis Procedure

Combining Exact Intervals

- For a given confidence level η,
  - one may obtain *S* study specific one-sided η-level confidence intervals for the risk difference.
  - each interval is constructed based on the data only from its corresponding study.

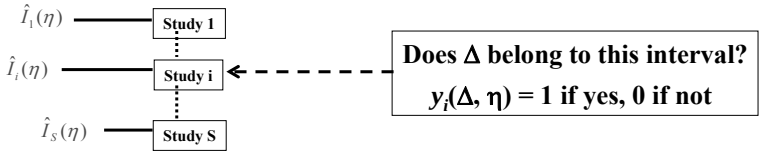


- For any given value of Δ, we examine whether Δ is the true value. If $\Delta = \Delta_0$, then by the definition of η-level confidence intervals
  - any given such interval should contain Δ with probability η
  - on average Δ should belong to at least 100η% of the above S independent intervals.

Exact Meta-Analysis Procedure

Combining Exact Intervals

- To determine whether a given value Δ should be included in the one-sided confidence interval (a, ∞)
  - we examine whether the probability that η-level study specific confidence intervals contain Δ is indeed at least η



  - Under the null that $\Delta = \Delta_0$,

$$P\{y_i(\Delta,\eta) = 1\} \geq \eta$$

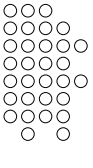
Exact Meta-Analysis Procedure
Combining Exact Intervals

- Thus, we propose to include Δ in the $100(1-\alpha)\%$ level confidence interval (a, ∞) if

$$t(\Delta,\eta)=\sum_{i=1}^{S} w_i\{y_i(\Delta,\eta)-\eta\}\geq c$$

  - where w_i is a study specific positive weight (e.g. sample size)
  - c is chosen such that $P(T(\eta) < c) \leq \alpha$,
  - $T(\eta)=\sum_{i=1}^{S} w_i(B_i-\eta)$ is the null counterpart of $t(\Delta,\eta)$
  - $\{B_i, i = 1, \ldots, S\}$ are n independent Bernoulli random variable with "success" probability of η.
- We repeat this process will all other possible values for Δ and obtain the final interval.

Exact Meta-Analysis Procedure
Combining Exact Intervals



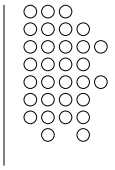
- One may further improve the interval estimate by using multiple intervals with a range of η levels from each study.
- Let $0 < \eta_1 < \ldots < \eta_K < 1$ and $\hat{I}_i(\eta_1) \subseteq \cdots \subseteq \hat{I}_i(\eta_K)$
- For any given Δ, we include Δ in the final interval (a, ∞) if

$$t_{\text{comb}}(\Delta)=\sum_{j=1}^{K} v_j t(\Delta,\eta_j)\geq d \quad \text{where } P\left\{\sum_{k=1}^{K} v_j T_j(\eta_j)<d\right\}\leq\alpha$$

  - v_j is a positive weight for the η_j level intervals
  - $\{T_1(\eta_1),\ldots,T_K(\eta_K)\}$ is constructed based on correlated Bernoulli random vectors such that $\sum_{j=1}^{K} v_j T_j(\eta_j)$ represents the null counterpart of $t_{\text{comb}}(\Delta)$

Exact Meta-Analysis Procedure

Combining Exact Intervals



- Similarly, we obtain combined $(1 - \alpha)$% one sided interval $(-\infty, b)$ based on the corresponding one-sided study specific intervals.
- Thus, (a, b) would be a $(1 - 2\alpha)$% two-sided interval for the risk difference.
- A point estimator for Δ_0 may be obtained as $\hat{\Delta}$, the mid-point of the intersection of all two-sided intervals for Δ_0 across all possible values of α
  - $\hat{\Delta}$ is the value of Δ with the least evidence of being rejected as the truth.

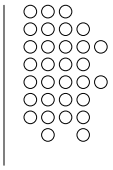
ID	Rosiglitazone Group			Control Group		
	Patient	MI	Death	Patient	MI	Deat
1	357	2	1	176	0	0
2	391	2	0	207	1	0
3	774	1	0	185	1	0
4	213	0	0	109	1	0
5	232	1	1	116	0	0
6	43	0	0	47	1	0
7	121	1	0	124	0	0
8	110	5	3	114	2	2
9	382	1	0	384	0	0
10	284	1	0	135	0	0
11	294	0	2	302	1	1
12	563	2	0	142	0	0
13	278	2	0	279	1	1
14	418	2	0	212	0	0
15	395	2	2	198	1	0
16	203	1	1	106	1	1
17	104	1	0	99	2	0
18	212	2	1	107	0	0
19	138	3	1	139	1	0
20	196	0	1	96	0	0
21	122	0	0	120	1	0
22	175	0	0	173	1	0
23	56	1	0	58	0	0
24	39	1	0	38	0	0

ID	Rosiglitazone Group			Control Group		
	Patient	MI	Death	Patient	MI	Deatl
25	561	0	1	276	2	0
26	116	2	2	111	3	1
27	148	1	2	143	0	0
28	231	1	1	242	0	0
29	89	1	0	88	0	0
30	168	1	1	172	0	0
31	116	0	0	61	0	0
32	1172	1	1	377	0	0
33	706	0	1	325	0	0
34	204	1	0	185	2	1
35	288	1	1	280	0	0
36	254	1	0	272	0	0
37	314	1	0	154	0	0
38	162	0	0	160	0	0
39	442	1	1	112	0	0
40	394	1	1	124	0	0
41	2635	15	12	2634	9	10
42	1456	27	2	2895	41	5
43*	101	0	0	51	0	0
44*	232	0	0	115	0	0
45*	70	0	0	75	0	0
46*	25	0	0	24	0	0
47*	196	0	0	195	0	0
48*	676	0	0	225	0	0

- **Event Rates from 0% to 2.70% for MI**
- **Event Rates from 0% to 1.75% for CVD Death**

Example

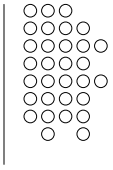
Effect of Rosiglitazone on MI or CVD Deaths



- For RR or OR effect measures
  - unless prior information about the underlying event rates is available, it is not clear how to utilize studies with zero events without continuity correction.
- RD may be used as an alternative effect measure
  - appealing interpretation
  - exact inference may be used
- We examine the effect of Rosiglitazone on MI or CVD deaths based on Δ = RD (Rosiglitazone – Control).

Example

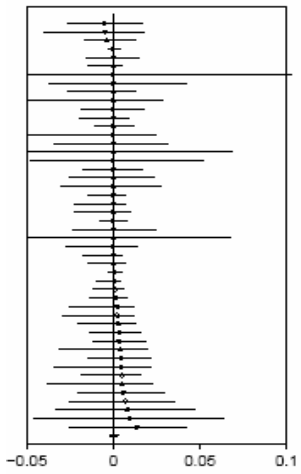
Effect of Rosiglitazone on MI or CVD Deaths



- Each study, we construct 20 exact confidence intervals at levels $\{\eta_1, \ldots \eta_{20}\}$ which are equally spaced from 0.1 to 0.95.
- Based on these individual intervals, we then construct the final combined interval based on the hypothesis testing procedure.

CVD Death

Exact Inference

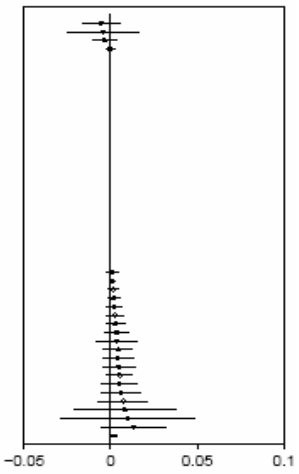


$\hat{\Delta} = 0.063\%$

95% CI: (-0.13, 0.23)%

P-value = 0.83

Asymptotic Inference



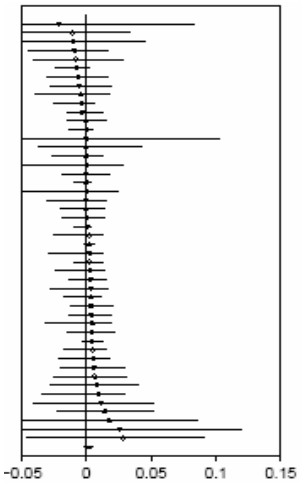
$\hat{\Delta} = 0.11\%$

95% CI: (0.00, 0.31)%

P-value = 0.05

MI

Exact Inference

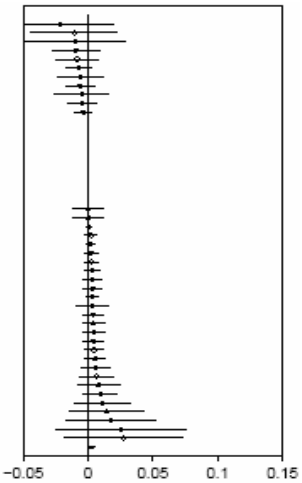


$\hat{\Delta} = 0.18\%$

95% CI: (-0.08, 0.38)%

P-value = 0.27

Asymptotic Inference

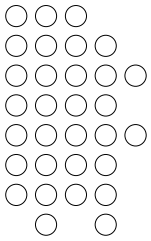


$\hat{\Delta} = 0.19\%$

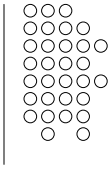
95% CI: (0.02, 0.42)%

P-value = 0.03

Non-parametric Inference for the Random Effects Distribution in Meta Analysis

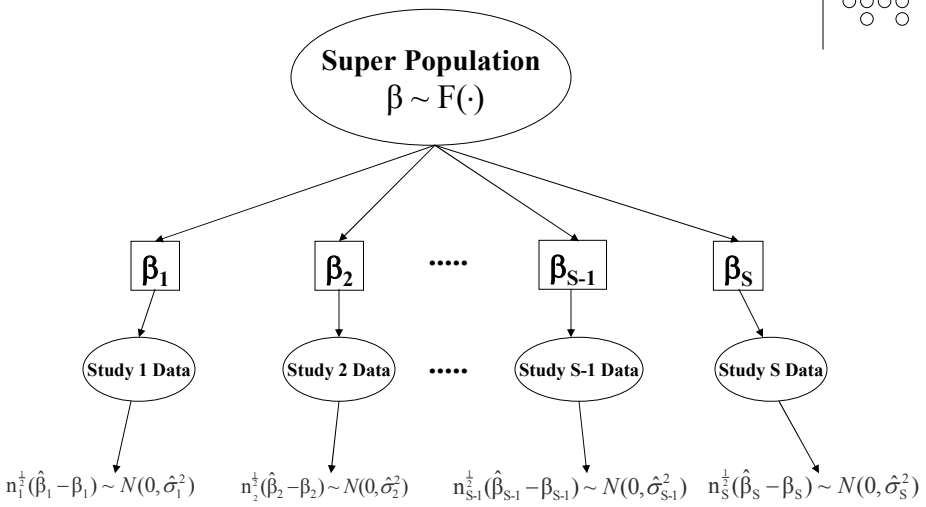


Random Effects Meta Analysis



- Fixed effects meta analysis methods assume that the true effects of interest are the same across all primary studies
  - The estimated study specific effect $\hat{\beta}_i$ converges to the same quantity as study sample size $n_i \rightarrow \infty$
- The underlying effects across studies may be heterogeneous
  - Each study effect size $\hat{\beta}_i$ is estimating an individual population effect β_i with $\hat{\beta}_i \rightarrow \beta_i$ as $n_i \rightarrow \infty$
  - Some of the β_i may be the same, but not all of them.

Random Effects Meta Analysis



Random Effects Meta Analysis

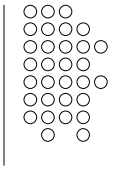
- Under random effects model with normal prior, DerSimonian and Laird (DL) (1986) proposed procedures for estimating the ***mean*** of the random effects β_0:

$$\beta_i = \beta_0 + \tau_i,\ n_i^{\frac{1}{2}}(\hat{\beta}_i - \beta_i) \sim N(0, \hat{\sigma}_i^2),\ i = 1,..., S$$

  - $E(\tau_i) = 0$ and $var(\tau_i)$ represents between study variation.
  - The study specific variance $\hat{\sigma}_i^2 / n_i \to 0$ as $n_i \to \infty$
  - **May not work well with when the number of studies is small!**

Random Effects Meta Analysis

Non-parametric Estimation of the Median



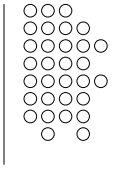
- Wang et al (2008) proposed interval estimation procedures for the quantiles of β_i without requiring the number of studies to be large.
- Suppose we are interested in estimating the median of β_i denoted by μ_0.
  - If β_i known, exact confidence interval for μ_0 can be obtained by inverting a sign test:

$$T(\mu) = \sum_{i=1}^{S} \{I(\beta_i < \mu) - 0.5\}$$

  - The null distribution of $T(\mu_0)$ + S/2 is a Binomial(S,0.5).

Random Effects Meta Analysis

Non-parametric Estimation of the Median



- If $\hat{\beta}_i$'s consistently estimate β_i, then one may consider the test statistic

$$\tilde{T}(\mu) = \sum_{i=1}^{S} \{I(\hat{\beta}_i < \mu) - 0.5\}$$

  - Unconditional null distribution of $\tilde{T}(\mu_0)$ + S/2 is approximately Binomial(S, 0.5).
  - However, the Bernoulli variable $I(\hat{\beta}_i < \mu)$ may not be a good surrogate for $I(\beta_i < \mu)$
    - If the variance of $\hat{\beta}_i$ is not small relative to the distance between β_i and μ.

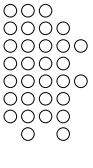
Random Effects Meta Analysis

Non-parametric Estimation of the Median

- Alternatively, one may replace $I(\hat{\beta}_i < \mu)$ with a measure of likelihood for the event $\beta_i < \mu$
  - Example: the observed coverage level of the interval (-∞,μ) for β_i which is
    $$\Phi\{(\mu - \hat{\beta}_i)/\hat{\sigma}_i\}$$
  - The test statistic based on the coverage level is
    $$\hat{T}(\mu) = \sum_{i=1}^{S}[\Phi\{(\mu - \hat{\beta}_i)/\hat{\sigma}_i\} - 0.5]$$
    - Studies with data that are more informative for the event $\beta_i < \mu$ would yield coverage level closer to either 0 or 1 and thus carry more weight in the test statistic.

Random Effects Meta Analysis

Non-parametric Estimation of the Median

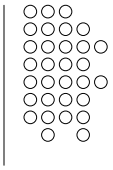


- Since
  $$\Phi\{(\mu - \hat{\beta}_i)/\hat{\sigma}_i\} - I(\beta_i < \mu) \to 0 \text{ in probability}$$
  $$\Phi\{(\mu - \hat{\beta}_i)/\hat{\sigma}_i\} \text{ symmetric around } 0.5$$

  one may approximate the null distribution of
  $$\hat{T}(\mu) = \sum_{i=1}^{S}[\Phi\{(\mu - \hat{\beta}_i)/\hat{\sigma}_i\} - 0.5]$$
  based on
  $$T^*(\mu) = \sum_{i=1}^{S}|\Phi\{(\mu - \hat{\beta}_i)/\hat{\sigma}_i\} - 0.5|(2\Delta_i - 1)$$
  - $\{\Delta_i, i=1,\ldots,S\} \sim$ Bernoulli(0.5) independent of data

Random Effects Meta Analysis

Non-parametric Estimation of the 100pth Percentile



- For the 100pth percentile, the test statistic is

$$\hat{T}_p(\mu) = \sum_{i=1}^{S} [\Phi\{(\mu - \hat{\beta}_i)/\hat{\sigma}_i\} - 0.5]$$

  - Unconditionally $\hat{T}_p(\mu)$ is asymptotically Binomial(S,p)

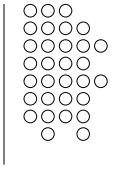
- The null distribution of $\hat{T}_p(\mu)$ can be approximated by

$$T_p^*(\mu) = \sum_{i=1}^{S} |\Phi\{(\mu - \hat{\beta}_i)/\hat{\sigma}_i\} - 0.5|\,(2\varepsilon_i - 1)$$

  - $\{\varepsilon_i, i=1,\ldots,S\} \sim$ Bernoulli(p) independent of data

Example

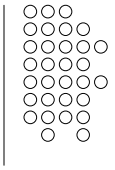
Effect of ESA on the Risk Mortality



- Bennett et al (2008): meta analysis to examine whether the erythropoietin-stimulating agents (ESA) for treating cancer chemotherapy-associated anemia would increase the risk of mortality
  - 51 phase III comparative trials (ESA vs placebo or standard of care)
  - Effect measure: hazard ratio
    - From each study, $\hat{\beta}_i$ is a consistent estimator of the underlying study specific hazard ratio β_i. Confidence intervals of β_i are also available to infer the within study variation.

Example

Effect of ESA on the Risk Mortality



95% Confidence Intervals for the Hazard Ratio

	Median	25th Percentile	75th Percentile
DL	(1.01, 1.20)		
$\widetilde{T}(\cdot)$ with I(·)	(0.90, 1.26)	(0.49, 0.93)	(1.25, 1.72)
$\hat{T}(\cdot)$ with Φ(·)	(0.94, 1.21)	(0.70, 0.99)	(1.18, 1.48)