PSYC30013 Research Methods for Human Inquiry
Week 11 (Day 1): Meta-Analysis

A typical research publication

The standard research publication in a psychology journal usually contains the following features:

1. Theoretical basis of the research.

2. _Narrative review_ of the relevant literature.

3. Rationale for undertaking the research.

4. Set of research questions (or hypotheses).

5. Description of the participant recruitment methods.

6. Explanation of the research design and measurement methods.

7. Justification of the statistical analysis to be used.

8. Summary of the findings.

9. Discussion of findings in relation to theory, previous literature, the rationale for the research, and limitations.

What's a narrative review?

Borenstein et al. (2021) explain:

"An expert in a given field would read the studies that addressed a question, summarize the findings, and then arrive at a conclusion – for example, that the treatment in question was, or was not, effective."

Key features:
- Expert subjectively decides which studies should be included in the review, and how they should be weighted
- Strategy of search for studies generally not disclosed
- Focus on qualitative description of prior results, rather than on quantification.
- Expert subjectively researches a conclusion about the overall trend of prior results

Problems with narrative reviews

Shadish, Cook, & Campbell (2002) list four problems with narrative reviews:

1. The greater the number of studies under review, the greater the difficulty in forming a coherent narrative.

2. They encourage a comparison between the number of statistically significant vs non-significant studies ('vote counting').

3. Narrative reviews tend to describe results in an imprecise way

4. Narrative reviews struggle to account for more studies involving complex relationships between variables.

What is a systematic review?

Cochrane Handbook (2023): "A systematic review attempts to collate all the empirical evidence that fits pre-specified eligibility criteria in order to answer a specific research question"

Systematic reviews adopt transparent and predefined

– Inclusion/exclusion criteria

– Search strategies

– Coding of studies

Most systematic reviews include *statistical* synthesis of past research (i.e. they include a meta-analysis)

What is meta-analysis?

Glass (1976): "[T]he statistical analysis of a large collection of analysis results from individual studies for the purpose of integrating the findings."

Cochrane Handbook (2023): "Meta-analysis is the statistical combination of results from two or more separate studies."

Unlike systematic review, meta-analysis is inherently statistical. Typically, the key statistic will be a summary estimate of the true effect of some variable.

- Example: a meta-analysis combines 50 studies on drug X, and finds that:
  - Heart attack patients randomly assigned to take drug X live on average 70 days longer than patients randomly assigned to take a placebo. 95% Confidence Interval: 60-80 days.

Meta-analysis and systematic review

They often go together. Lots of papers have titles like:

– Van Nostrand, M. (2024). Uromycitisis poisoning: a systematic review and meta-analysis.

To do a meta-analysis, you pretty much need to do a systematic review first

Not true in reverse - people often do systematic reviews and don't go on to a meta-analysis

Meta-analysis was initially controversial

In Part III I briefly examine the issue of the relevance of statistical inference, of whatever kind, to research in social sciences. I devote a section to a new *bête noire*, meta-analysis. It should be stifled at birth.

AMERICAN PSYCHOLOGIST • MAY 1978

H. J. EYSENCK
Institute of Psychiatry
London, England

American Journal of
EPIDEMIOLOGY

Meta-analysis/Shmeta-analysis

Samuel Shapiro

An Exercise in Mega-Silliness

The article by Smith and Glass (September 1977) begins promisingly by referring to my "tendentious diatribes" (p. 752) on the outcome problem in psychotherapy, inviting

Meta-analysis is alive and well

Google "meta-analysis"

Web   Books   Images   Videos   Apps   More ▾   Search

About 15,800,000 results (0.24 seconds)

Meta-analysis - Wikipedia, the free encyclopedia
en.wikipedia.org/wiki/Meta-analysis ▾ Wikipedia ▾
In statistics, meta-analysis comprises statistical methods for contrasting and
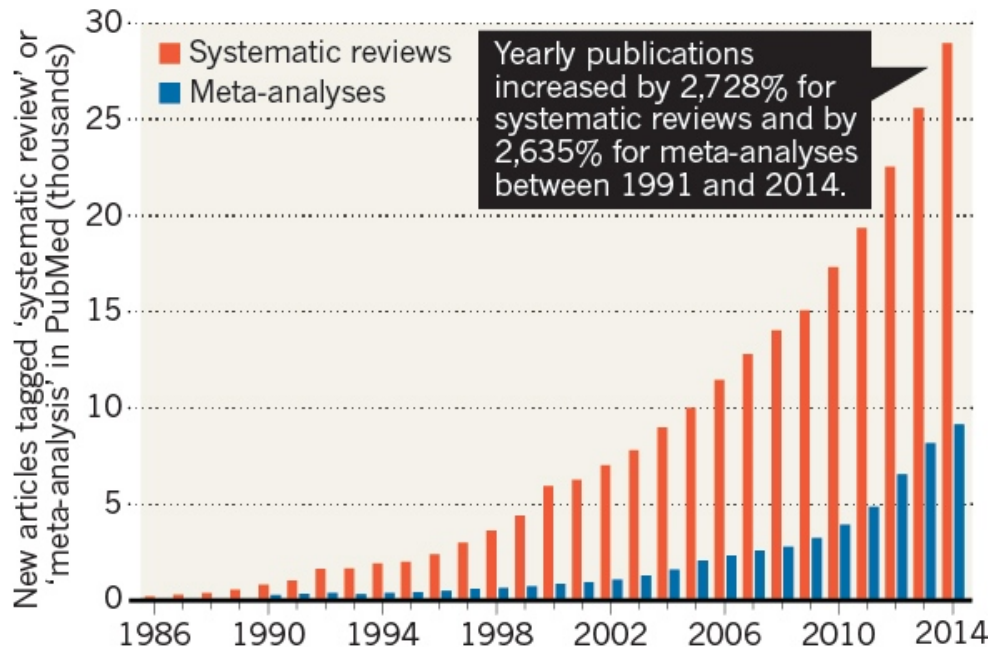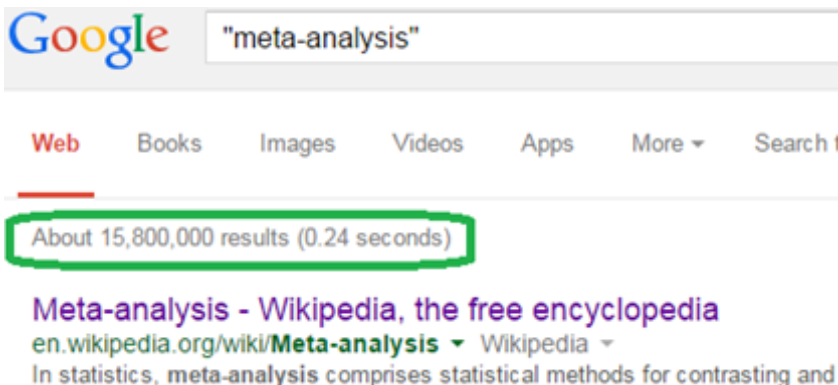
MANY META-ANALYSIS STUDIES INCLUDE THE PHRASE "WE SEARCHED MEDLINE, EMBASE, AND COCHRANE FOR STUDIES..."

THIS HAS LED TO META-META-ANALYSES COMPARING META-ANALYSIS METHODS.
e.g. M SAMPSON (2003), PL ROYLE (2005) E LEE (2011), AR LEMESHOW (2005)

WE PERFORMED A META-META-META-ANALYSIS OF THESE META-META-ANALYSES.

METHODS: WE SEARCHED MEDLINE, EMBASE, AND COCHRANE FOR THE PHRASE "WE SEARCHED MEDLINE, EMBASE, AND COCHRANE FOR THE PHRASE "WE SEARCHED MEDLINE EMBASE AND

LIFE GOAL #28: GET A PAPER REJECTED WITH THE COMMENT "TOO META"
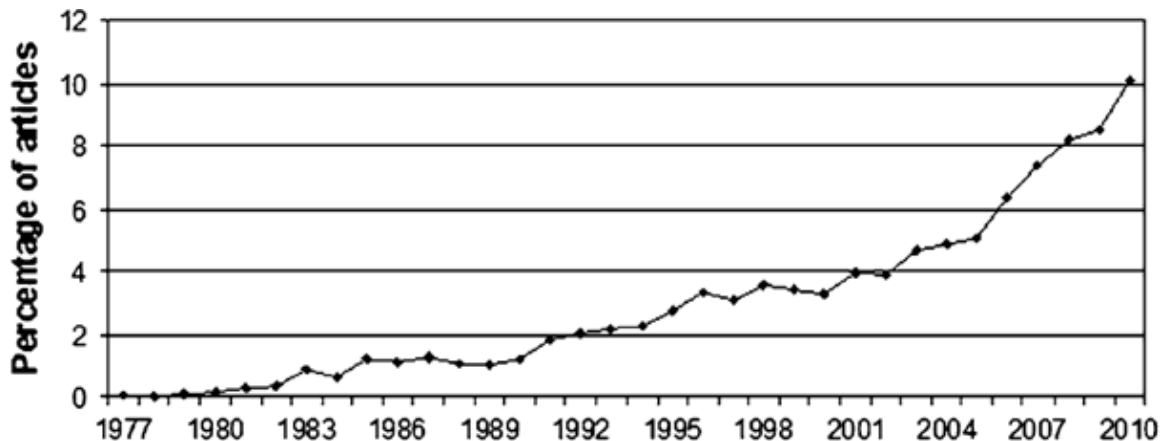
Figure from Guilera et al's (2013) study of psychology journals.

9

Forest plot (blobbogram)

Study name and sample size

Effect size from Dr Lucky's study

Lucky (Total $N = 44$)

Noluck (Total $N = 36$)

MA combination

−2   0   2   4   6   8

Difference between the means
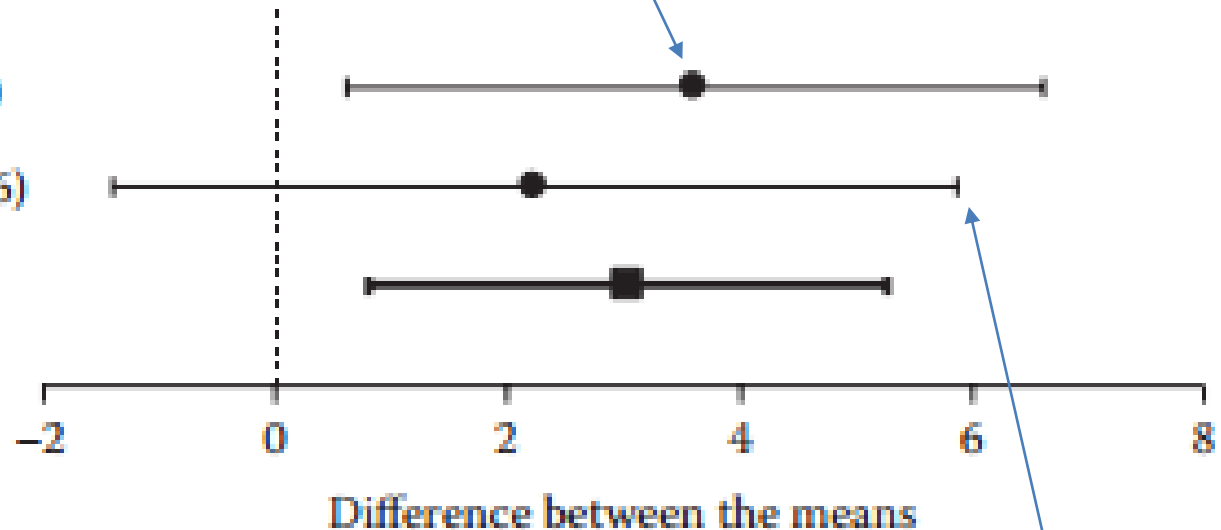
Meta-analysis combination, also known as "the pooled effect"

Lines represent 95% Confidence Intervals

Example: combining two studies in which *p* = .10

What will the meta-analysis CI look like?

Example: combining two studies in which *p* = .10

Study 1

Study 2

MA combination

−2   0   2   4   6   8

Difference between the means
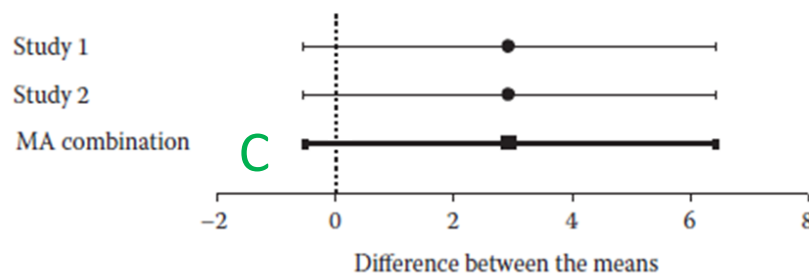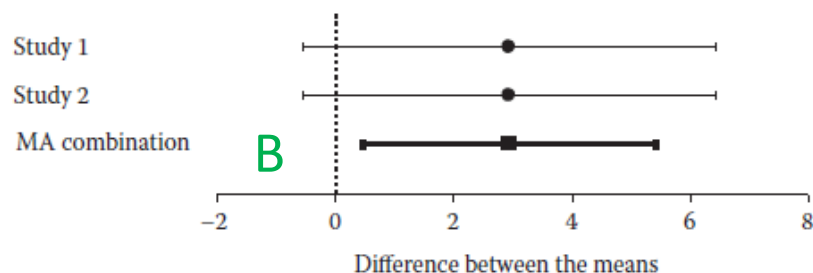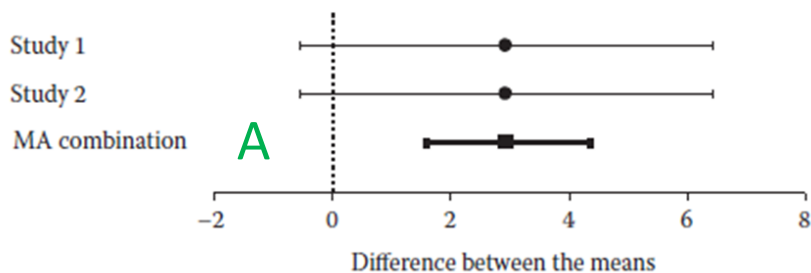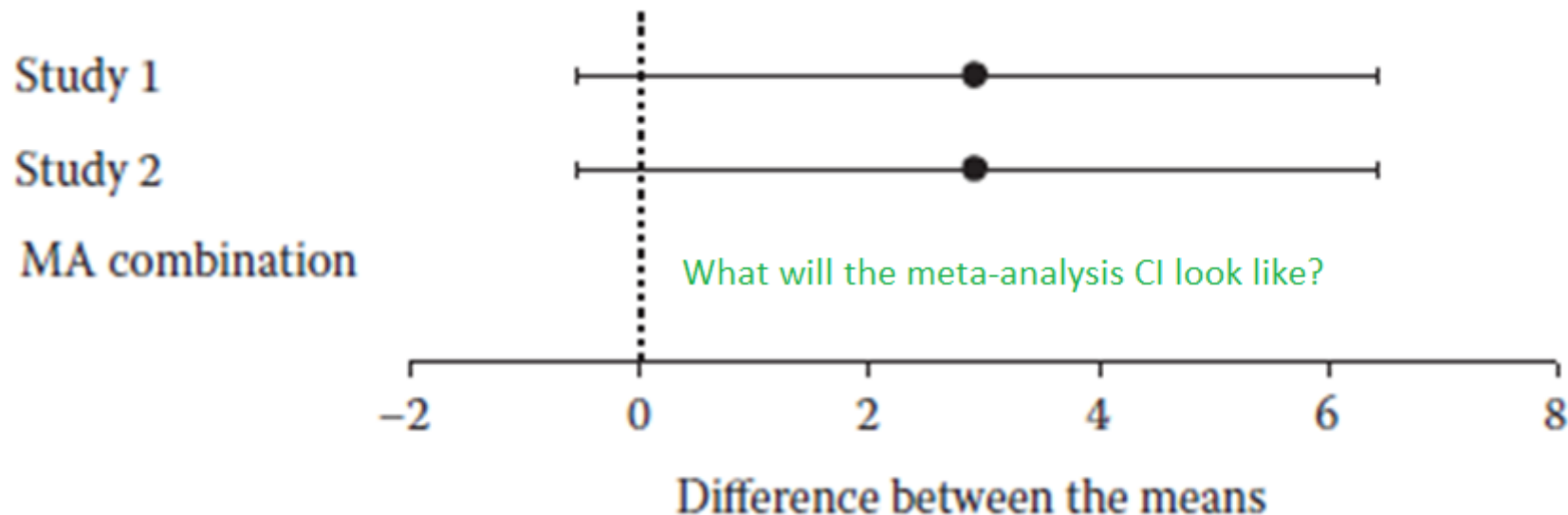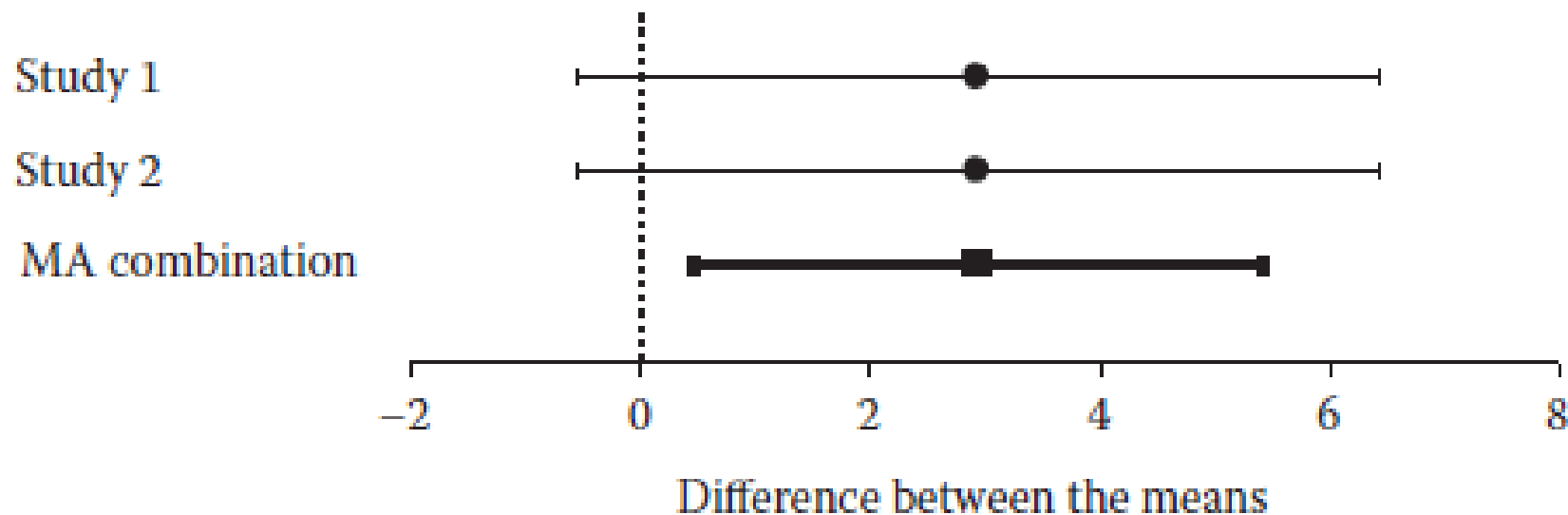
The meta-analysis confidence interval is about 30% shorter than those of the two original studies

- Exact reduction in length depends on the sample sizes of the original studies, and on the similarity of the confidence intervals in the original studies.

Risk Ratio (also known as "Relative Risk")

	Outcome	No outcome	N
Exposed	A	B	$n_1 = A + B$
Unexposed	C	D	$n_2 = C + D$

$$Risk\ Ratio = \frac{\left(A/n_1\right)}{\left(C/n_2\right)}$$

Risk ratios tell us about the strength of association between an exposure and an outcome. To get the risk ratio, take the risk (i.e. probability) of the outcome in the exposed group divided by the risk of the outcome in the unexposed group.

Risk Ratio (also known as "Relative Risk")

	Infected	Not infected	N
Treatment	A = 20	B = 20	n_1 = 40
Control	C = 25	D = 15	n_2 = 40

$$Risk\ Ratio = \frac{(A/n_1)}{(C/n_2)} = \frac{(20/40)}{(25/40)} = 0.8$$

The risk of infection was 62.5% (25/40) in the control group
The risk of infection was 50.0% (20/40) in the treatment group

People taking the treatment had 0.8 times the risk of infection of those in the control group.

Why vote counting is bad

Study	Risk Ratio and 95% Confidence Interval	p-value	N	Risk Ratio [95% CI]
Study 1		0.265	80	0.80 [0.54, 1.18]
Study 2		0.265	80	0.80 [0.54, 1.18]
Study 3		0.265	80	0.80 [0.54, 1.18]
Study 4		0.265	80	0.80 [0.54, 1.18]
Study 5		0.265	80	0.80 [0.54, 1.18]
Meta-Analysis		0.013	400	0.80 [0.67, 0.95]

0.5   1   2

Favours Treatment   Favours Control
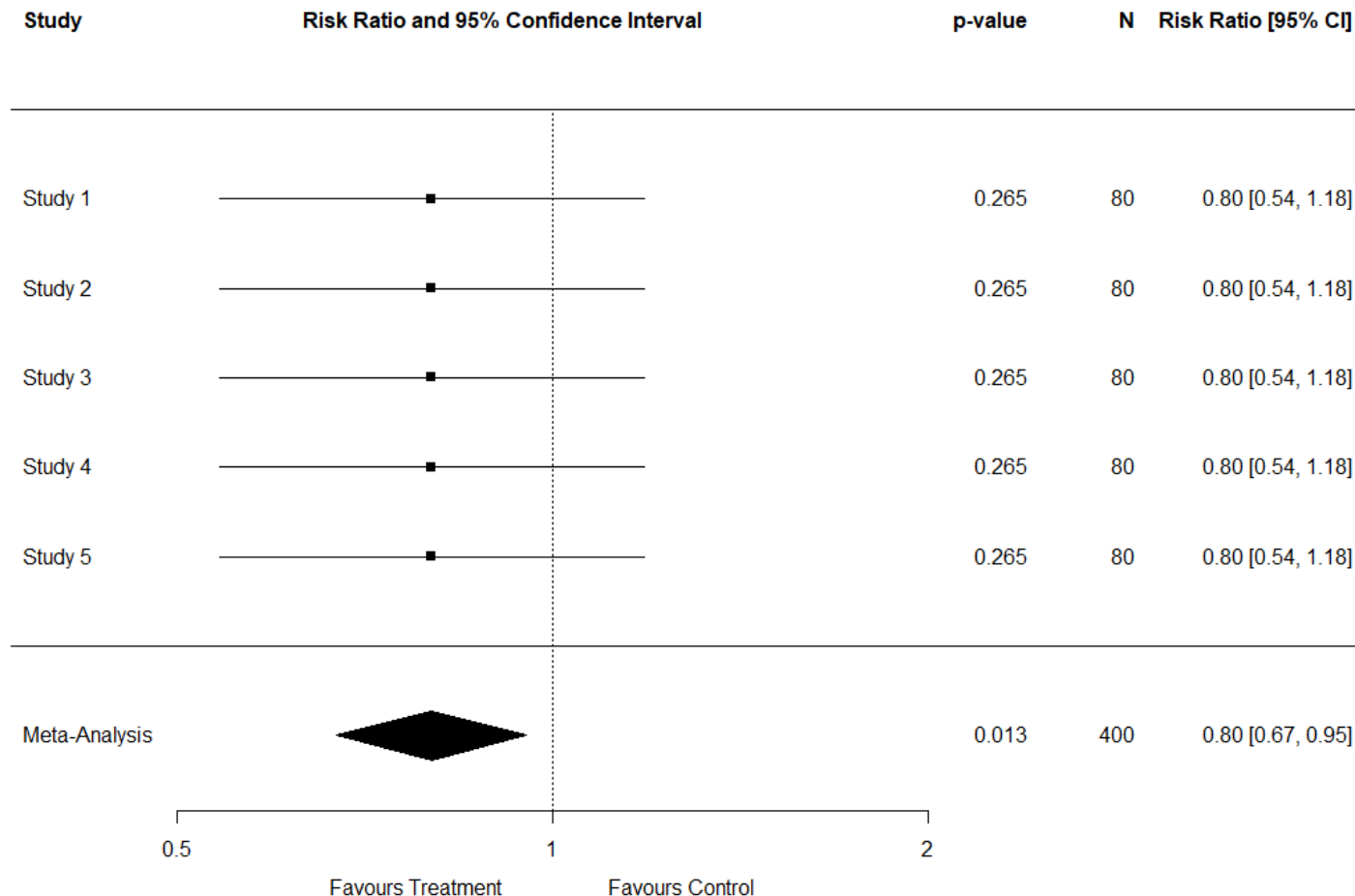
According to vote counting: "All five previous studies failed to achieve a statistically significant result."

According to meta-analysis: "The summary p-value is statistically significant, *p* = .013. The estimate of the effect size is…"

15

Streptokinase studies

Study	Year	# of patients	p-value < .05?
Fletcher	1959	23	NO
Dewar	1963	42	NO
European 1	1969	167	NO
European 2	1971	730	YES
Heikinheimo	1971	426	NO
Italian	1971	321	NO
Australian 1	1973	517	NO
Frankfurt 2	1973	206	YES
NHLBI SMIT	1974	107	NO
Frank	1975	108	NO
Valere	1975	91	NO
Klein	1976	23	NO
UK Collab	1976	595	NO
Austrian	1977	728	YES
Australian 2	1977	230	NO
Lasierra	1977	24	NO
N Ger Collab	1977	483	NO
Witchitz	1977	58	NO
European 3	1979	315	NO
ISAM	1986	1,741	NO
GISSI-1	1986	11,712	YES
Olson	1986	52	NO
Baroffio	1986	59	NO
Schreiber	1986	38	NO
Criber	1986	44	NO
Sainsous	1986	98	NO
Durand	1987	64	NO
White	1987	219	YES
Bassand	1987	107	NO
Vlay	1988	25	NO
Kennedy	1988	368	NO
ISIS-2	1988	17,187	YES
Wisenberg	1988	66	NO
TOTAL		36,974	6 vs 27

- Streptokinase is a medicine that dissolves blood clots

- Particularly useful after myocardial infarction (heart attack) has occurred

- Extremely lengthy period of clinical trials from late 1950s to late 1980s.

Streptokinase meta-analysis

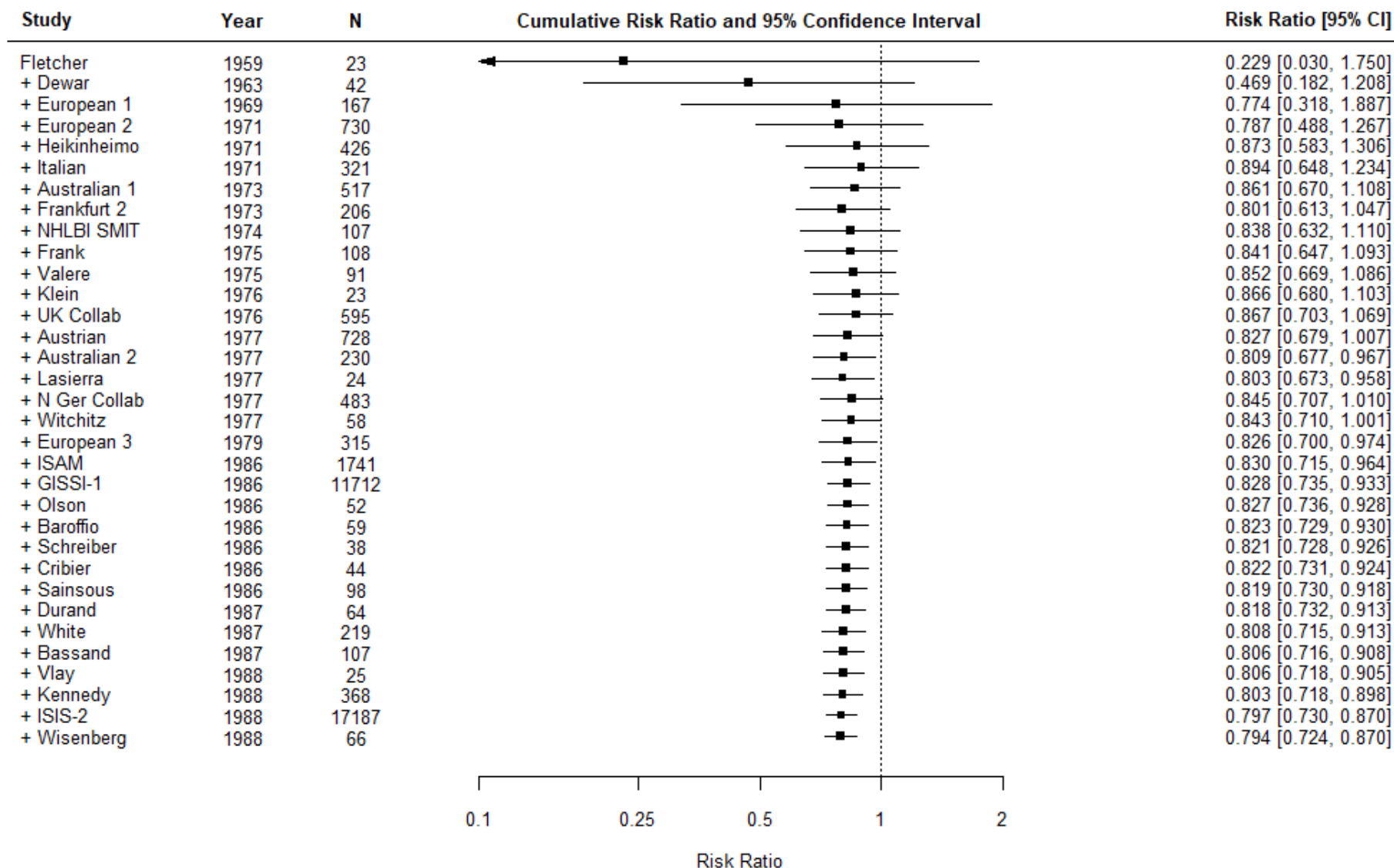| Study	Year	N	p-value		Risk Ratio [95% CI]
Fletcher	1959	23	0.155		0.229 [0.030, 1.750]
Dewar	1963	42	0.305		0.571 [0.196, 1.665]
European 1	1969	167	0.325		1.349 [0.743, 2.451]
European 2	1971	730	0.012		0.703 [0.534, 0.925]
Heikinheimo	1971	426	0.513		1.223 [0.669, 2.237]
Italian	1971	321	0.973		1.011 [0.551, 1.853]
Australian 1	1973	517	0.315		0.779 [0.478, 1.268]
Frankfurt 2	1973	206	0.01		0.457 [0.252, 0.828]
NHLBI SMIT	1974	107	0.191		2.377 [0.649, 8.709]
Frank	1975	108	0.946		0.964 [0.332, 2.801]
Valere	1975	91	0.907		1.048 [0.481, 2.282]
Klein	1976	23	0.361		2.571 [0.339, 19.481]
UK Collab	1976	595	0.699		0.922 [0.609, 1.394]
Austrian	1977	728	0.01		0.608 [0.417, 0.886]
Australian 2	1977	230	0.13		0.702 [0.443, 1.110]
Lasierra	1977	24	0.241		0.282 [0.034, 2.340]
N Ger Collab	1977	483	0.366		1.161 [0.840, 1.604]
Witchitz	1977	58	0.718		0.812 [0.263, 2.506]
European 3	1979	315	0.075		0.612 [0.356, 1.050]
ISAM	1986	1741	0.476		0.880 [0.619, 1.250]
GISSI-1	1986	11712	0		0.827 [0.749, 0.914]
Olson	1986	52	0.477		0.429 [0.041, 4.439]
Baroffio	1986	59	0.08		0.079 [0.005, 1.350]
Schreiber	1986	38	0.322		0.333 [0.038, 2.925]
Cribier	1986	44	0.948		1.095 [0.073, 16.427]
Sainsous	1986	98	0.306		0.500 [0.132, 1.887]
Durand	1987	64	0.51		0.621 [0.151, 2.555]
White	1987	219	0.02		0.174 [0.040, 0.761]
Bassand	1987	107	0.398		0.604 [0.188, 1.944]
Vlay	1988	25	0.504		0.462 [0.048, 4.461]
Kennedy	1988	368	0.241		0.654 [0.322, 1.331]
ISIS-2	1988	17187	0		0.769 [0.704, 0.839]
Wisenberg	1988	66	0.077		0.244 [0.051, 1.164]
Meta-Analysis			0.0000008		0.794 [0.724, 0.870]

Risk Ratio and 95% Confidence Interval

0.01   0.1   1   10   100

Favours Treatment   Favours Control

17

Streptokinase cumulative meta-analysis

Study	Year	N	Cumulative Risk Ratio and 95% Confidence Interval	Risk Ratio [95% CI]
Fletcher	1959	23		0.229 [0.030, 1.750]
+ Dewar	1963	42		0.469 [0.182, 1.208]
+ European 1	1969	167		0.774 [0.318, 1.887]
+ European 2	1971	730		0.787 [0.488, 1.267]
+ Heikinheimo	1971	426		0.873 [0.583, 1.306]
+ Italian	1971	321		0.894 [0.648, 1.234]
+ Australian 1	1973	517		0.861 [0.670, 1.108]
+ Frankfurt 2	1973	206		0.801 [0.613, 1.047]
+ NHLBI SMIT	1974	107		0.838 [0.632, 1.110]
+ Frank	1975	108		0.841 [0.647, 1.093]
+ Valere	1975	91		0.852 [0.669, 1.086]
+ Klein	1976	23		0.866 [0.680, 1.103]
+ UK Collab	1976	595		0.867 [0.703, 1.069]
+ Austrian	1977	728		0.827 [0.679, 1.007]
+ Australian 2	1977	230		0.809 [0.677, 0.967]
+ Lasierra	1977	24		0.803 [0.673, 0.958]
+ N Ger Collab	1977	483		0.845 [0.707, 1.010]
+ Witchitz	1977	58		0.843 [0.710, 1.001]
+ European 3	1979	315		0.826 [0.700, 0.974]
+ ISAM	1986	1741		0.830 [0.715, 0.964]
+ GISSI-1	1986	11712		0.828 [0.735, 0.933]
+ Olson	1986	52		0.827 [0.736, 0.928]
+ Baroffio	1986	59		0.823 [0.729, 0.930]
+ Schreiber	1986	38		0.821 [0.728, 0.926]
+ Cribier	1986	44		0.822 [0.731, 0.924]
+ Sainsous	1986	98		0.819 [0.730, 0.918]
+ Durand	1987	64		0.818 [0.732, 0.913]
+ White	1987	219		0.808 [0.715, 0.913]
+ Bassand	1987	107		0.806 [0.716, 0.908]
+ Vlay	1988	25		0.806 [0.718, 0.905]
+ Kennedy	1988	368		0.803 [0.718, 0.898]
+ ISIS-2	1988	17187		0.797 [0.730, 0.870]
+ Wisenberg	1988	66		0.794 [0.724, 0.870]

Risk Ratio: 0.1 0.25 0.5 1 2

Risk Ratio

# Meta-analysis and Sudden Infant Death Syndrome

This well-known book contained advice now believed to be incorrect

- SIDS refers to the sudden unexplained death of a child < 1 year old

- 100,000+ deaths from 1950s-1990s

- Gilbert et al. (2005) write that

  - "Advice to put infants to sleep on the front for nearly half a century was contrary to evidence available from 1970 that this was likely to be harmful. Systematic review of preventable risk factors for SIDS from 1970 would have led to earlier recognition of the risks of sleeping on the front and might have prevented over 10,000 infant deaths in the UK and at least 50,000 in the Europe, the USA and Australasia."

Note: nothing in this lecture should be taken as medical advice. If any of the lecture examples become relevant to you in real life, seek advice from health professionals with relevant expertise

SIDS Rate and Back Sleeping
(1988 – 2006)
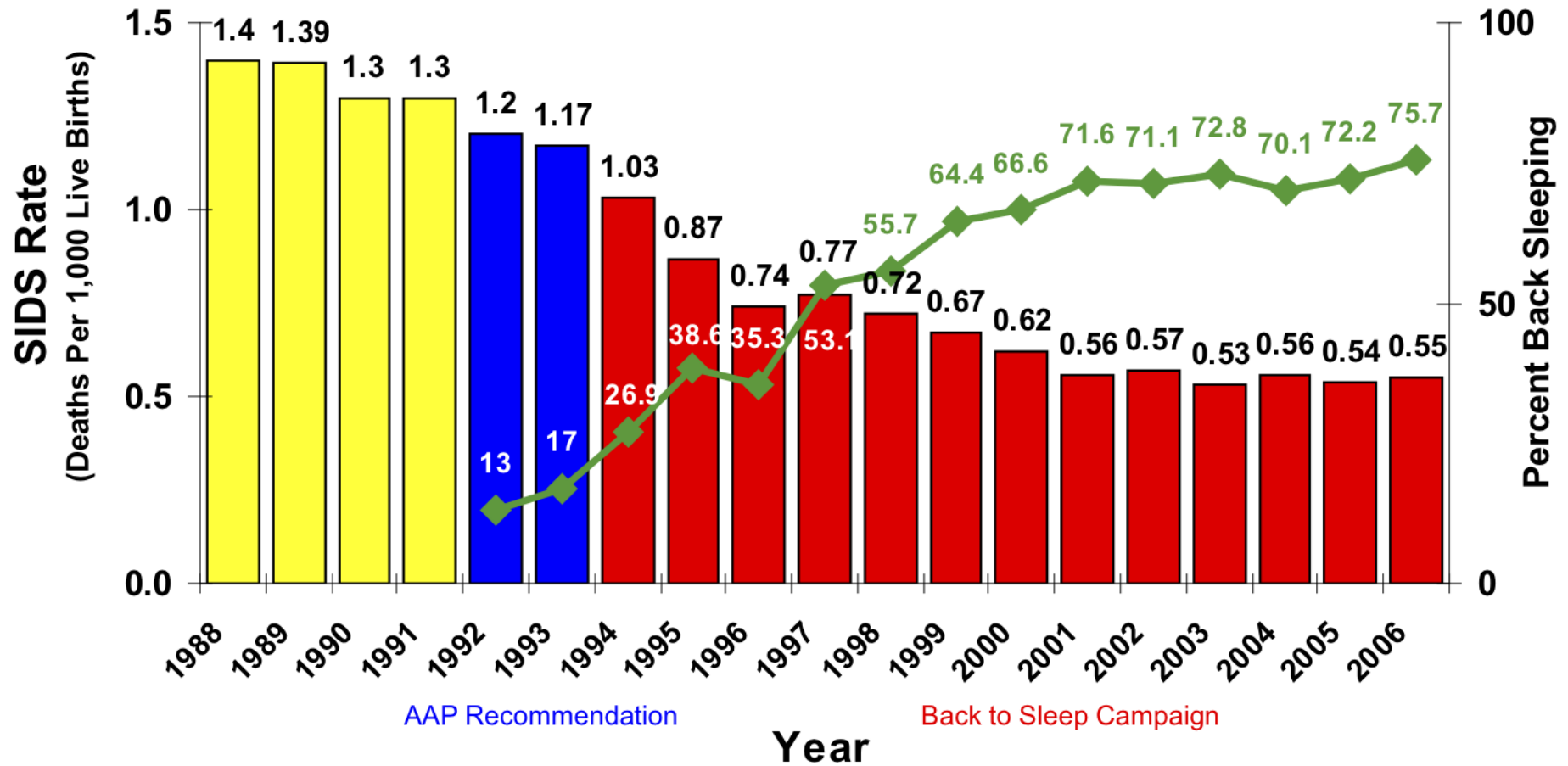
SIDS Rate Source: CDC, National Center for Health Statistics,
Sleep Position Data: NICHD, National Infant Sleep Position Study.

Criticisms of meta-analysis

1. The file drawer problem invalidates meta-analysis

2. One number cannot summarize a research field

3. Mixing apples and oranges

4. Garbage in, garbage out

5. Meta-analyses are done poorly
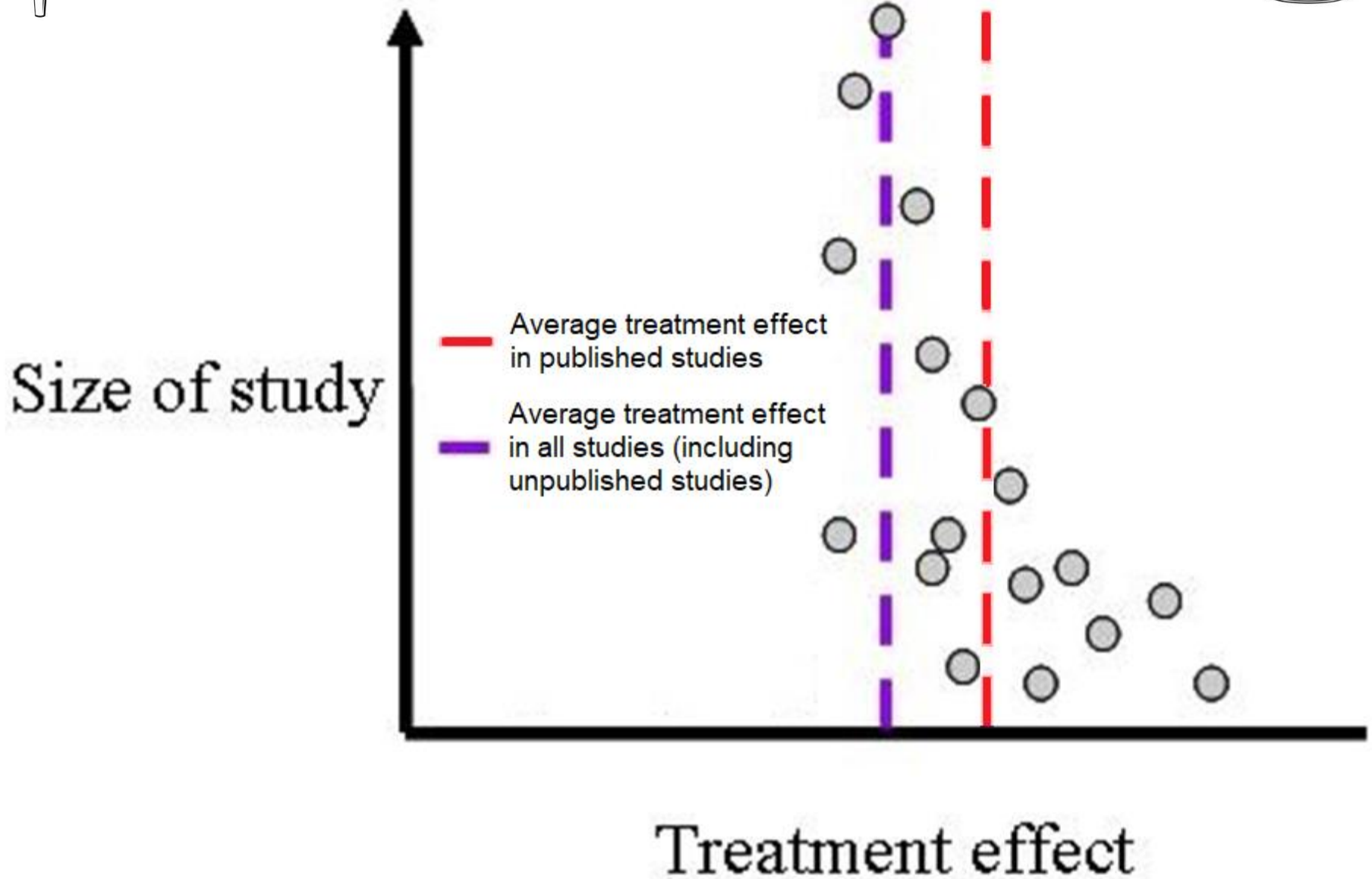
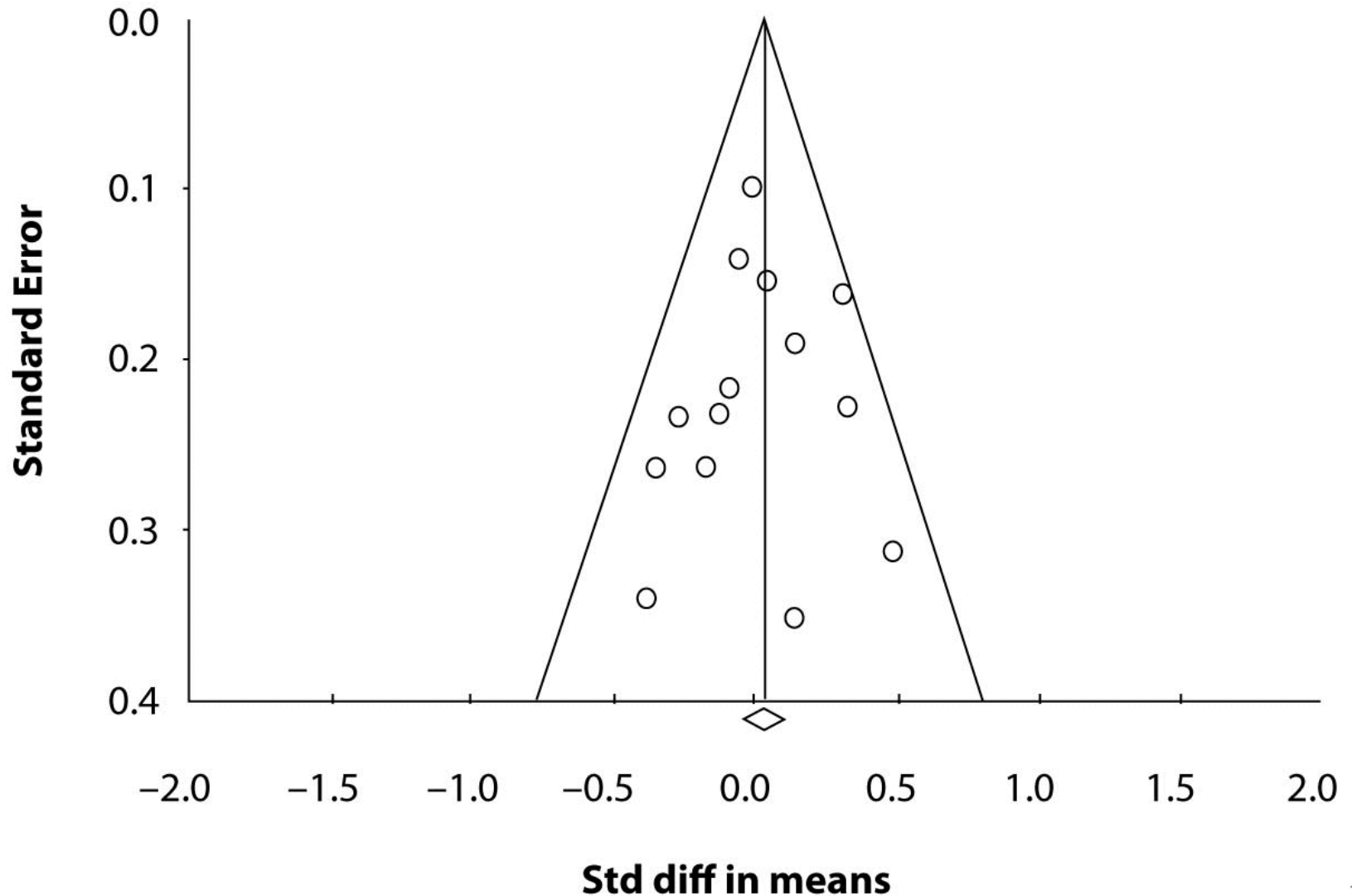6. Meta-analyses are prone to conflicts of interest

Funnel plot with no file drawer problem

Size of study

Treatment effect
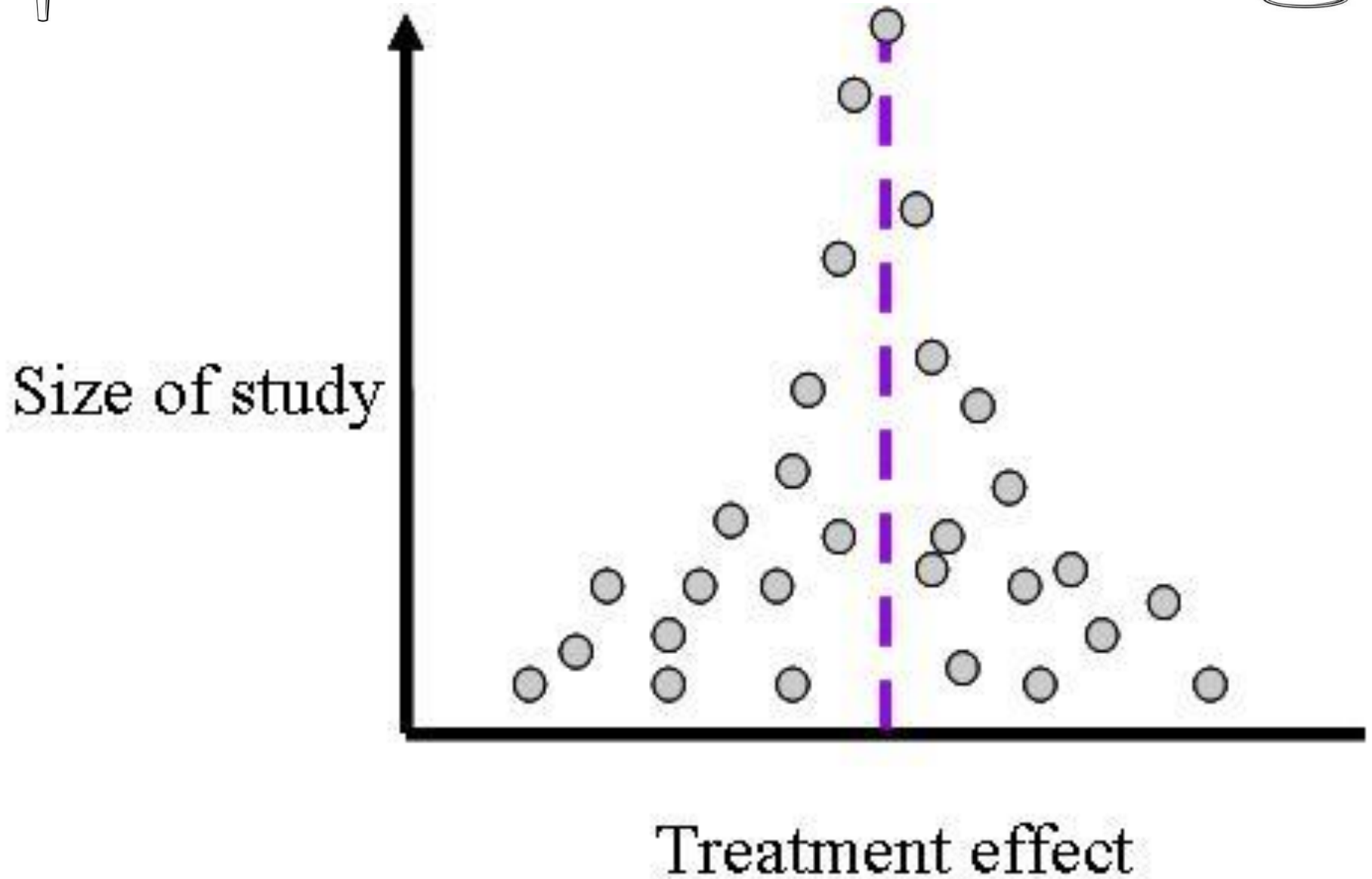
Funnel plot with file drawer problem

Size of study

Average treatment effect in published studies

Average treatment effect in all studies (including unpublished studies)

Treatment effect

Funnel Plot: std error and std mean differences

Funnel Plot of Standard Error by Std diff in means

Funnel Plot: std error and log risk ratio

Funnel Plot of Standard Error by Log risk ratio
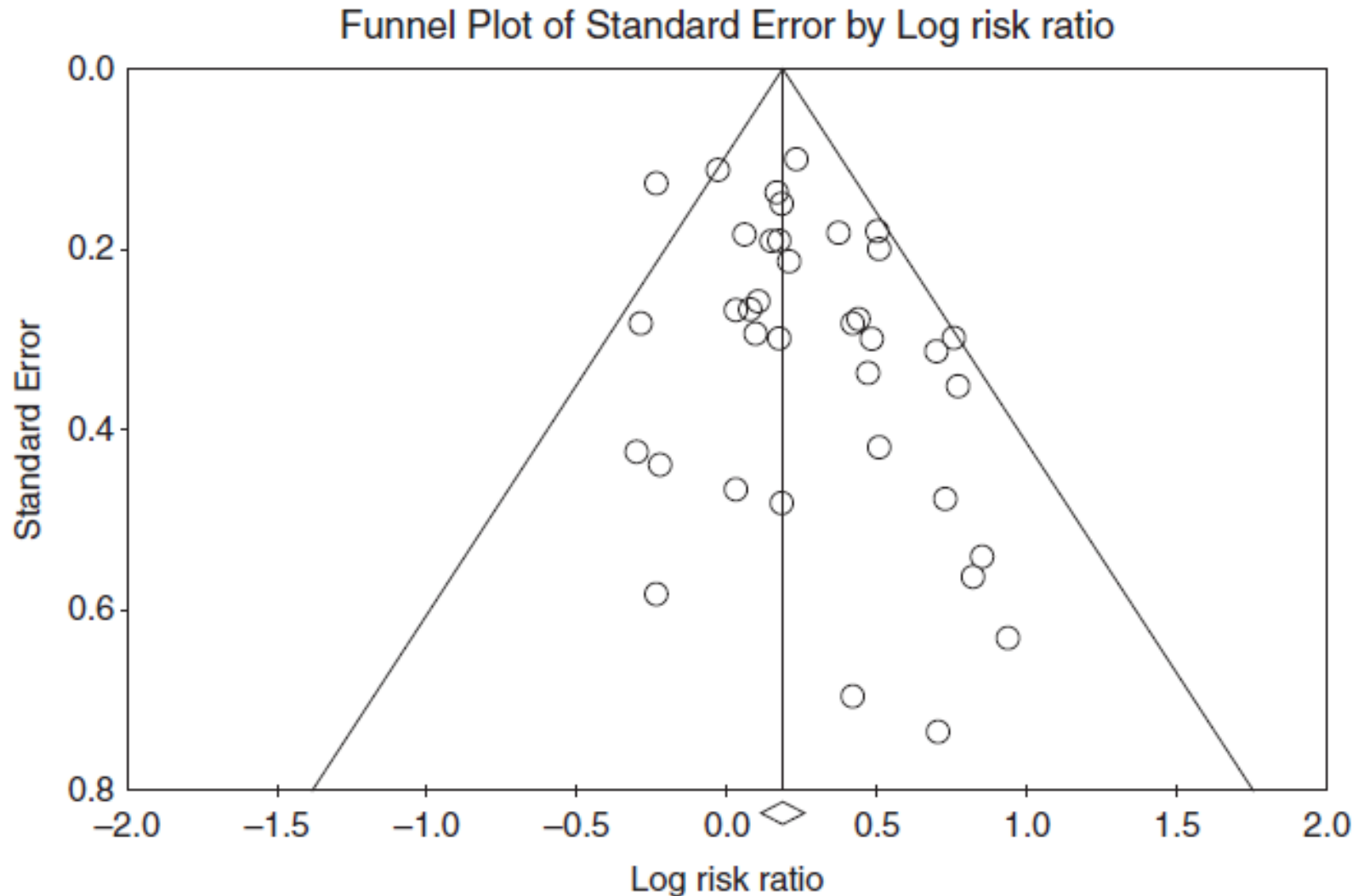
Studies imputed with the 'trim and fill' method

Funnel Plot of Standard Error by Log risk ratio
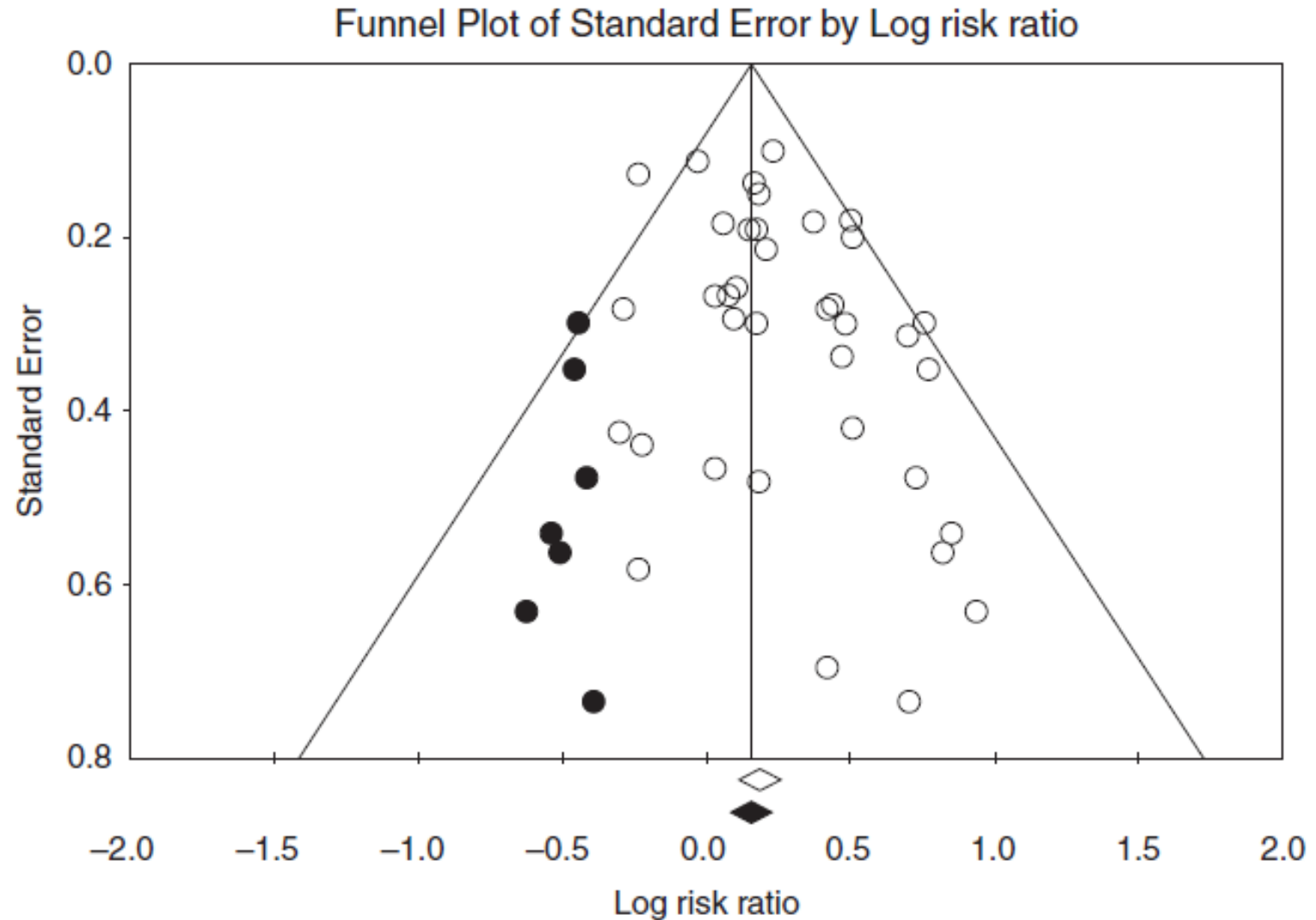
Criticisms of meta-analysis

Criticism 1: The file drawer problem invalidates meta-analysis.

Response:

- "Trim and fill" method can help

- Problem also arises in narrative review

- Unpublished studies can be sought out

- Preregistration of studies can also help

Criticisms of meta-analysis

Criticism 2: One number cannot summarize a research field.

Response: Meta-analysis seeks synthesis and not (necessarily) a summary

Criticisms of meta-analysis

Criticism 3: Meta-analysis may ignore important differences across studies, and hence be 'mixing apples and oranges'.

Response:

- Meta-analysis allows us to address the differences formally

- Mixing apples and oranges is good if you want to make a fruit salad

Criticisms of meta-analysis

Criticism 4: Garbage in, garbage out. By including many studies meta-analyses include low quality studies.

Response:

- The problem is addressed by adopting appropriate inclusion criteria.

- Meta-analysis allows a better understanding of the properties of (allegedly) garbage studies

Criticisms of meta-analysis

Criticism 5: Meta-analyses are done poorly.

Response:
- The problem is addressed by improving standards of meta-analysis

- No evidence narrative reviews are done better

- The APA Publication Manual publishes guidelines to help improve the quality of meta-analysis reporting, as does the Cochrane handbook.

Criticisms of meta-analysis

Criticism 6: Meta-analyses are prone to conflicts of interest

Case study of anti-depressants:

- 185 meta-analyses of antidepressants published during 2007-2014

- 54 (29%) had authors who were employees of the assessed drug's manufacturer
- 147 (79%) had some other industry-related conflict of interest

Some suggestions from Ioaniddes (2016):

- More transparency re: author conflicts of interest
- More widespread sharing of data
- Better enforcement of preregistration

Meta-analysis and the APA Publication Manual

Recommends providing detailed statistical information, and the complete set of raw data where possible.

– *"Mention all relevant results … be sure to include small effect sizes (or statistically nonsignificant findings)"* (p. 32).

– *"Your work will more easily become a part of the cumulative knowledge of the field if you include enough statistical information to allow its inclusion in future meta-analyses"* (p. 34).

Sets out Meta-Analysis Reporting Standards

Provides a sample paper to illustrate a meta-analysis write-up.

References

Optional reading, not required for this course but potentially useful for your future career:

American Psychological Association. (2009). *Publication manual of the American psychological association*. American Psychological Association (APA).

Bartolucci, A. A. (2009). Describing and interpreting the methodological and statistical techniques in meta-analyses. *Biochemia Medica*, *19*(2), 127-136.

Borenstein, M., Hedges, L. V., Higgins, J. P., & Rothstein, H. R. (2021).*Introduction to meta-analysis*. John Wiley & Sons.

Cumming, G. (2012). *Understanding the new statistics: Effect sizes, confidence intervals, and meta-analysis*. Routledge.

Eysenck, H. (1978). An exercise in mega-silliness. *American Psychologist*, 33, 517.

Gilbert, R., Salanti, G., Harden, M., & See, S. (2005). Infant sleeping position and the sudden infant death syndrome: systematic review of observational studies and historical review of recommendations from 1940 to 2002. *International journal of epidemiology*, *34*(4), 874-887.

Glass, G. V. (1976). Primary, secondary, and meta-analysis of research. *Educational researcher*, 3-8.

Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch VA (editors). *Cochrane Handbook for Systematic Reviews of Interventions version 6.4* (updated August 2023). Cochrane, 2023. Available from www.training.cochrane.org/handbook.

Ioannidis, J. (2016). The mass production of redundant, misleading, and conflicted systematic reviews and meta-analyses. *The Milbank Quarterly*, 94(3), 485-514.

Lau, J., Antman, E. M., Jimenez-Silva, J., Kupelnick, B., Mosteller, F., & Chalmers, T. C. (1992). Cumulative meta-analysis of therapeutic trials for myocardial infarction. *New England Journal of Medicine*, *327*(4), 248-254.

Oakes, M. W. (1986). *Statistical inference*. Epidemiology Resources.

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Wadsworth Cengage learning.

Shapiro, S. (1994). Meta-analysis/Shmeta-analysis. American Journal of Epidemiology, 140(9), 771-778.