

An Introduction to Meta-analysis

Maggie Chen

Univeristy of Toronto

Dec 8nd, 2011

Outlines

- ▶ What is meta-analysis?
- ▶ Data type and effect measurement
- ▶ Combining studies: fixed effect vs. random effect
- ▶ Meta-regression
- ▶ Publication bias
- ▶ Advanced statistical methods
- ▶ Summary

Definition of meta-analysis

Quantitative approach for **systematically combining** results of previous research to arrive at conclusions about the body of research

What does this mean?

- ▶ Quantitatives: numbers
- ▶ Systematic: methodological
- ▶ Combining: putting together
- ▶ Previous research: what has been done before
- ▶ Conclusions: new knowledge

Why do a meta-analysis?

- ▶ Decisions about the utility of an intervention or the validity of a hypothesis cannot be based on the results of a single study
- ▶ Meta-analysis can be used to synthesize data across studies
 1. to improve reliability of estimates of treatment effect
 2. to improve statistical power
 3. to inform specific decision

Steps to perform a meta-analysis

- ▶ Define a question of interest
- ▶ Collect studies that provide data on the question of interest
- ▶ Abstract data and compute the effect sizes
- ▶ Examine the effects and potential confounding variables
- ▶ Interpret the result and report

Generic outcomes for meta-analysis

Binary data (Event)

- Incidence, Odds Ratio, Relative Risk, Risk Difference and etc.

Continuous data

- Mean Difference, Standardized Mean Difference and etc.

Other outcome

- Hazard Ratio, Correlation coefficient and etc.

Binary outcome - Incidence rates

- ▶ The log incidence is given by:

$$\ln(\text{incidence rate}) = \ln(d/q)$$

where d events are observed in q person-years

- ▶ The approximate variance of the outcome is given by:

$$\text{var}(\ln(\text{incidence rate})) = 1/d$$

- ▶ The confidence interval for incidence rate is therefore:

$$\exp(\ln(\text{incidence rate}) \pm 1.96\sqrt{1/d})$$

Binary data - risk difference

	Yes	No	Total
Treatment	a_1	b_1	n_1
Control	a_2	b_2	n_2

Risk for each group: $\hat{r}_1 = \frac{a_1}{n_1}$ $\hat{r}_2 = \frac{a_2}{n_2}$

Risk difference: $\hat{RD} = \frac{a_1}{n_1} - \frac{a_2}{n_2}$

Standard error: $SE(\hat{RD}) = \sqrt{\frac{a_1 b_1}{n_1^2} + \frac{a_2 b_2}{n_2^2}}$

100(1- α)% CI: $\hat{RD} \pm Z_{\alpha/2} SE(\hat{RD})$

Binary data - relative risk

	Yes	No	Total
Treatment	a_1	b_1	n_1
Control	a_2	b_2	n_2

Risk for each group: $\hat{r}_1 = \frac{a_1}{n_1}$ $\hat{r}_2 = \frac{a_2}{n_2}$

Relative Risk: $\hat{RR} = \frac{a_1/n_1}{a_2/n_2}$

Standard error (logRR): $\sqrt{\frac{1-a_1/n_1}{a_1} + \frac{1-a_2/n_2}{a_2}}$

100(1- α)% CI for logRR: $\log(\hat{RR}) \pm Z_{\alpha/2} SE(\log(\hat{RR}))$

100(1- α)% CI for RR: $\exp(\log(\hat{RR}) \pm Z_{\alpha/2} SE(\log(\hat{RR})))$

Binary data - odds ratio

	Yes	No	Total
Treatment	a_1	b_1	n_1
Control	a_2	b_2	n_2

Odds for each group: $\hat{o}_1 = \frac{a_1}{b_1}$ $\hat{o}_2 = \frac{a_2}{b_2}$

Odds ratio: $\hat{OR} = \frac{a_1 b_2}{a_2 b_1}$

Standard error (logOR): $\sqrt{\frac{1}{a_1} + \frac{1}{a_2} + \frac{1}{b_1} + \frac{1}{b_2}}$

100(1- α)% CI for logOR: $\log(\hat{OR}) \pm Z_{\alpha/2} SE(\log(\hat{OR}))$

100(1- α)% CI for OR: $\exp(\log(\hat{OR}) \pm Z_{\alpha/2} SE(\log(\hat{OR})))$

OR, RR and RD

- ▶ Relative risk is often used in cohort study, but odds ratio is often used in case-control study
- ▶ Relative risk and risk difference are easy to interpret, but odds ratio is not
- ▶ Risk difference takes prevalence of event into account
- ▶ Odds ratio is equivalent to relative risk when prevalence of the disease is rare
- ▶ Odds ratio and relative risk are more stable compared to risk difference when applied to different population
- ▶ Odds ratio has good statistical properties and is more preferred by statisticians

Continuous data - mean difference

	Sample size	Mean	Standard deviation
Treatment	n_1	\bar{x}_1	s_1
Control	n_2	\bar{x}_2	s_2

Mean difference: $\bar{x}_1 - \bar{x}_2$

$$SE(\bar{x}_1 - \bar{x}_2) = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

Continuous data - standardized mean difference

	Sample size	Mean	Standard deviation
Treatment	n_1	\bar{x}_1	s_1
Control	n_2	\bar{x}_2	s_2

Standardized mean difference: $d = f \frac{\bar{x}_1 - \bar{x}_2}{s}$

where: $s = \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}}$ $f = \frac{4(n_1 + n_2 - 1) - 4}{4(n_1 + n_2 - 1) - 1}$

$$SE(d) = \sqrt{\frac{n_1 + n_2}{n_1 n_2} + \frac{d^2}{2(n_1 + n_2)}}$$

Mean difference and standardized mean difference

- ▶ Use mean difference when studies' outcomes are on same scale
- ▶ Use standardized mean difference when studies' outcomes are on different scale
- ▶ Standardized mean difference can also be derived from z , t , f and χ^2 statistics
- ▶ Cohen's conventional criteria is useful to interpret standardized mean difference

Correlation coefficient

- ▶ Suppose the correlation coefficient of two variables is r from a study with sample size n
- ▶ After adjust for bias, an approximation of population correlation coefficient is $G_{(r)} = r + \frac{r(1-r^2)}{2(n-3)}$
- ▶ It is conventional to convert r to a z-score in meta-analysis:
$$z_r = \frac{1}{2} \ln\left(\frac{1+r}{1-r}\right)$$
- ▶ If one wishes to estimate an unbiased population correlation coefficient, $G_{(r)}$ should be transformed to z-score
- ▶ z_r is normally distributed with standard deviation $\sqrt{\frac{1}{n-3}}$

Correlation coefficient

- ▶ Once the necessary computation is done, r can be transformed back as $r = \frac{e^{2zr} - 1}{e^{2zr} + 1}$
- ▶ Meta-analysts have formulas to calculate r from z , t , f and χ^2 statistics
- ▶ One should correct the correlation for its sample size bias, then convert this to a z -score before analyzing and finally transform it back

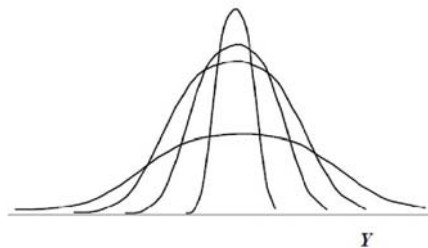
Combining studies

- ▶ Parameter of interest: θ .
- ▶ Data to collect from each study i :
 - Outcome of interest (OR, RR, MD...): Y_i
 - Its standard error: S_i
 - And sample size: n_i
- ▶ Methods: Weighting each study to overall summary effect by a function of standard error or sample size

Fixed effect model

- ▶ Assume the true effect θ is normally distributed with $N(\theta, S^2)$, where S^2 is the within study variance. Note there is no between studies variation allowed in fixed effect model
- ▶ Observe outcome Y_i and its variance S_i^2 from study i

Fixed effect model in picture



Fixed effect model

Study	Outcome	Standard Error	Weight
1	y_1	s_1	$w_1 = 1/s_1^2$
2	y_2	s_2	$w_2 = 1/s_2^2$
.	.	.	.
.	.	.	.
k	y_k	s_k	$w_k = 1/s_k^2$

Overall estimate: $\hat{\theta} = \frac{\sum_i w_i y_i}{\sum_i w_i}$ $SE(\hat{\theta}) = \frac{1}{\sqrt{\sum_i w_i}}$
 $100(1 - \alpha)\% CI : \hat{\theta} \pm Z_{\alpha/2} SE(\hat{\theta})$

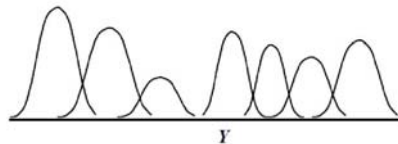
Other weighting methods for fixed effect

- ▶ Peto method for pooled odds ratio - asymptotically normal test statistic
- ▶ Mantel Haenszel method for pooled odds ratio across the strata of 2 by 2 tables - χ^2 test
- ▶ For Mantel Haenszel method, if any cell count in a table is zero then a continuity correction is applied to each cell in that table

Random effect model (DerSimonian-Laird)

- ▶ Assume outcome Y_i is normally distributed with $N(\theta_i, S_i^2)$, where S_i^2 is the within study variation
- ▶ θ_i is a realization from a distribution of the true effect θ distributed with $N(\theta, \tau^2)$
- ▶ Observe outcome Y_i and its variance S_i^2 from study i .

Random effect model in picture



Random effect model

Study	Outcome	Standard Error	Weight
1	y_1	s_1	$w_{n1} = 1/(s_1^2 + \tau^2)$
2	y_2	s_2	$w_{n2} = 1/(s_2^2 + \tau^2)$
.	.	.	.
.	.	.	.
k	y_k	s_k	$w_{nk} = 1/(s_k^2 + \tau^2)$

Where τ^2 is the between study variance and also called heterogeneity

$$\text{Overall estimate: } \hat{\theta} = \frac{\sum_i w_{ni} y_i}{\sum_i w_{ni}} \quad \text{se}(\hat{\theta}) = \frac{1}{\sqrt{\sum_i w_{ni}}}$$

$$100(1 - \alpha)\% \text{ CI} : \hat{\theta} \pm Z_{\alpha/2} \text{se}(\hat{\theta})$$

Testing heterogeneity

- ▶ A simple statistic to test the hypothesis that $\tau^2 = 0$ is

$$Q = \sum_i w_i (Y_i - \hat{\theta})^2$$

- ▶ Q has a χ^2 distribution with degree of freedom of $k - 1$
- ▶ Q statistic measures deviation of observed effect sizes from an underlying overall effect size
- ▶ The test is not very powerful if the number of studies is small or if the conditional variances are large
- ▶ Even if the test does not reject the hypothesis, it is unlikely the between study variance is zero
- ▶ The cut-off point is often set to 0.10

Testing heterogeneity

- ▶ Another statistic to test the hypothesis that $\tau^2 = 0$ is

$$I^2 = \max\left\{0, \frac{Q - k + 1}{Q} 100\%\right\}$$

- ▶ Interpreted as approximately the proportion of total variation due to heterogeneity rather than sampling error
- ▶ A rough guide to interpret I^2 : low 25%; medium 50% ; high 75%

Estimate between study variance

- ▶ τ can be estimated based on different distribution assumptions of random effect using MLE
- ▶ A simple distribution free estimate of τ is given by:

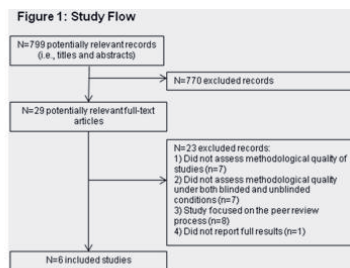
$$\tau^2 = \max\left\{0, \frac{Q - k + 1}{\sum_i w_i - \sum_i w_i^2 / \sum_i w_i}\right\}$$

- ▶ Where Q is defined before.

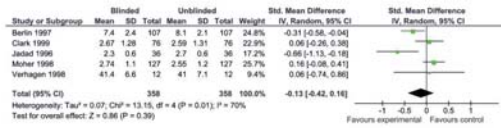
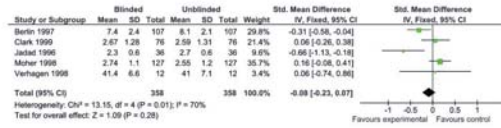
Compare fixed effect vs. random effect

- ▶ Fixed effect model
 1. There is one real value for the outcome and all studies estimate this one
 2. It is weighted by inverse of the variance
 3. The confidence interval is too narrow
- ▶ Random effect model
 1. There are many possible values for the outcome and each study estimate its own value
 2. It is weighted by inverse of the sum of variance and heterogeneity
 3. The confidence interval is more conservative

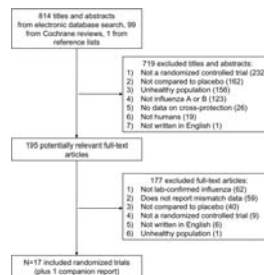
An example : blinded vs. unblinded assessment of risk of bias in studies



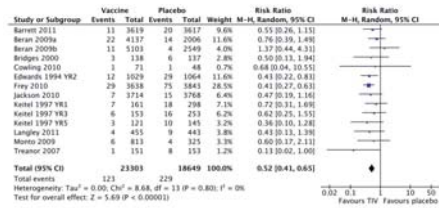
An example : fixed effect vs. random effect



An example : Influenza Vaccine Study



An example : Meta-analysis



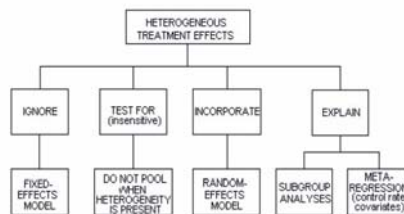
Heterogeneity

- ▶ Clinical heterogeneity
 1. PICO - patients, intervention, comparison and outcome
- ▶ Methodological heterogeneity
 1. Design: RCT, cohort study, survey and etc.
 2. Conduct: data abstract, analysis approaches, handling missing data
- ▶ Statistical heterogeneity
 1. Different methods
 2. Unknown or unrecorded study characteristic associated with outcome

How to identify heterogeneity?

- ▶ Common sense
Are the patients, interventions and outcomes in each of the included studies sufficiently similar
- ▶ Graphic aid
Forest plot, overlapping confidence intervals
- ▶ Statistical tests
Q test and I^2 statistic

How to deal with heterogeneity



Subgroup analysis

- ▶ Subsets of trials
- ▶ Subsets of patients
- ▶ Subgroups should be pre-specified
- ▶ Subgroup analysis should be few

Meta-regression

- ▶ Meta-regression investigates whether particular covariates explain any of the heterogeneity of treatment effects between studies
- ▶ A meta-regression is either a linear or logistic regression
- ▶ It could be fixed-effect or random-effect model
- ▶ The unit of analysis is a study, sometimes it could be a specific treatment arm
- ▶ Predictors in the regression are the study-level covariates

The pitfalls of meta-regression

- ▶ There are few trials in a meta-analysis and it is usually lack of statistical power
- ▶ Meta-regression is sometimes used to relate the results of the trials to published averages of patient characteristics within trials, which are difficult to interpret
- ▶ False positive conclusions deriving from meta-analyses are particularly hazardous

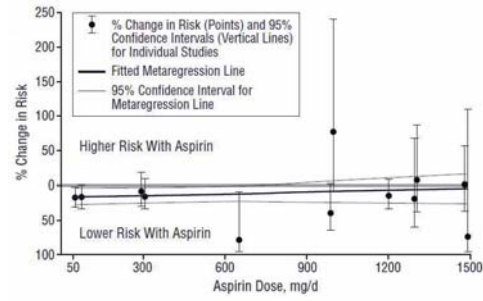
In a meta-analysis, which has accumulated the totality of evidence on a particular issue, there is no possibility of such external testing or validation. Thus we are left in the dangerous position of having speculative conclusions, without the ability to investigate them further

An example for meta-regression

Table 2. Risk Ratio Estimates and 95% Confidence Intervals for Stroke by Study and Aspirin Dose

Study No.	Aspirin Dose, mg/d	Risk Ratio	95% Confidence Interval	% Risk Reduction†
1	50	0.82	0.69-0.98	18
2	75	0.84	0.65-1.08	16
3	300	0.91	0.71-1.18	9
4	300	0.85	0.66-1.09	15
5	650	0.21	0.05-0.91	79
6	990	0.59	0.35-1.01	41
7	1000	1.77	0.92-3.40	-77
4	1200	0.85	0.66-1.08	15
8	1300	0.80	0.39-1.67	20
9	1300	1.06	0.61-1.86	-6
10	1500	0.25	0.03-2.10	75
11	1500	1.00	0.63-1.57	0

An example for meta-regression



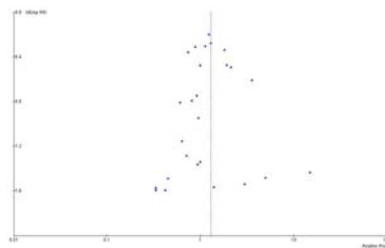
Publication bias

- ▶ The research that appears in the published literature is systematically unrepresentative of the population of completed studies
- ▶ Funnel plot to detect publication bias
- ▶ Formal test with low power:
 - Rank correlation test
 - Linear regression test

Funnel plot

- ▶ Scatter plot of effect estimate (x-axis) against inverse of its variance (y-axis)
- ▶ If no bias, expect symmetric and inverted funnel
- ▶ If bias, expect asymmetric or skewed shape

An example for funnel plot



Funnel plot

An asymmetric funnel plot might be caused by:

- ▶ Publication bias
- ▶ Variation among the quality of the studies
- ▶ Different intervention duration, or underlying risks
- ▶ Poor design
- ▶ Inadequate analysis

Cumulative meta-analysis

- ▶ It is the retrospective undertaking of a series of meta-analysis, each one repeated upon the inclusion of one additional study
- ▶ Cumulative meta-analysis is most often used to track evidence over time
- ▶ It is recommended to use sequential methods attempt to retain type I and type II error rates over repeated analysis

Mixed treatment comparison

- ▶ Traditional meta-analysis is a standardized pairwise comparison between A and B
- ▶ Mixed comparison can include data structures such as A vs. B, B vs. C and A vs. C
- ▶ It strengthens inference concerning the relative efficacy of two treatments, by including both direct and indirect comparisons
- ▶ It facilitates simultaneous inference regarding all treatments, in order for example to select the best treatment
- ▶ It is conducted using a Bayesian approach

Multivariate meta-analysis

- ▶ More than one outcome is of the interest
- ▶ One has to consider the variance and covariance structure for the outcomes
- ▶ It can be extended to mixed comparison multivariate meta-analysis

Software

- ▶ Revman
- ▶ Stata
- ▶ Winbugs
- ▶ R and SAS

Summary

- ▶ Meta-analysis is a statistical literature review of magnitude of an effect
- ▶ Meta-analysis uses the magnitude of the effect, its precision and other information to produce a weighted mean
- ▶ It is important to examine heterogeneity and publication bias
- ▶ A weakness of meta-analysis is that sources of bias are not controlled by the method; Garbage In and Garbage Out

Reference

- ▶ DerSimonian R, Laird N. Meta-analysis in clinical trials. *Controlled Clinical Trials* 1986; 7:177–188.
- ▶ Berkey CS, Hoaglin DC, Mosteller F, Colditz GA. A random-effects regression model for meta-analysis. *Statistics in Medicine* 1995; 14:395–411.
- ▶ Van Houwelingen HC, Arends LR, Stijnen T. Tutorial in biostatistics. Advanced methods in meta-analysis: multivariate approach and meta-regression. *Statistics in Medicine* 2002; 21:589–624.
- ▶ Sutton A, Abrams K, Jones D, Sheldon T, Song F, *Methods for Meta-Analysis in Medical Research* statistics, Wiley, 2000