A Framework of Causal Inference in Meta-Analysis

Presentation for G8325 Causal Inference

Meta-Analysis

- Meta-analysis, which collect and combine findings from several different but related studies, is becoming increasingly popular in fields such as evidence-based medicine.

Meta-Analysis

- Meta-analysis, which collect and combine findings from several different but related studies, is becoming increasingly popular in fields such as evidence-based medicine.
- However, several pitfalls are well-known in the practice of meta-analysis, such as publication bias, uneven qualities and heterogeneous results of of the included studies.
- We will deal with the heterogeneity of the estimates of effect size of the studies.

Visualizing Heterogeneity of Meta-Analysis

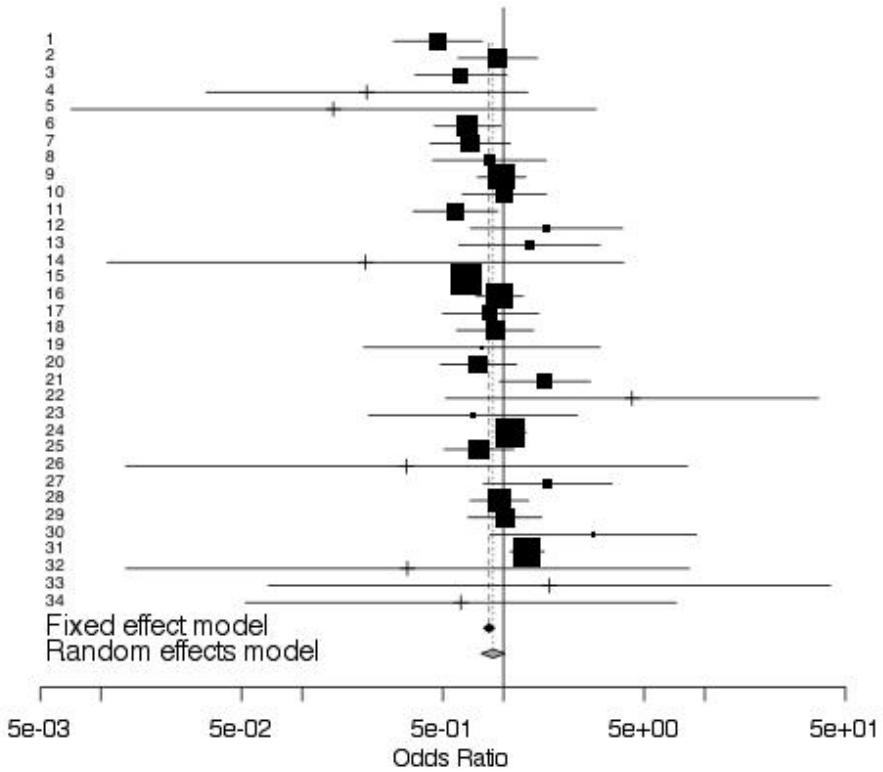

Figure: The typical approach of visualizing a meta-analysis: forest plot. The last two rows give Fixed Effects and Random Effects estimates.

Dealing with Heterogeneity: Fixed Effects model

- The most natural way of dealing with heterogeneity is to model it as sampling errors, mathematically

$$\hat{\theta}_i = \theta + \varepsilon_i$$

in which θ is the true underlying effect of interest.

- Then a re-weighted method based on the standard errors of the study-level estimates can be used to estimate θ

$$\hat{\theta} = \frac{\sum w_i \hat{\theta}_i}{\sum w_i}$$

in which

$$w_i = 1/\mathrm{Var}(\hat{\theta})$$

- Unfortunately, this method is dubbed Fixed Effects model, which might cause some confusions.

Dealing with Heterogeneity: Random Effects model

- The predominant way of dealing with heterogeneity of meta-analysis is through Random Effect model, made popular by DeSimonian and Laird (1986).

- Let $\hat{\theta}_i$ denote the effect size estimate from the $i-$th study, Random Effect model assumes that it is the sum of the true effect size for the $i-$th study, θ_i, and sampling error ε_i

$$\hat{\theta}_i = \theta_i + \varepsilon_i$$

- And the true effect size of each study comes from some distribution with mean θ and sd σ. θ is considered as the ultimate quantity of interest.

$$\theta_i \overset{\text{iid}}{\sim} G(\theta, \sigma^2)$$

- The estimate is given by

$$\hat{\theta} = \frac{\sum w_i \hat{\theta}_i}{\sum w_i} \text{ in which } w_i = 1/[\text{Var}(\hat{\theta}) + \hat{\sigma^2}]$$

A Potential Outcomes Framework for Meta-Analysis

- To base the meta-analysis on a solid causal foundation, a Potential Outcome framework can be constructed.

- Assume that there are S trials, indexed by $1, 2, 3, \ldots, s$, and 2 treatments with $Z = 0$ as the control arm and $Z = 1$ as the treatment arm, and the outcome of interest is Y. Then set of the potential outcomes for the individual i will be

$$\vec{Y}_i = \begin{pmatrix} Y_i(Z=0, S=1), & \ldots, & Y_i(Z=0, S=s) \\ Y_i(Z=1, S=1), & \ldots, & Y_i(Z=1, S=s) \end{pmatrix}$$

Identification Assumptions

- A1.**Extended SUTVA:** For all **Z** and **S**, $\vec{Y}_i(\mathbf{Z}, \mathbf{S}) \equiv \vec{Y}_i(Z_i, S_i)$, in which Z_i and S_i are the treatment and trial assignments of individual i.

- A2. **Ignorability within Trials**

$$\vec{Y}_i \perp\!\!\!\perp Z_i, \ \forall i \in \{S_i = s\}$$

  So within each trial, it is a randomized experiments.

- With these two assumptions, apparently we can identify treatment effect inside any trial.

Further Modeling of the Response

- We can write the potential outcome as

$$Y_i(z,s) = \alpha_i + \beta_i(s) + \tau_i(z) + \gamma_i(z,s) + \varepsilon_i(z,s)$$

- Then the unit level treatment effect in study s is

$$Y_i(1,s) - Y_i(0,s) = \tau_i(1) - \tau_i(0) + \gamma_i(1,s) - \gamma_i(0,s) + \varepsilon_i(1,s) - \varepsilon_i(0,s)$$

- Aggregating to the study level, the treatment effect in study s is

$$\theta_s \stackrel{\triangle}{=} E_s[Y_i(1,s) - Y_i(0,s)] = E_s[\tau_i(1) - \tau_i(0)] + E_s[\gamma_i(1,s) - \gamma_i(0,s)]$$

Study-by-Treatment Interaction and Selection into Studies

- From the last formula, we can see two factors contribute to the heterogeneity of study level estimated effect sizes. We can make further Identification Assumptions about them.

Study-by-Treatment Interaction and Selection into Studies

- From the last formula, we can see two factors contribute to the heterogeneity of study level estimated effect sizes. We can make further Identification Assumptions about them.

- A3. **No Study-by-Treatment Interaction**

$$\gamma_i(z, s) = 0$$

- A4. **No Selection into Studies**

$$\vec{Y}_i \perp\!\!\!\perp S_i, \forall i$$

- If A1-A4 all hold, then study level treatment effect θ_s should not depend on study, which means the heterogeneity solely results from sampling errors. Thus we should use Fixed Effect model.

Interpretation of the "Treatment Effect" Estimate from Random Effects model

- If A3 or A4 doesn't hold, then Fixed Effects model is not valid. But how much sense does the Random Effects model estimate make?

- Some justification of Random Effects model is based on exchangeability, Higgins et al. (2009), i.e., we cannot tell a priori which θ_i is larger or smaller. Thus the existence of a distribution G from which, conditioned on some parameters, θ_i's can be seen as iid sample.

- But the pure theoretical nature of this argument also threatens its practical relevance. The mean of distribution G might not has any causal interpretation.

A Selection into Studies Example

- In a hypothetical example, where the entire population can be divided into four sub-populations by age (young/elderly) and gender (male/female). The response is continuous, and the treatment effects by sub-population are given

	Male	Female
Young	1	2
Elderly	4	6

Table: Treatment Effects by Sub-populations.

- We conduct a meta-analysis based on 4 studies. However, the first two studies focused on the treatment effects on elderly, and the second two studies focused on the treatment effects on female.

	Sub-population				
Study	YM	YF	EM	EF	Average Effect
1	0%	0%	40%	60%	5.2
2	0%	0%	60%	40%	4.8
3	0%	40%	0%	60%	4.4
3	0%	60%	0%	40%	3.6

Table: Compositions of sub-population and average effects of the 4 studies

Example Cont'd

- Further, we can assume that the research on elderly is better funded, and thus has larger sample size and smaller standard errors. And the effect sizes of all trials are accurately estimated.

Study	Estimate	Standard Error
1	5.2	0.21
2	4.8	0.18
3	4.4	0.42
4	3.6	0.40

Table: Estimated Effect Sizes and Standard Errors from each Studies

- The pooled treatment effect estimate based on Random Effects model is 4.60, based on Fixed Effects model is 4.79, and based on unweighted average is 4.5.

- But these estimates don't correspond to a treatment effect for any well-defined population. In fact, it is even less informative than the estimates from individual trials.

Individual-Level Data Explaining the Heterogeneity away

- Vioxx Trials Data: Vioxx is a drug that was withdrawn from market because of severe adverse impact on cardiovascular system. We conduct a Individual-Level meta-analysis with data from 30 randomized trials to examine the impact of Vioxx on risk of cardiovascular adverse event.

- For ease of computation, we apply a Poisson regression model to the survival data, which is equivalent to a Cox Proportional Hazard model with constant baseline hazard function.

Individual-Level Data Explaining the Heterogeneity away

- Fitting a random effect model on individual-level data but without individual-level demographic covariates gives a large estimate of random effect for treatment.

```
Random effects:
 Groups Name        Variance Std.Dev. Corr
 trial  (Intercept) 0.402776 0.634646
        treatment1  0.009353 0.096711 -1.000
Number of obs: 17254, groups: trial, 30

Fixed effects:
            Estimate Std. Error z value Pr(>|z|)
(Intercept) -10.1515     0.2075  -48.92  < 2e-16 ***
treatment1    0.4622     0.1246    3.71 0.000209 ***
```

Individual-Level Data Explaining the Heterogeneity away

- However, once the demographic covariates are included, the estimate of random effect becomes zero.

```
Random effects:
 Groups Name        Variance   Std.Dev.   Corr
 trial  (Intercept) 0.0000e+00 0.0000e+00
        treatment1  1.0398e-10 1.0197e-05   NaN
Number of obs: 17254, groups: trial, 30

Fixed effects:
             Estimate Std. Error z value Pr(>|z|)
(Intercept) -15.343832   0.484044  -31.70  < 2e-16 ***
raceblack     0.259301   0.390801    0.66 0.507004
racewhite     0.244987   0.238631    1.03 0.304593
age           0.071338   0.005902   12.09  < 2e-16 ***
genderM       0.415846   0.124012    3.35 0.000799 ***
treatment1    0.431808   0.118046    3.66 0.000254 ***
```

Summary

- Standard practice of meta-analysis lacks a solid causal foundation, although researchers are mostly asking causal questions. Establishing a potential framework for meta-analysis will help us better understand the questions that are being asked and hopefully guide us in finding better answers.
- Predominant methods such as Random Effects model, made popular by DeSimonian and Laird (1986), typically fail to provide us an estimate of causal interpretation.
- Treating meta-analysis naïvely as number crunching is a bad practice that should be discouraged. Instead, we should try to dig into the original studies for possible explanations of heterogeneity. We demonstrate an example where including individual demographic covariates explains heterogeneity away.