

# **OBSERVATIONAL STUDY DESIGN PART III**

## **META-ANALYSIS: EXAMPLE**

Welcome to this third part of the observational study design series. In this module, we will focus on an example of a meta-analysis and specifically, on the interpretation of the results.

## Example

- Thrombolytic therapy in acute myocardial infarction
- Reference: Olkin I (1995), Statistical and theoretical considerations in meta-analysis. *Journal of Clinical Epidemiology*, **48**, 133–146.

2

Next, I would like to present an example meta-analysis to demonstrate the interpretation of the results and to identify potential pitfalls in the meta-analysis process. The example involves clinical trial data from a series of studies investigating the effect of thrombolytic therapy on 30-day mortality among patients with acute myocardial infarction.

## Example

- Consider data from early studies (prior to 1975):

	author	year	event.e	n.e	event.c	n.c
1	Fletcher	1959	1	12	4	11
2	Dewar	1963	4	21	7	21
3	Lippschutz	1965	6	43	7	41
4	European 1	1969	20	83	15	84
5	European 2	1971	69	373	94	357
6	Heikinheimo	1971	22	219	17	207
7	Italian	1971	19	164	18	157
8	Australian 1	1973	26	264	32	253
9	Frankfurt 2	1973	13	102	29	104
10	Gormsen	1973	2	14	3	14
11	NHLBI SMIT	1974	7	53	3	54

3

We will focus on the analysis of data from 11 studies reported in the cited paper. The first two columns provide information about the study, including the first author and year of publication. The next two columns summarize the number of events, meaning, the number of deaths within 30 days, among the total number (n) patients treated in the experimental arm. The last two columns provide similar information (number of events and total number treated) for the control arm. Across the studies we see that the sample sizes are quite variable, ranging from approximately 10 per arm to roughly 350 per arm.

## Modeling probabilities

- Denote the probability of an event by  $p$ , where  $p$  ranges from 0 to 1.

- Notation:

$$\begin{aligned} p &= \text{probability of an event} \\ &= \text{prob}(y=1) \end{aligned}$$

$$\begin{aligned} 1-p &= \text{probability of no event} \\ &= \text{prob}(y=0) \end{aligned}$$

4

Before we derive a pooled estimate of the intervention effect on 30-day mortality, let's review the calculation of an odds ratio.

We will let  $p$  denote the probability of an event, in this case, death within 30 days of an MI, and then  $(1-p)$  equals the probability of not observing the event of interest, i.e., remaining alive at 30 days following the MI.

## Modeling probabilities

- Notation continued:

$$\frac{p}{1-p} = \text{odds}$$
$$\frac{\left(\frac{p_1}{1-p_1}\right)}{\left(\frac{p_2}{1-p_2}\right)} = \frac{p_1(1-p_2)}{(1-p_1)p_2} = \text{odds ratio}$$

5

The odds of an event can then be calculated as the probability of the event divided by the probability of the event not occurring. The odds ratio is then the ratio of the odds of death for those in the intervention group divided by the odds of death for the control group. Values greater than one indicate a higher odds of death under the experimental arm while values less than one indicate a lower odds of death under the experimental arm relative to the control arm.

## Modeling probabilities

- Example: 1959 study
  - Experimental arm:
    - The probability of 30-day mortality is  $1/12=0.083$
    - The odds of 30-day mortality are  $0.083/(1-0.083) = 0.091$
  - Control arm:
    - The probability of 30-day mortality is  $4/11=0.36$
    - The odds of 30-day mortality are  $0.36/(1-0.36) = 0.57$
  - The odds ratio for 30-day mortality for patients treated with the intervention relative to control is  $0.091/0.57 = 0.16$
  - Interpretation: The odds of 30-day mortality are 84% lower for the intervention relative to the control group.

6

Let's consider data from a specific study, the 1959 trial where 1 of 12 experimental patients died within 30 days and 4 of 11 control patients died within 30 days. The probability of death within 30 days is  $1/12 = 0.083$  in the experimental arm, resulting in an odds of death of 0.091. Similarly, the probability of death under the control therapy is 0.36, with an odds of death of 0.57. The odds ratio is then 0.16. This value can be interpreted to mean that the odds of death for the experimental group are 16% of the odds of death in the control arm, or, in other words, when comparing the odds ratio to a value of 1, the odds of death are reduced by 84% among the intervention patients compared to the control patients.

## Example

	OR	95%-CI	%W(fixed)	%W(random)
Fletcher 1959	0.1591	[0.0146; 1.7318]	1.94	1.55
Dewar 1963	0.4706	[0.1140; 1.9422]	2.88	4.01
Lippschutz 1965	0.7876	[0.2407; 2.5779]	3.13	5.42
European 1 1969	1.4603	[0.6887; 3.0963]	5.75	10.63
European 2 1971	0.6350	[0.4467; 0.9028]	39.76	21.45
Heikinheimo 1971	1.2481	[0.6428; 2.4235]	7.98	12.39
Italian 1971	1.0119	[0.5099; 2.0081]	8.26	11.92
Australian 1 1973	0.7545	[0.4358; 1.3063]	14.96	15.21
Frankfurt 2 1973	0.3778	[0.1834; 0.7782]	12.72	11.17
Gormsen 1973	0.6111	[0.0854; 4.3706]	1.31	2.22
NHLBI SMIT 1974	2.5870	[0.6316; 10.5961]	1.31	4.04

Number of trials combined: 11

7

Before estimating a pooled intervention effect, we will first review the estimated odds ratio and confidence interval for each individual study. Recall that odds ratio values less than one indicate that the odds of death are reduced under the experimental agent compared to the control while values greater than one indicate that the odds of death are increased under the experimental agent compared to the control. When reviewing the odds ratio values, we see that most studies resulted in an odds ratio less than one, indicating that the odds of death are lower under the experimental agent. Next, we will consider the confidence interval for the odds ratio. Recall that if the 95% confidence interval lies entirely above one or below one, the intervention effect is statistically significant at a two-sided 0.05 alpha level; however, if the interval includes a value of 1, the intervention effect is not statistically significant. We see that the confidence intervals are quite wide, particularly for small studies and that the reduction in the odds of death is statistically significant for only two of the 11 studies, the European 2 study and the Frankfurt 2 study. Finally, we will view the weighting that is applied through the fixed effects and random effects analysis methods.

## Example

	OR	95%-CI	%W(fixed)	%W(random)
Fletcher 1959	0.1591	[0.0146; 1.7318]	1.94	1.55
Dewar 1963	0.4706	[0.1140; 1.9422]	2.88	4.01
Lippschutz 1965	0.7876	[0.2407; 2.5779]	3.13	5.42
European 1 1969	1.4603	[0.6887; 3.0963]	5.75	10.63
European 2 1971	0.6350	[0.4467; 0.9028]	39.76	21.45
Heikinheimo 1971	1.2481	[0.6428; 2.4235]	7.98	12.39
Italian 1971	1.0119	[0.5099; 2.0081]	8.26	11.92
Australian 1 1973	0.7545	[0.4358; 1.3063]	14.96	15.21
Frankfurt 2 1973	0.3778	[0.1834; 0.7782]	12.72	11.17
Gormsen 1973	0.6111	[0.0854; 4.3706]	1.31	2.22
NHLBI SMIT 1974	2.5870	[0.6316; 10.5961]	1.31	4.04

Number of trials combined: 11

8

The fixed effects weights are primarily driven by sample size; smaller studies will have lower weights.

## Example

	OR		95%-CI	%W(fixed)	%W(random)
Fletcher 1959	0.1591	[0.0146;	1.7318]	1.94	1.55
Dewar 1963	0.4706	[0.1140;	1.9422]	2.88	4.01
Lippschutz 1965	0.7876	[0.2407;	2.5779]	3.13	5.42
European 1 1969	1.4603	[0.6887;	3.0963]	5.75	10.63
European 2 1971	0.6350	[0.4467;	0.9028]	39.76	21.45
Heikinheimo 1971	1.2481	[0.6428;	2.4235]	7.98	12.39
Italian 1971	1.0119	[0.5099;	2.0081]	8.26	11.92
Australian 1 1973	0.7545	[0.4358;	1.3063]	14.96	15.21
Frankfurt 2 1973	0.3778	[0.1834;	0.7782]	12.72	11.17
Gormsen 1973	0.6111	[0.0854;	4.3706]	1.31	2.22
NHLBI SMIT 1974	2.5870	[0.6316;	10.5961]	1.31	4.04

Number of trials combined: 11

9

The random effects weights are less extreme and reflect study-to-study variation in the effect estimate.

## Example

	OR	95%-CI	%W(fixed)	%W(random)
Fletcher 1959	0.1591	[0.0146; 1.7318]	1.94	1.55
Dewar 1963	0.4706	[0.1140; 1.9422]	2.88	4.01
Lippschutz 1965	0.7876	[0.2407; 2.5779]	3.13	5.42
European 1 1969	1.4603	[0.6887; 3.0963]	5.75	10.63
European 2 1971	0.6350	[0.4467; 0.9028]	39.76	21.45
Heikinheimo 1971	1.2481	[0.6428; 2.4235]	7.98	12.39
Italian 1971	1.0119	[0.5099; 2.0081]	8.26	11.92
Australian 1 1973	0.7545	[0.4358; 1.3063]	14.96	15.21
Frankfurt 2 1973	0.3778	[0.1834; 0.7782]	12.72	11.17
Gormsen 1973	0.6111	[0.0854; 4.3706]	1.31	2.22
NHLBI SMIT 1974	2.5870	[0.6316; 10.5961]	1.31	4.04

Number of trials combined: 11

10

We see that the study that receives the largest weight is the European 2 study. This study had the largest sample size.

## Example

	OR	95%-CI	z	p.value
Fixed effect model	0.7637	[0.6182; 0.9435]	-2.4986	0.0125
Random effects model	0.7994	[0.5894; 1.0842]	-1.4400	0.1499

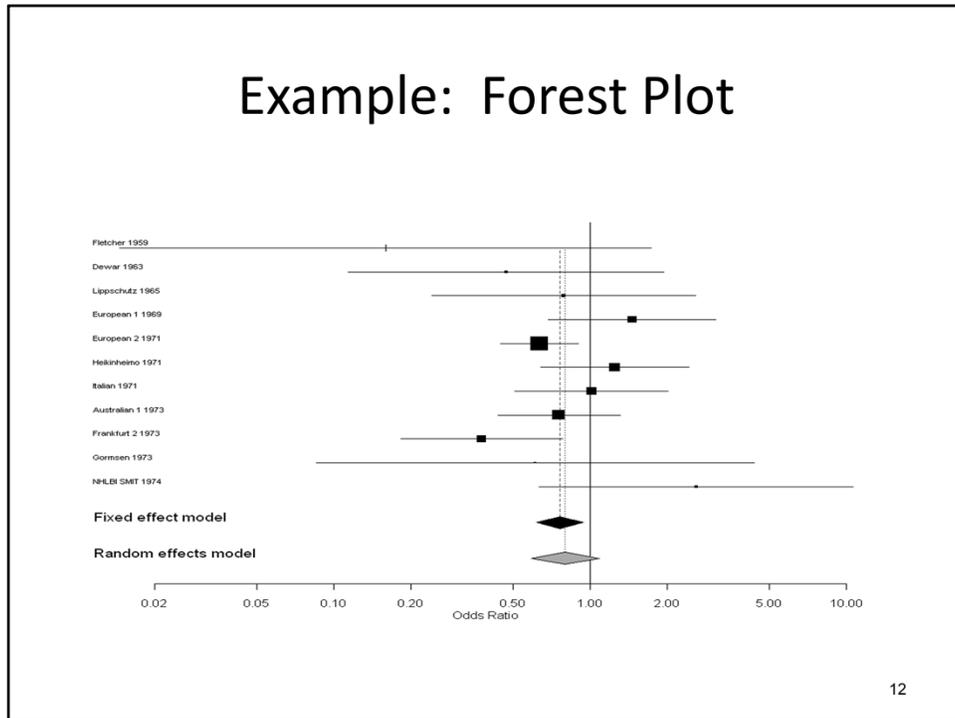
Quantifying heterogeneity:  
 $\tau^2 = 0.0805$ ;  $H = 1.24$  [1; 1.77];  $I^2 = 34.8\%$  [0%; 68%]

Test of heterogeneity:  
Q d.f. p.value  
15.35 10 0.1199

11

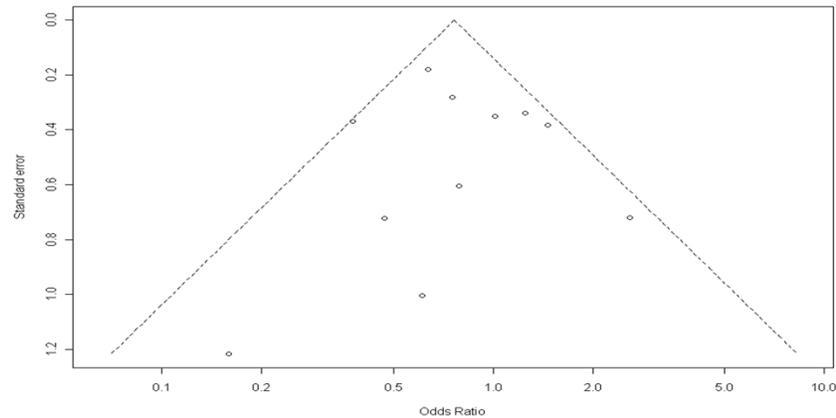
After viewing the individual, study-specific estimates, we will derive the pooled estimate using both a fixed and a random effects modeling approach. We see that regardless of the estimation method, the estimated odds ratio value is fairly stable, with roughly a 20% reduction in the odds of death associated with 30-day mortality. As expected, the confidence interval under the random effects method is wider. In this case, the estimated odds ratio under the fixed effects method is statistically significant while the estimated odds ratio under the random effects model is not statistically significant. To determine which estimates should be reported, we will then review the statistics related to tests of heterogeneity. In this case, the  $I^2$  index indicates that 35% of the variation in estimated treatment effects is due to between-study variation. The p-value for the Q statistic is very close to the alpha level of 0.10, which is typically used to define statistically significant heterogeneity. In this case, given the variability in odds ratio values that is seen on the previous slide and the results of the  $I^2$  index and Q statistic, I would report the results from the random effects model and conclude that the odds of death are estimated to be 20% lower under the experimental agent but this reduction is not statistically significant ( $p=0.15$ ). The confidence interval ranges from a 41% reduction in the odds to an 8% increase in the odds of death associated with experimental treatment.

## Example: Forest Plot



A Forest plot is typically presented in a meta-analysis to graphically display the study-specific point estimates for the odds ratio values and the corresponding confidence interval. The size of the box corresponds to the sample size and the width of the horizontal lines corresponds to the confidence interval. The pooled estimates can also be displayed, where the center of the diamond corresponds to the point estimate and the width of the diamond corresponds to the confidence interval around the point estimate.

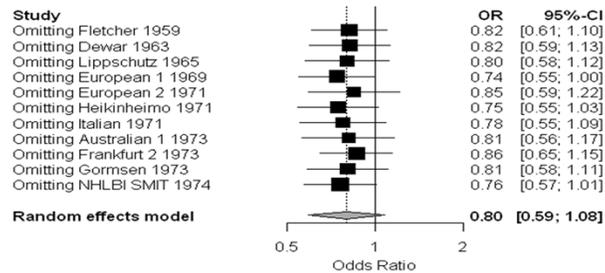
## Example: Funnel Plot



13

A funnel plot can be created to explore the potential for publication bias. To orient you to the figure, the horizontal axis reflects the odds ratio. The vertical axis reflects standard error of the estimated odds ratio where smaller values are at the top of the figure. Note that smaller standard error values will typically arise when the sample size is larger. If we have complete reporting of all studies, we should see that the funnel is filled in for all areas. When considering small studies with high standard error values, we should see a wide range of odds ratio values above and below one. As the sample size increases and the standard error decreases, moving to the peak of the funnel, we should see estimates on either side of the peak. If an area of the funnel is empty, there is an indication of potential publication bias. In this case, we did not observe small studies with large standard error values coupled with odds ratio values greater than 1. It is likely that if a small study was conducted that suggested increased odds of death under the investigational agent, the results were not published.

## Example: influence analysis



An influence analysis can be performed whereby each individual study is deleted from the set of studies and the pooled estimate is recalculated. In this example dataset, individual studies do not have a major influence on the point estimates.

## Summary

- Limitations
  - Information that can be retrieved
  - Publication bias
  - Differential reporting
- Inference is dependent on the quality of the systematic review and meta-analysis

In summary, be aware of the difficulties in implementing and reporting systematic reviews and meta-analyses. The primary limitation is that the systematic review and meta-analysis can only reflect data that are available. Publication bias and differential or variable reporting are important concerns. In this series we have discussed methodological approaches to ensure that the review is comprehensive and exhaustive and that bias is minimized. When reading or conducting systematic reviews and meta-analyses, it is important to keep in mind these methodological guidelines. There are many assumptions and steps that were made in order to derive the overall pooled estimate that is reported.