ISCTM
THE INTERNATIONAL SOCIETY FOR CNS CLINICAL TRIALS AND METHODOLOGY

Frequentist methods of meta analysis

October 3, 2011

Amelia Island, Florida

Michael Borenstein

Michael@Meta-Analysis.com

Funded by NIH

- National Institute of Mental Health
- National Institute on Aging
- National Institute on Drug Abuse

Agenda

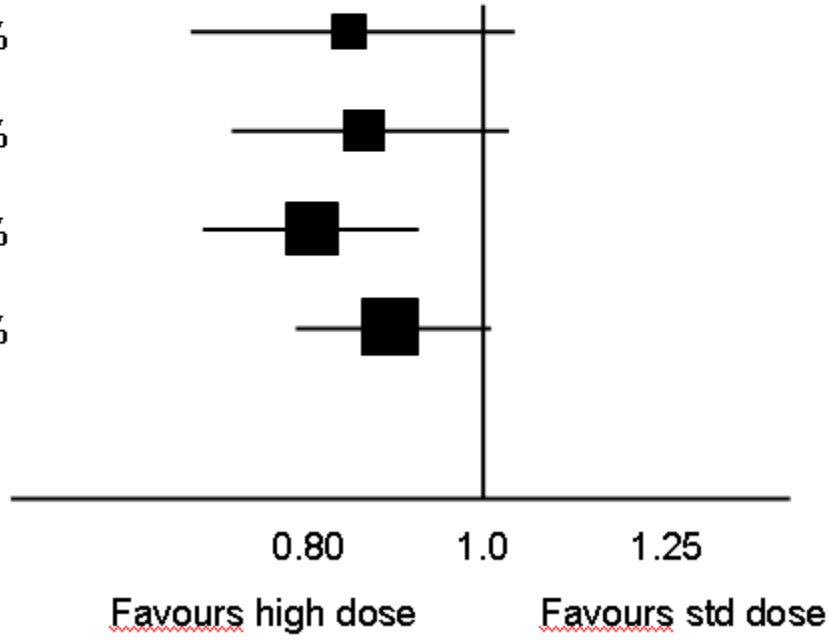
1. What is a meta-analysis
2. Why perform a meta-analysis
3. How a meta-analysis works
4. Using meta-analysis to compare subgroups
5. Criticisms of meta-analysis

What is a meta-analysis

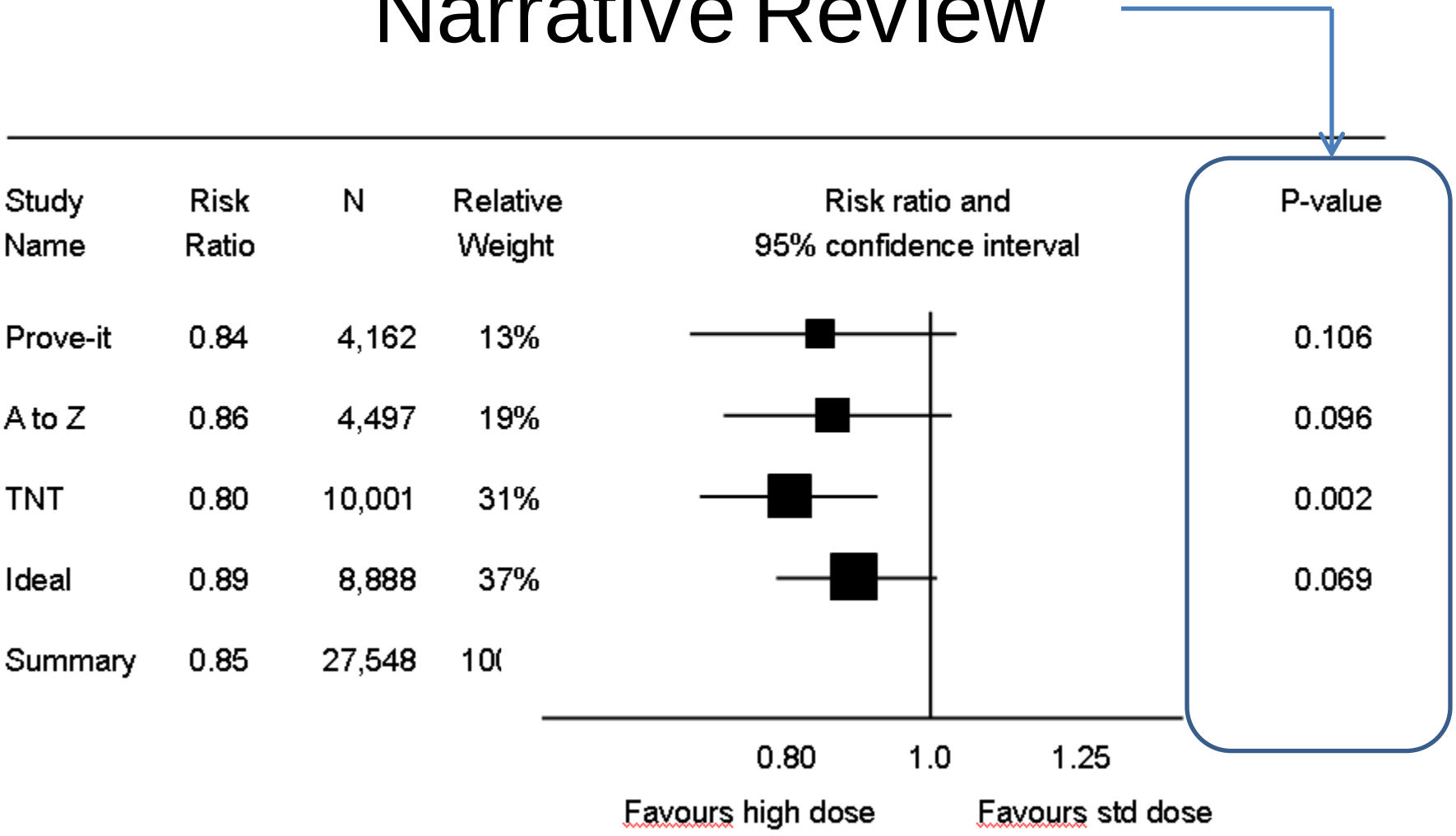
Example of Canon et al.

Impact of Statin Dose On Death and Myocardial Infarction

Study Name	Risk Ratio	N	Relative Weight	Risk ratio and 95% confidence interval	P-value
Prove-it	0.84	4,162	13%		0.106
A to Z	0.86	4,497	19%		0.096
TNT	0.80	10,001	31%		0.002
Ideal	0.89	8,888	37%		0.069

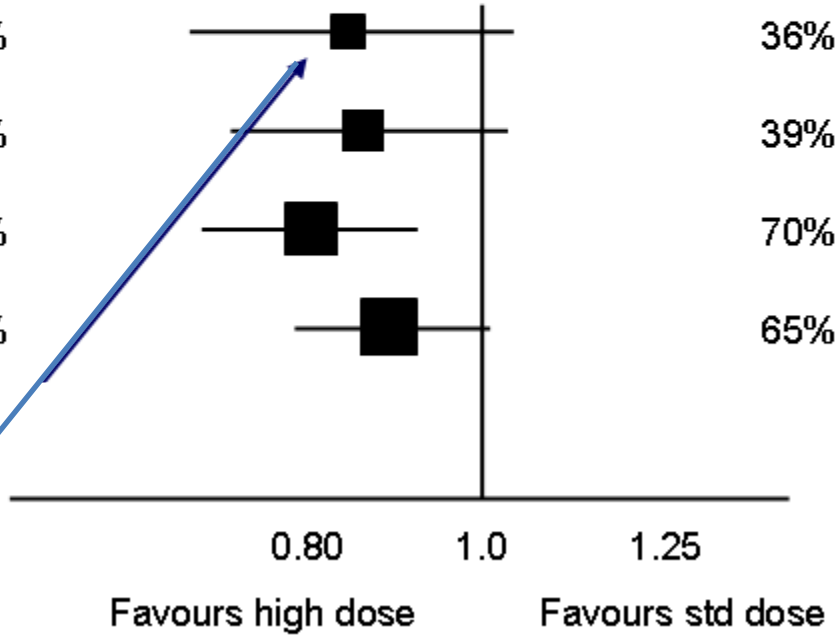


Narrative Review



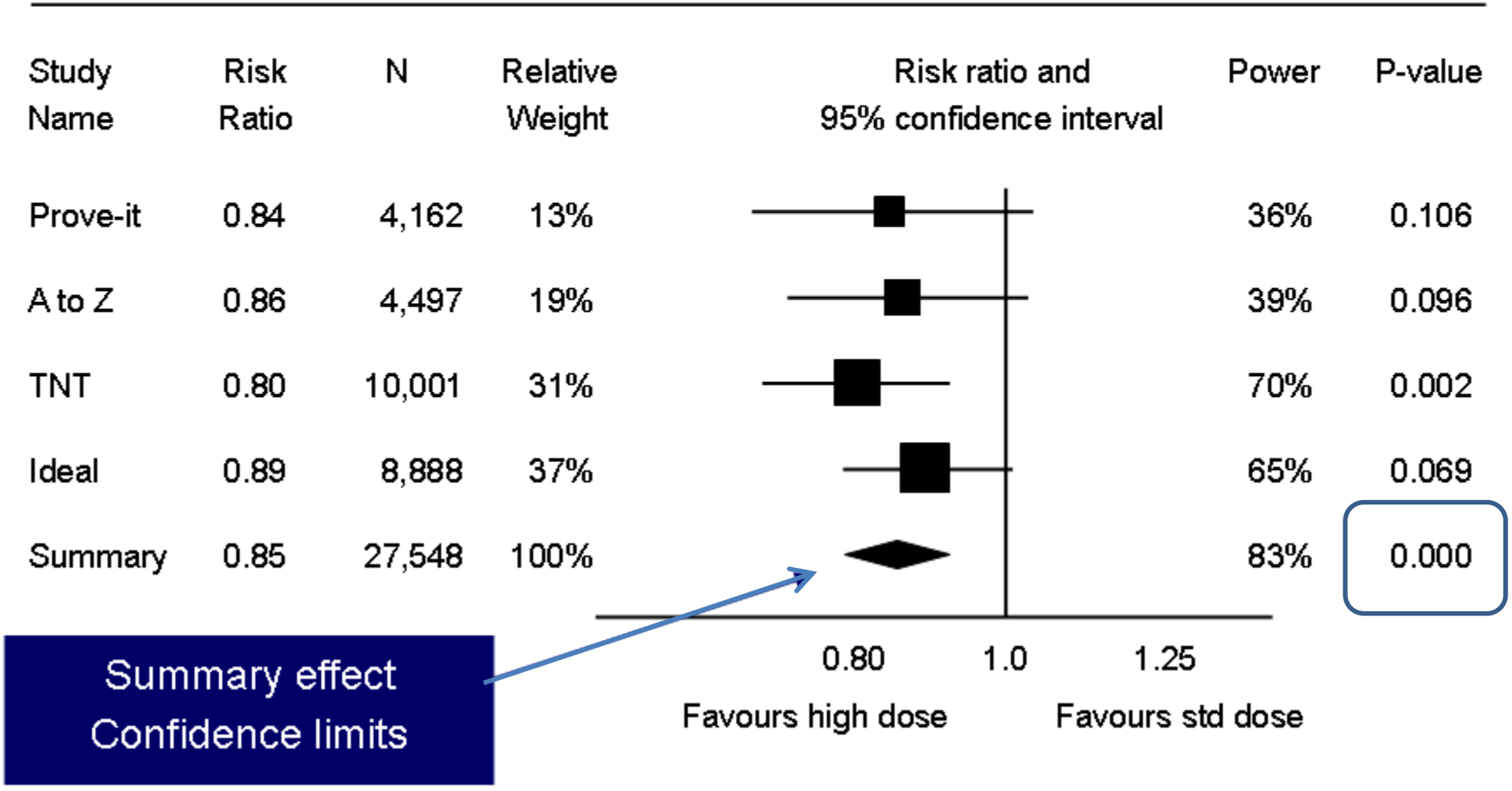
Meta-Analysis

Study Name	Risk Ratio	N	Relative Weight	Risk ratio and 95% confidence interval	Power	P-value
Prove-it	0.84	4,162	13%		36%	0.106
A to Z	0.86	4,497	19%		39%	0.096
TNT	0.80	10,001	31%		70%	0.002
Ideal	0.89	8,888	37%		65%	0.069



Effect size and weight
For each study

Meta-Analysis



Primary study

- Hypothesis
- Inclusion/exclusion criteria
- Collect data
- Statistical analysis
- Report

Systematic review

- Hypothesis
- Inclusion/exclusion criteria
- Collect data
- Statistical analysis (meta-analysis)
- Report

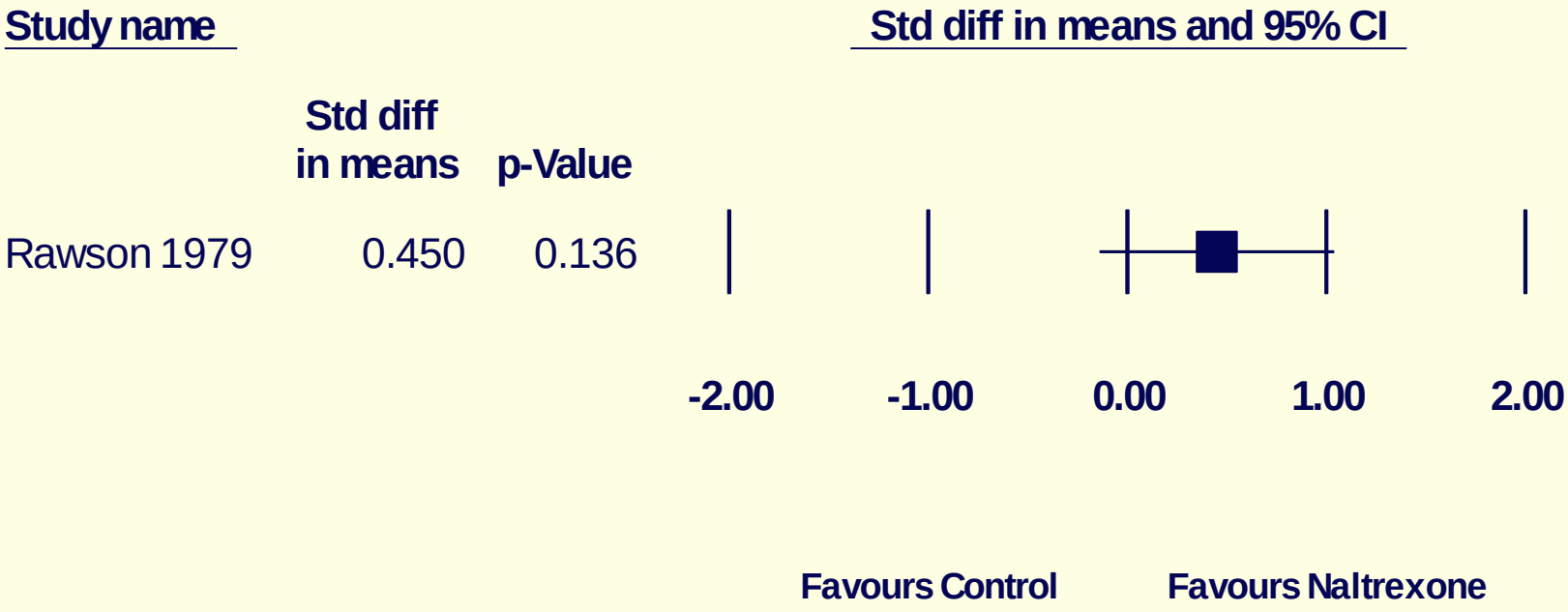
Why perform a meta-analysis

Why perform a meta-analysis

- It provides context
- It provides power
- It makes it possible to compare subgroups

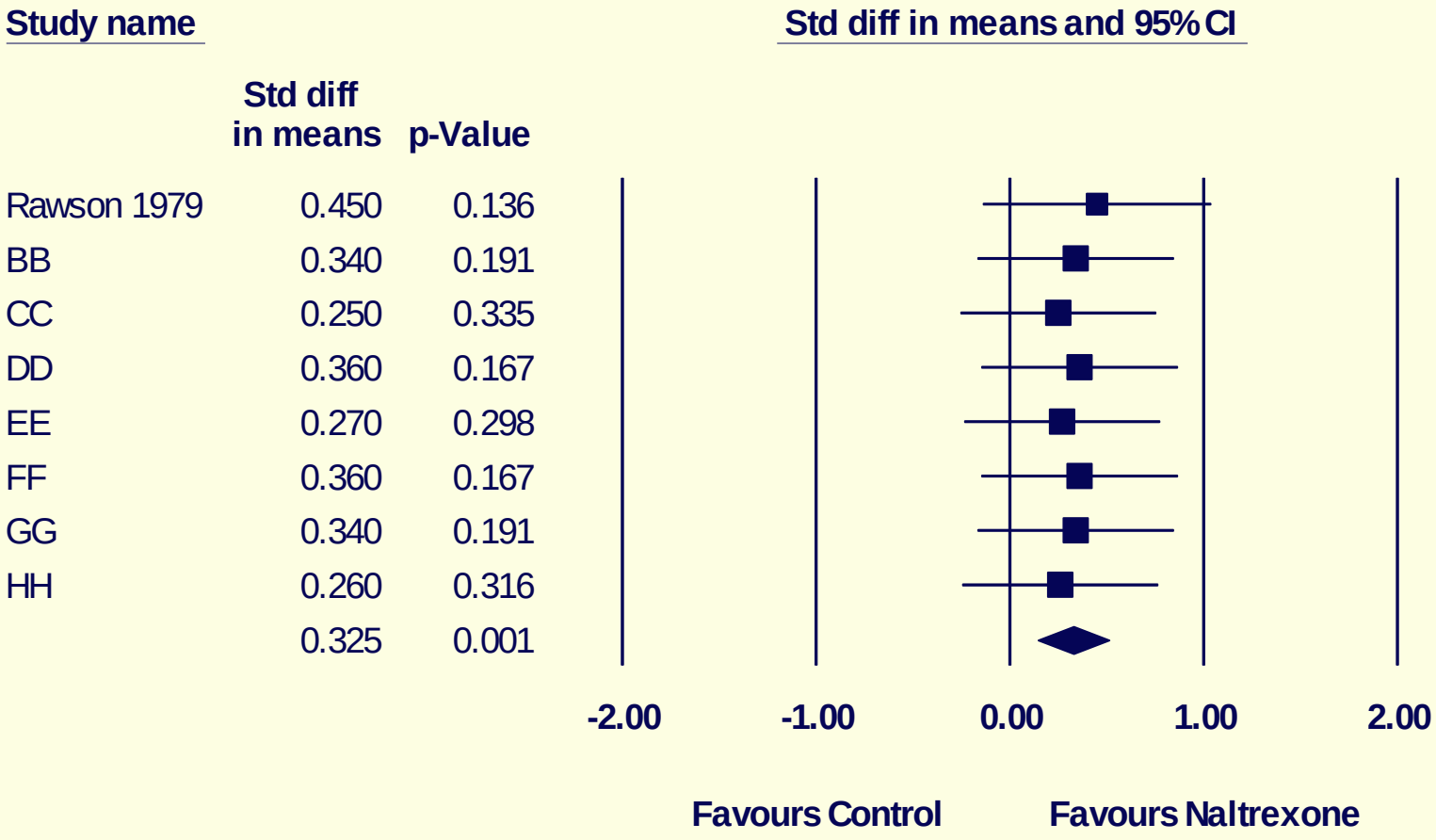
It provides context

Impact of Naltrexone on Opioid Use



Adapted from Johansson et al (1996)

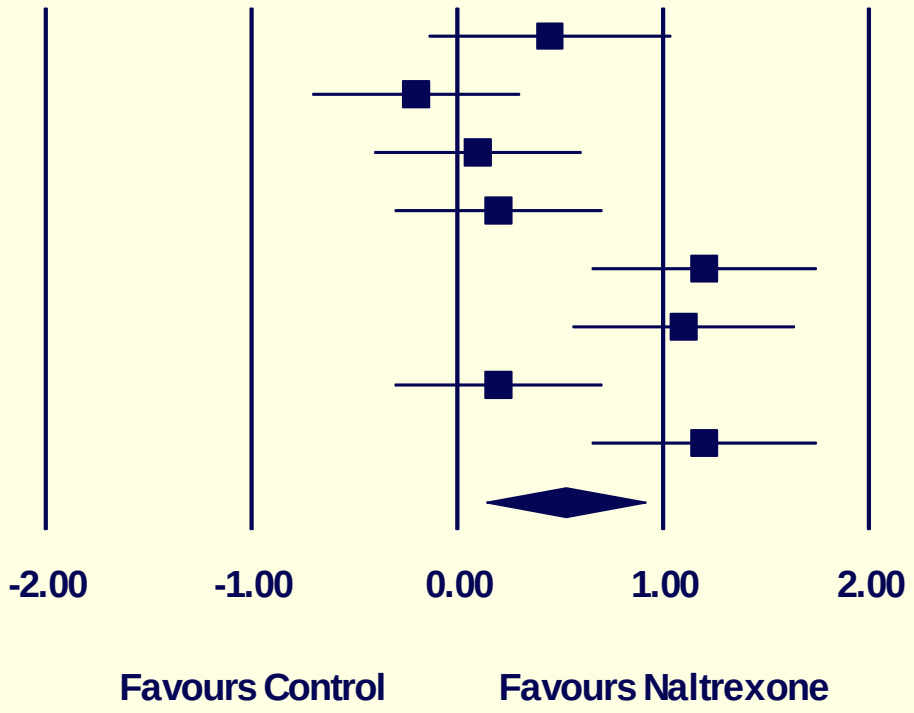
Impact of Naltrexone on Opioid Use ??



Adapted from Johansson et al (1996)

Impact of Naltrexone on Opioid Use ??

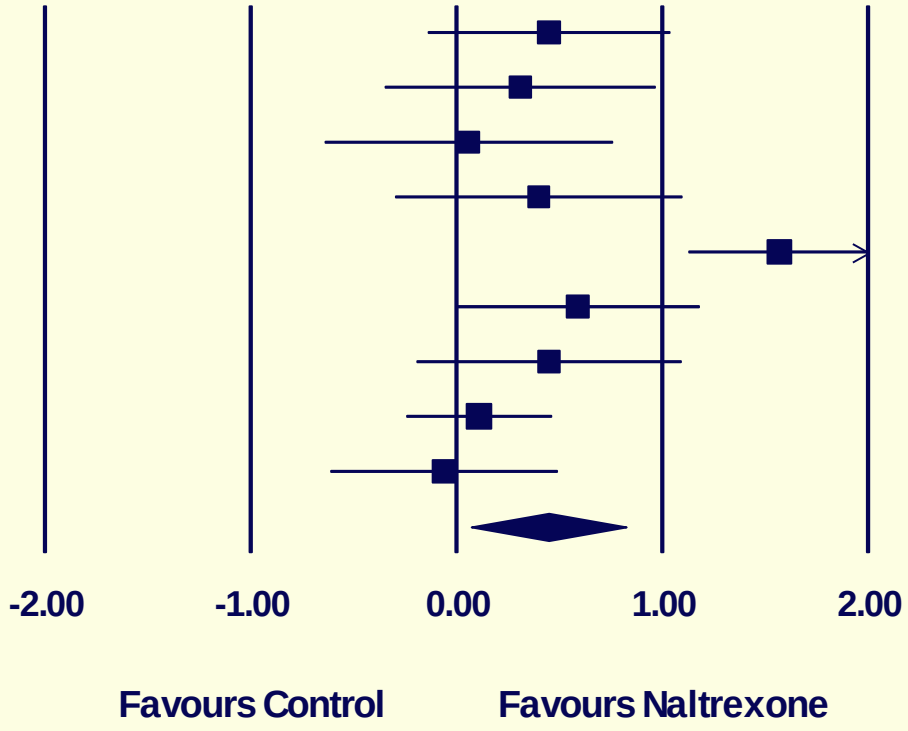
Study name	Std diff in means	p-Value
Rawson 1979	0.450	0.136
B	-0.200	0.440
C	0.100	0.699
D	0.200	0.440
E	1.200	0.000
F	1.100	0.000
G	0.200	0.440
H	1.200	0.000
	0.523	0.008



Adapted from Johansson et al (1996)

Impact of Naltrexone on Opioid Use

Study name	Std diff in means	p-Value
Rawson 1979	0.450	0.136
Guo 1992	0.310	0.358
Lerner 1992	0.060	0.867
Shufman 1994	0.400	0.263
Gerra 1995	1.570	0.000
Cornish 1997	0.590	0.052
Curran 1976	0.450	0.173
Hollister 1978	0.110	0.545
San 1991	-0.060	0.832
	0.444	0.021



Adapted from Johansson et al (1996)

It provides power

Journal of the American College of Cardiology
© 2006 by the American College of Cardiology Foundation
Published by Elsevier Inc.

Vol. 48, No. 3, 2006
ISSN 0735-1097/06/$32.00
doi:10.1016/j.jacc.2006.04.070

CLINICAL RESEARCH **Coronary Artery Disease**

Meta-Analysis of Cardiovascular Outcomes Trials Comparing Intensive Versus Moderate Statin Therapy

Christopher P. Cannon, MD, Benjamin A. Steinberg, BA, Sabina A. Murphy, MPH, Jessica L. Mega, MD, Eugene Braunwald, MD

Boston, Massachusetts

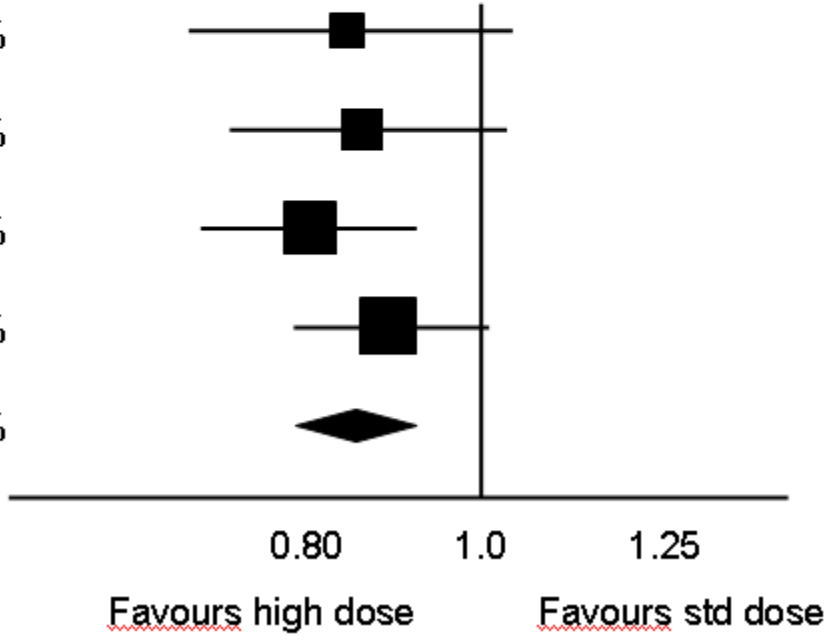
OBJECTIVES The purpose of this study was to conduct a meta-analysis that compares the reduction of cardiovascular outcomes with high-dose statin therapy versus standard dosing.

BACKGROUND Debate exists regarding the merit of more intensive lipid lowering with high-dose statin therapy as compared with standard-dose therapy.

METHODS We searched PubMed and article references for randomized controlled trials of intensive versus standard-dose statin therapy enrolling more than 1,000 patients with either stable coronary heart disease or acute coronary syndromes. Four trials were identified: the TNT (Treating to New Targets) and the IDEAL (Incremental Decrease in End Points Through Aggressive Lipid-Lowering) trials involved patients with stable cardiovascular disease, and the PROVE IT–TIMI-22 (Pravastatin or Atorvastatin Evaluation and Infection Therapy–Thrombolysis in Myocardial Infarction-22) and A-to-Z (Aggrastat-to-Zocor) trials involved patients with acute coronary syndromes. We carried out a meta-analysis of the relative odds on the basis of a fixed-effects model using the Mantel-Haenszel method for the major

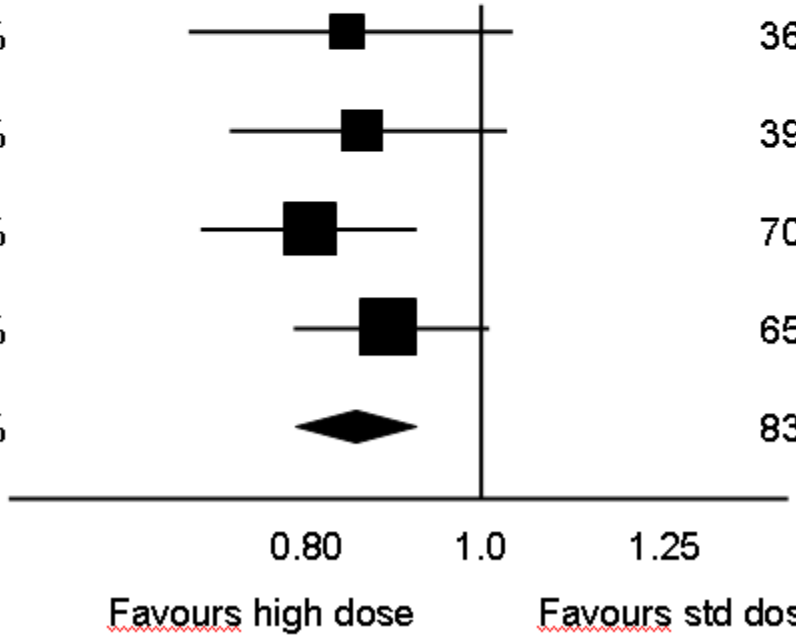
Impact of Statin Dose On Death and Myocardial Infarction

Study Name	Risk Ratio	N	Relative Weight	Risk ratio and 95% confidence interval	P-value
Prove-it	0.84	4,162	13%		0.106
A to Z	0.86	4,497	19%		0.096
TNT	0.80	10,001	31%		0.002
Ideal	0.89	8,888	37%		0.069
Summary	0.85	27,548	100%		0.000



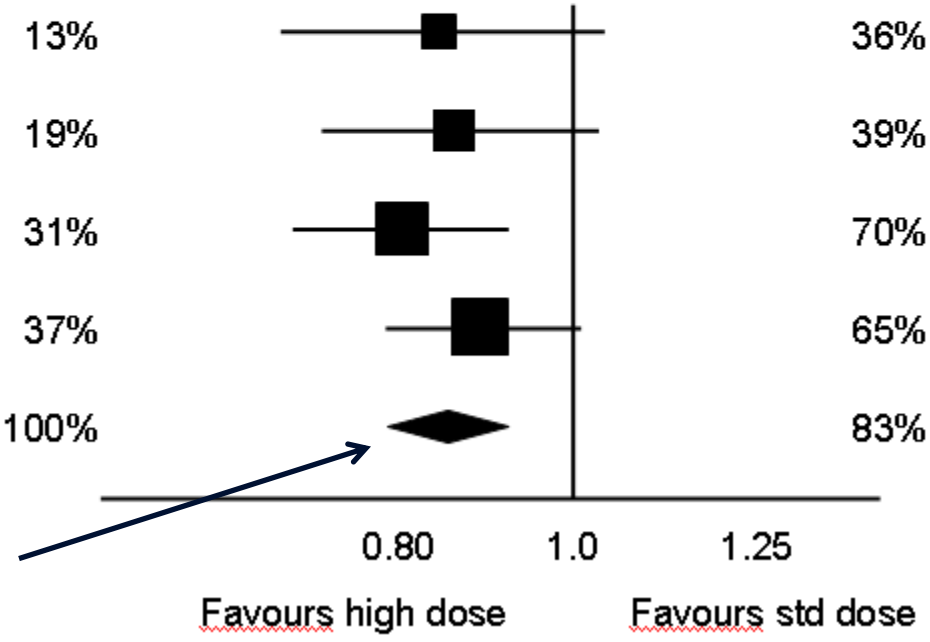
Impact of Statin Dose On Death and Myocardial Infarction

Study Name	Risk Ratio	N	Relative Weight	Risk ratio and 95% confidence interval	Power	P-value
Prove-it	0.84	4,162	13%		36%	0.106
A to Z	0.86	4,497	19%		39%	0.096
TNT	0.80	10,001	31%		70%	0.002
Ideal	0.89	8,888	37%		65%	0.069
Summary	0.85	27,548	100%		83%	0.000



Impact of Statin Dose On Death and Myocardial Infarction

Study Name	Risk Ratio	N	Relative Weight	Risk ratio and 95% confidence interval	Power	P-value
Prove-it	0.84	4,162	13%		36%	0.106
A to Z	0.86	4,497	19%		39%	0.096
TNT	0.80	10,001	31%		70%	0.002
Ideal	0.89	8,888	37%		65%	0.069
Summary	0.85	27,548	100%		83%	0.000



Meta-analysis
$p = 0.00003$

Avandia Meta-Analysis

The NEW ENGLAND JOURNAL of MEDICINE

Effect of Rosiglitazone on the Risk of Myocardial Infarction and Death from Cardiovascular Causes

Steven E. Nissen, M.D., and Kathy Wolski, M.P.H.

ABSTRACT

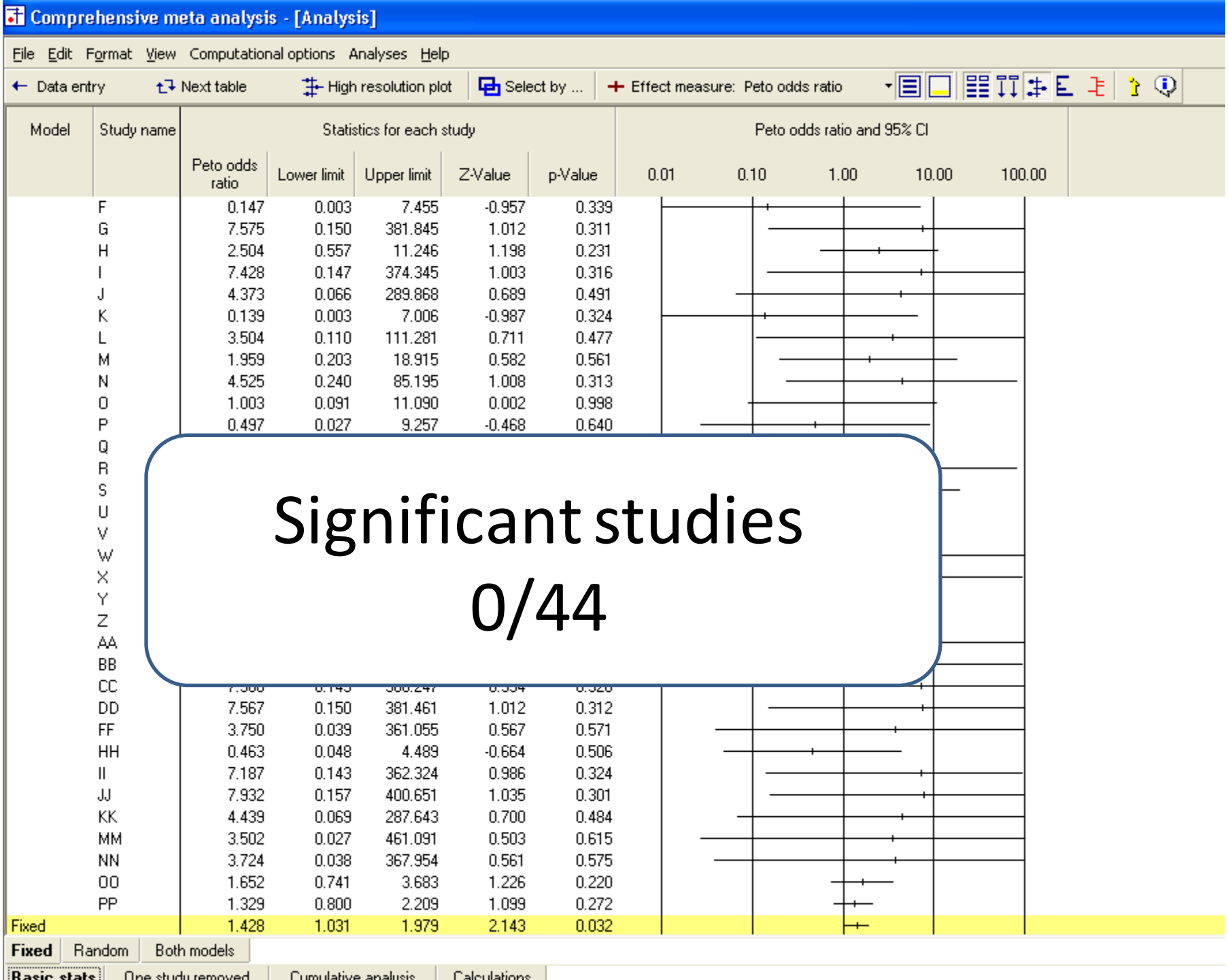
BACKGROUND

Rosiglitazone is widely used to treat patients with type 2 diabetes mellitus, but its effect on cardiovascular morbidity and mortality has not been determined.

METHODS

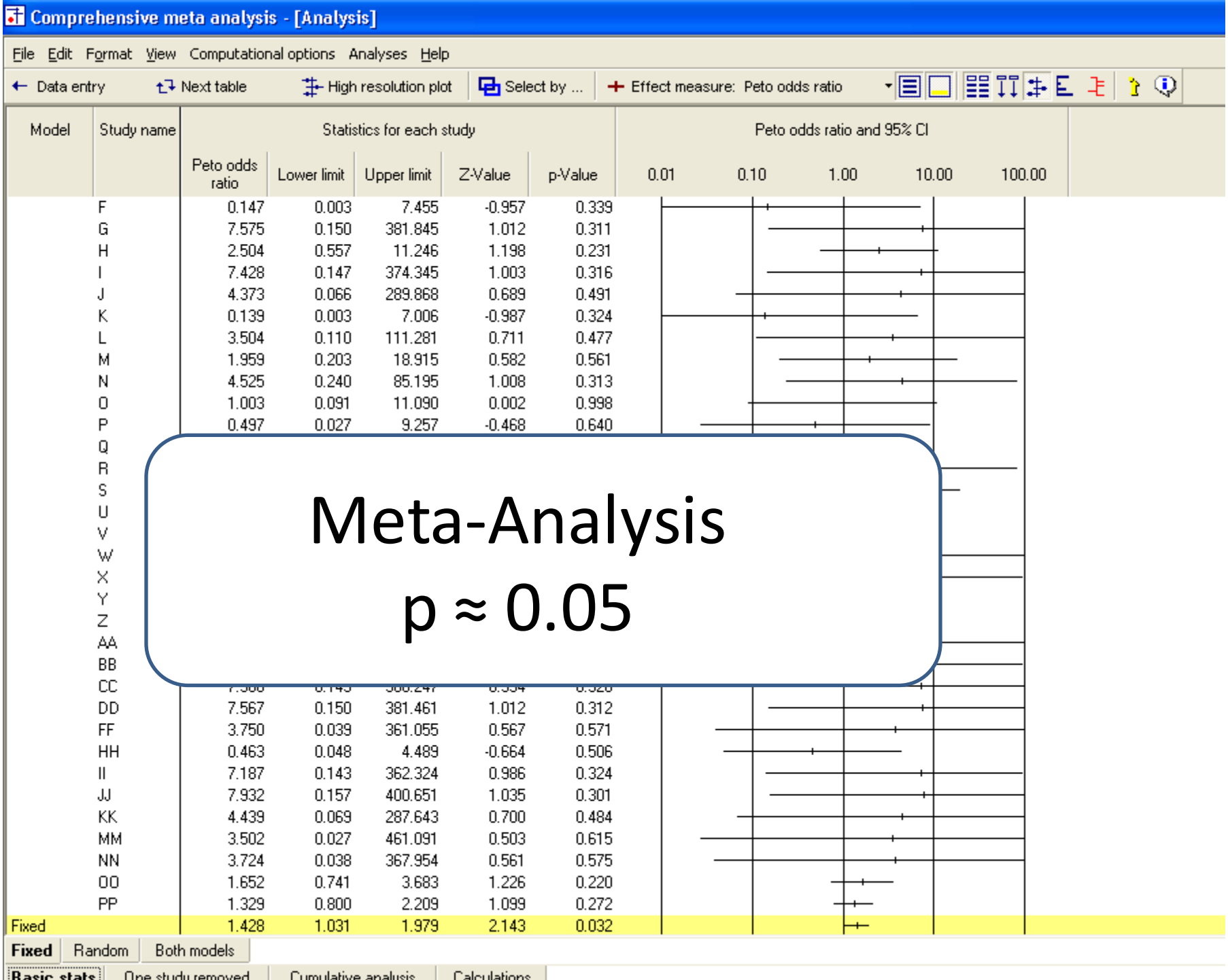
We conducted searches of the published literature, the Web site of the Food and

From the Cleveland Clinic, Cleveland. Address reprint requests to Dr. Nissen at the Department of Cardiovascular Medicine, Cleveland Clinic, 9500 Euclid Ave., Cleveland, OH 44195, or at nissens@ccf.org.



Significant studies

0/44



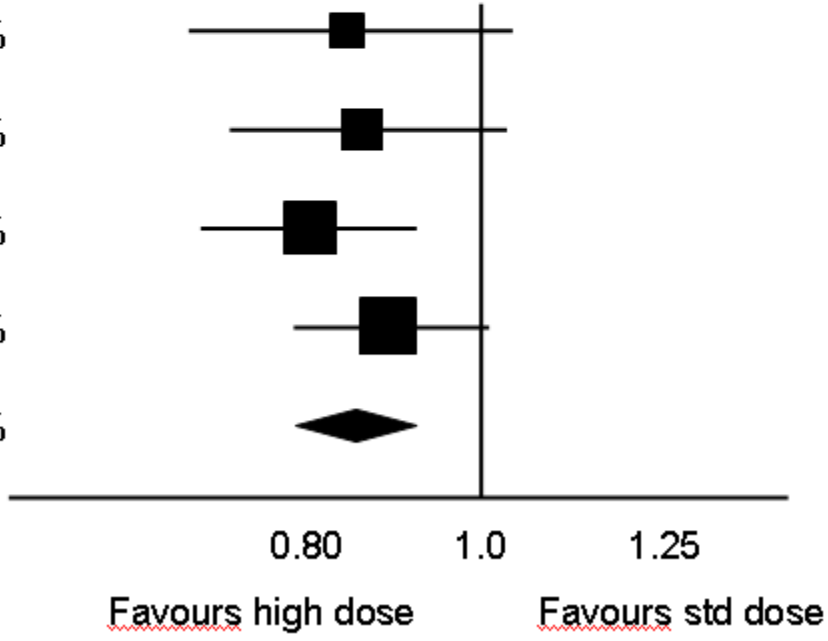
Meta-Analysis

$p \approx 0.05$

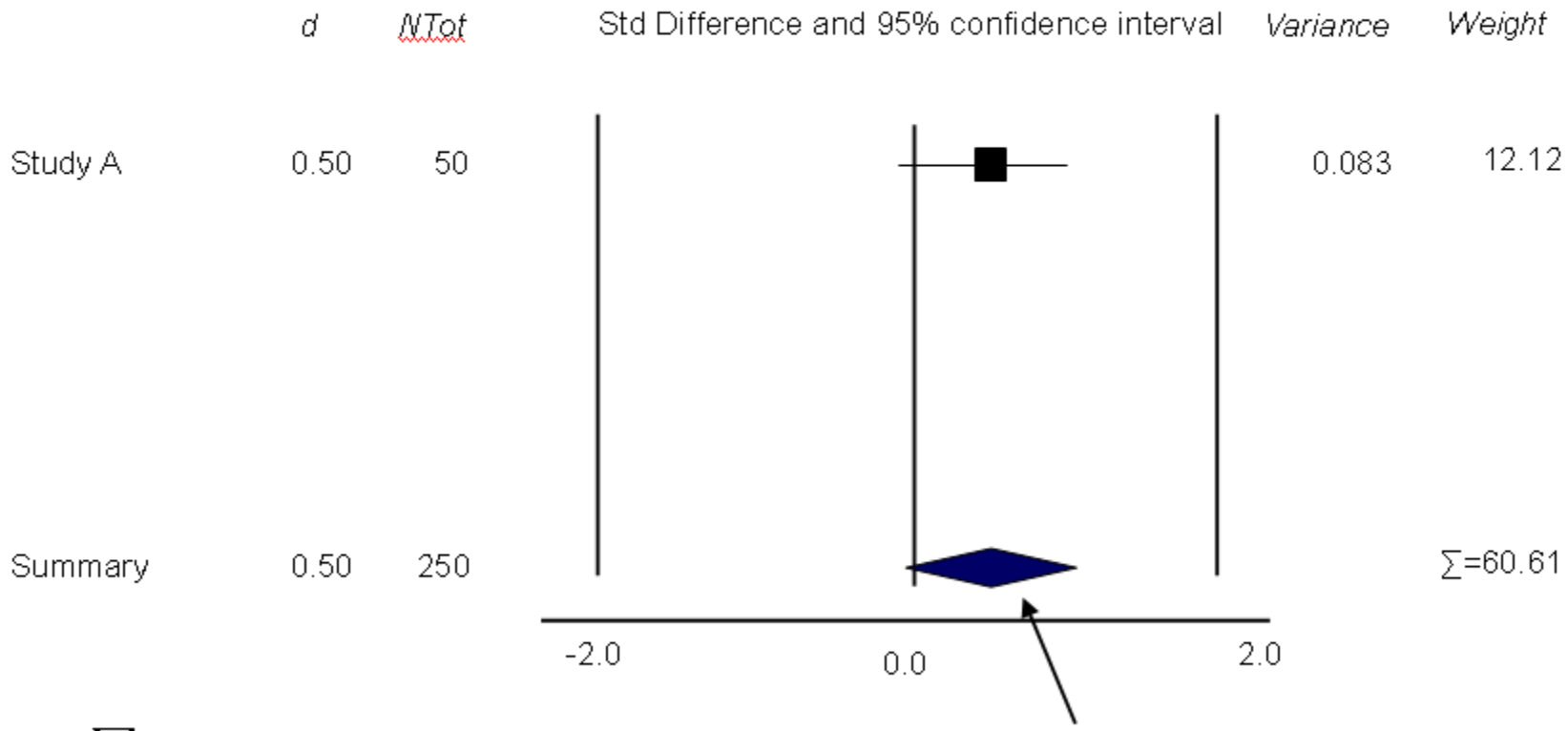
How a meta-analysis works

Impact of Statin Dose On Death and Myocardial Infarction

Study Name	Risk Ratio	N	Relative Weight	Risk ratio and 95% confidence interval	P-value
Prove-it	0.84	4,162	13%		0.106
A to Z	0.86	4,497	19%		0.096
TNT	0.80	10,001	31%		0.002
Ideal	0.89	8,888	37%		0.069
Summary	0.85	27,548	100%		0.000

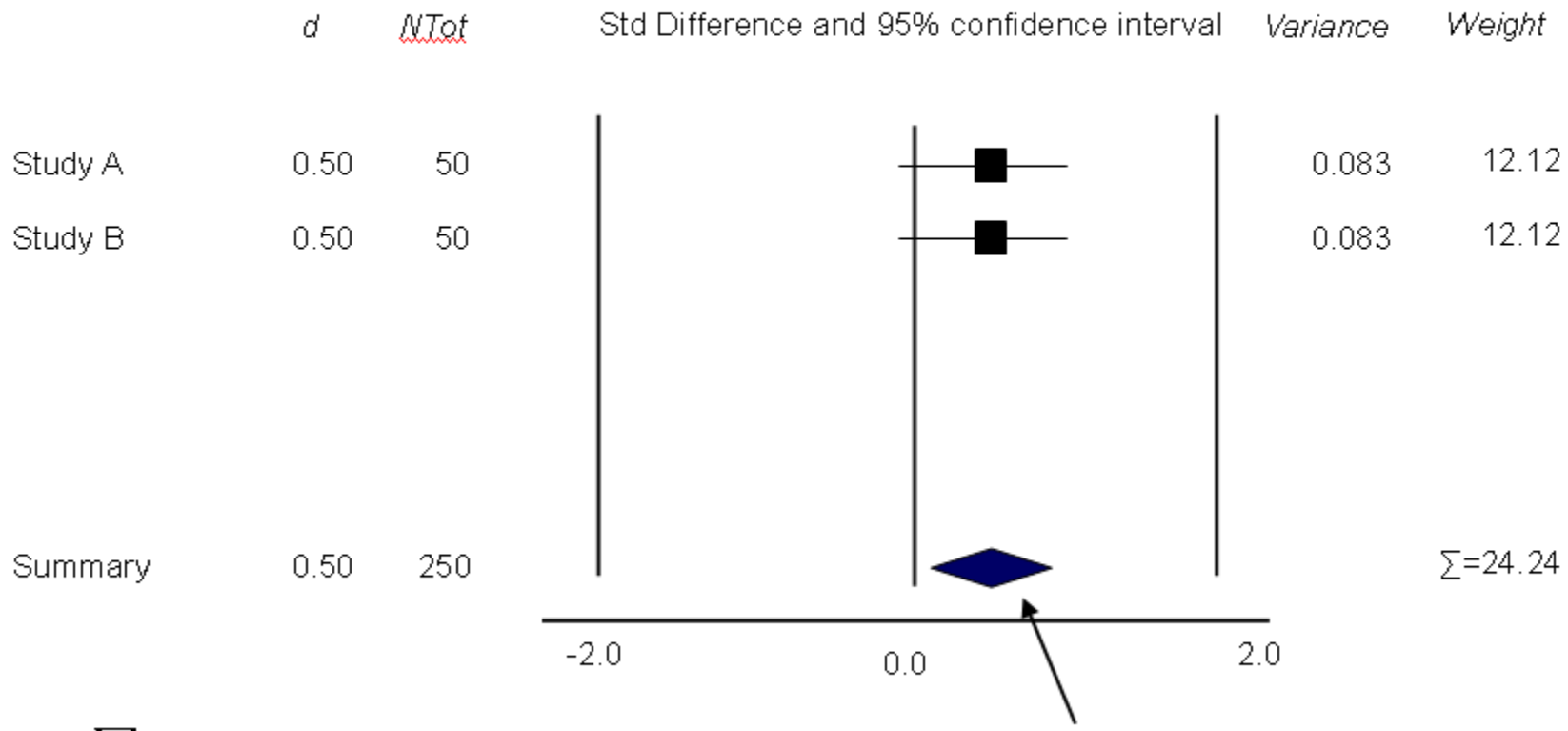


Meta-analysis with consistent effects $k = 1$



$$M = \frac{\sum W_i Y_i}{\sum W_i} = \frac{6.06}{12.12} = 0.50 \qquad V_M = \frac{1}{\sum W_i} = \frac{1}{12.12} = 0.083 \; SE = \sqrt{0.083} = 0.287$$

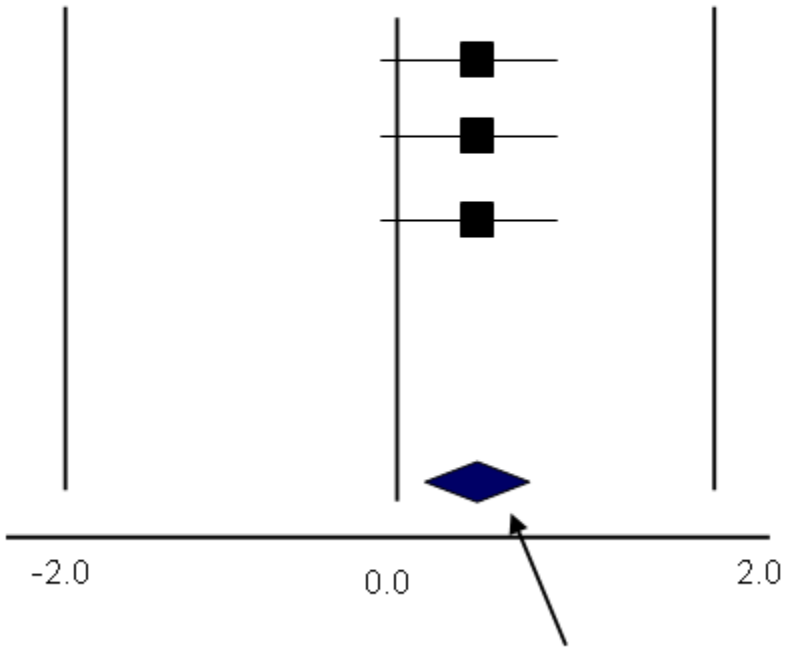
Meta-analysis with consistent effects $k = 2$



$$M = \frac{\sum W_i Y_i}{\sum W_i} = \frac{12.12}{24.24} = 0.50 \qquad V_M = \frac{1}{\sum W_i} = \frac{1}{24.24} = 0.041 \; SE = \sqrt{0.041} = 0.203$$

Meta-analysis with consistent effects $k = 3$

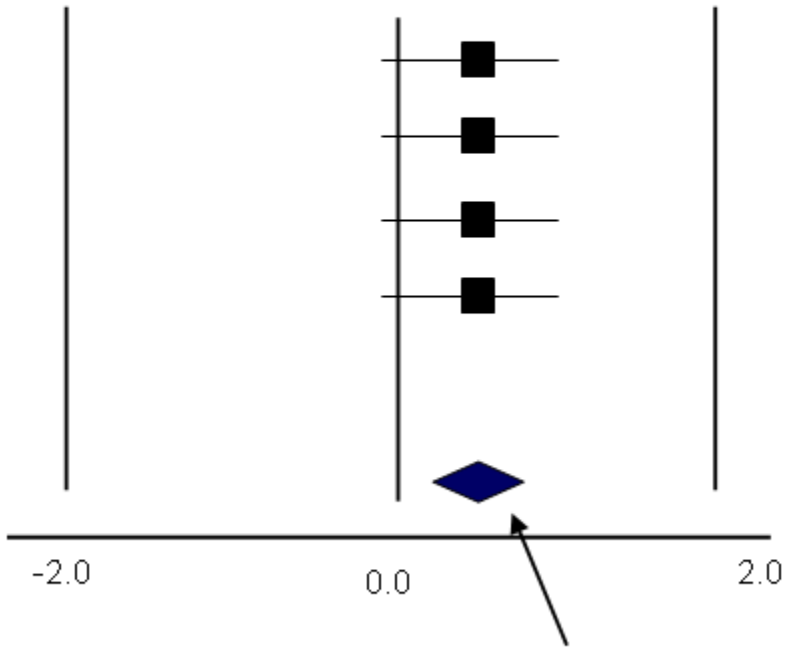
	d	*NTot*	Std Difference and 95% confidence interval	*Variance*	*Weight*
Study A	0.50	50		0.083	12.12
Study B	0.50	50		0.083	12.12
Study C	0.50	50		0.083	12.12
Summary	0.50	250			∑=36.36



$$M = \frac{\sum W_i Y_i}{\sum W_i} = \frac{18.18}{36.36} = 0.50 \qquad V_M = \frac{1}{\sum W_i} = \frac{1}{36.36} = 0.028 \; SE = \sqrt{0.028} = 0.166$$

Meta-analysis with consistent effects $k = 4$

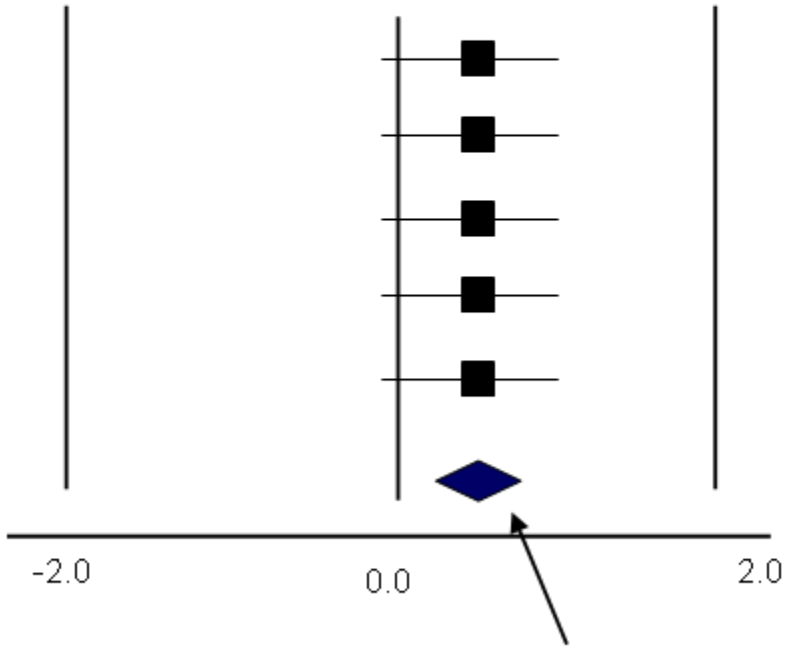
	d	*NTot*	Std Difference and 95% confidence interval	*Variance*	*Weight*
Study A	0.50	50		0.083	12.12
Study B	0.50	50		0.083	12.12
Study C	0.50	50		0.083	12.12
Study D	0.50	50		0.083	12.12
Summary	0.50	250			∑=48.49



$$M = \frac{\sum W_i Y_i}{\sum W_i} = \frac{24.24}{48.48} = 0.50 \qquad V_M = \frac{1}{\sum W_i} = \frac{1}{48.48} = 0.021 \quad SE = \sqrt{0.021} = 0.144$$

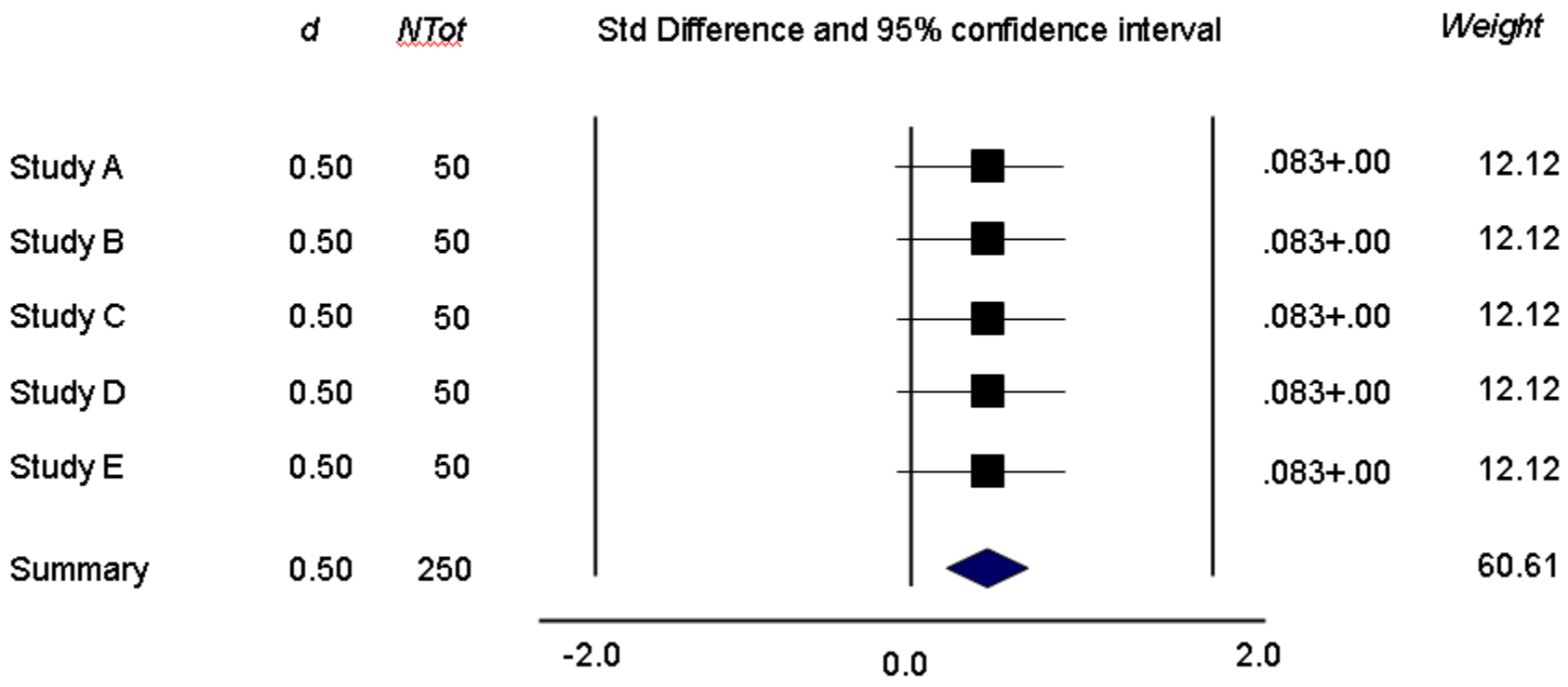
Meta-analysis with consistent effects $k = 5$

	d	*NTot*	Std Difference and 95% confidence interval	*Variance*	*Weight*
Study A	0.50	50		0.083	12.12
Study B	0.50	50		0.083	12.12
Study C	0.50	50		0.083	12.12
Study D	0.50	50		0.083	12.12
Study E	0.50	50		0.083	12.12
Summary	0.50	250			∑=60.61



$$M = \frac{\sum W_i Y_i}{\sum W_i} = \frac{30.30}{60.60} = 0.50 \qquad V_M = \frac{1}{\sum W_i} = \frac{1}{60.61} = 0.017 \qquad SE = \sqrt{0.017} = 0.128$$

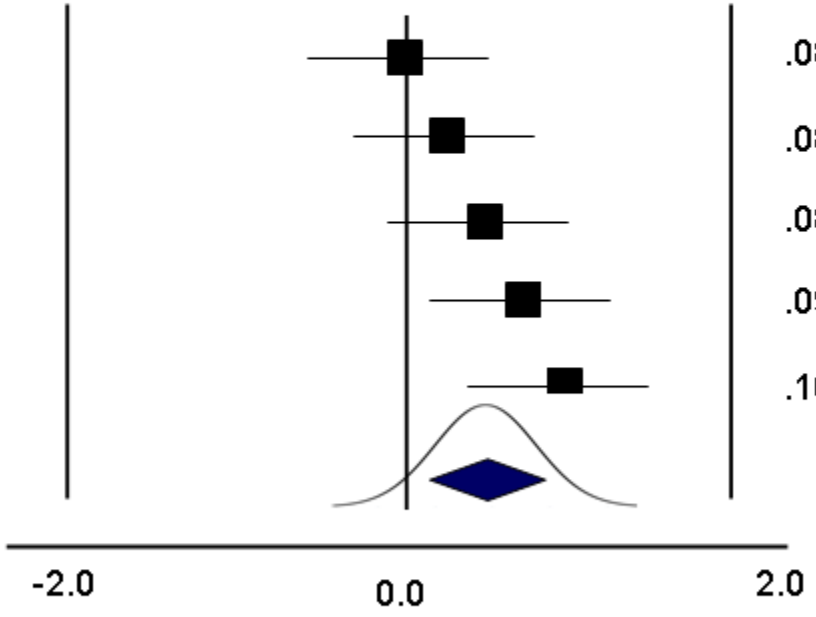
Meta-analysis with consistent effects



$$M = \frac{\sum W_i Y_i}{\sum W_i} = \frac{30.30}{60.60} = 0.50 \qquad V_M = \frac{1}{\sum W_i} = \frac{1}{60.61} = 0.017 \qquad SE = \sqrt{0.017} = 0.128$$

Meta-analysis with heterogeneous effects

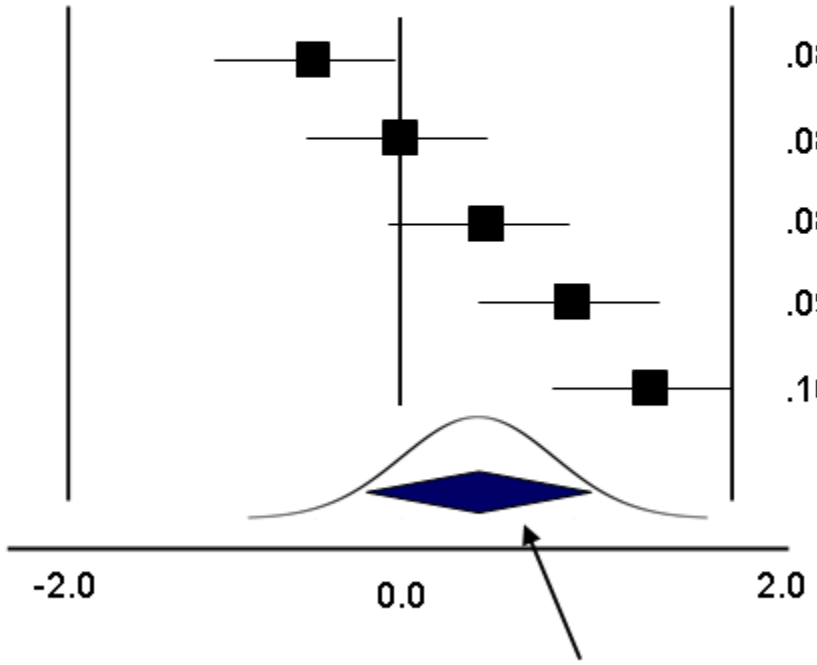
	d	*NTot*	Std Difference and 95% confidence interval	*Variance*	*Weight*
Study A	0.00	50		.083+.07	6.62
Study B	0.25	50		.080+.07	6.60
Study C	0.50	50		.083+.07	6.52
Study D	0.75	50		.090+.07	6.39
Study E	1.00	50		.103+.07	6.21
Summary	0.49	250			∑=32.33



$$M = \frac{\sum W_i Y_i}{\sum W_i} = \frac{15.91}{32.32} = 0.49 \qquad V_M = \frac{1}{\sum W_i} = \frac{1}{32.32} = 0.031 \quad SE = \sqrt{0.031} = 0.176$$

Meta-analysis with heterogeneous effects

	d	*NTot*	Std Difference and 95% confidence interval	*Variance*	*Weight*
Study A	-0.50	50		.083+.51	1.68
Study B	0.00	50		.080+.51	1.69
Study C	0.50	50		.083+.51	1.68
Study D	1.00	50		.090+.51	1.66
Study E	1.50	50		.103+.51	1.62
Summary	0.49	250			∑=8.322



$$M = \frac{\sum W_i Y_i}{\sum W_i} = \frac{4.09}{8.32} = 0.49 \qquad V_M = \frac{1}{\sum W_i} = \frac{1}{8.32} = 0.120 \quad SE = \sqrt{0.120} = 0.347$$

Weights

- Reflect the *total* amount of information
- Reflect the *relative* amount of information

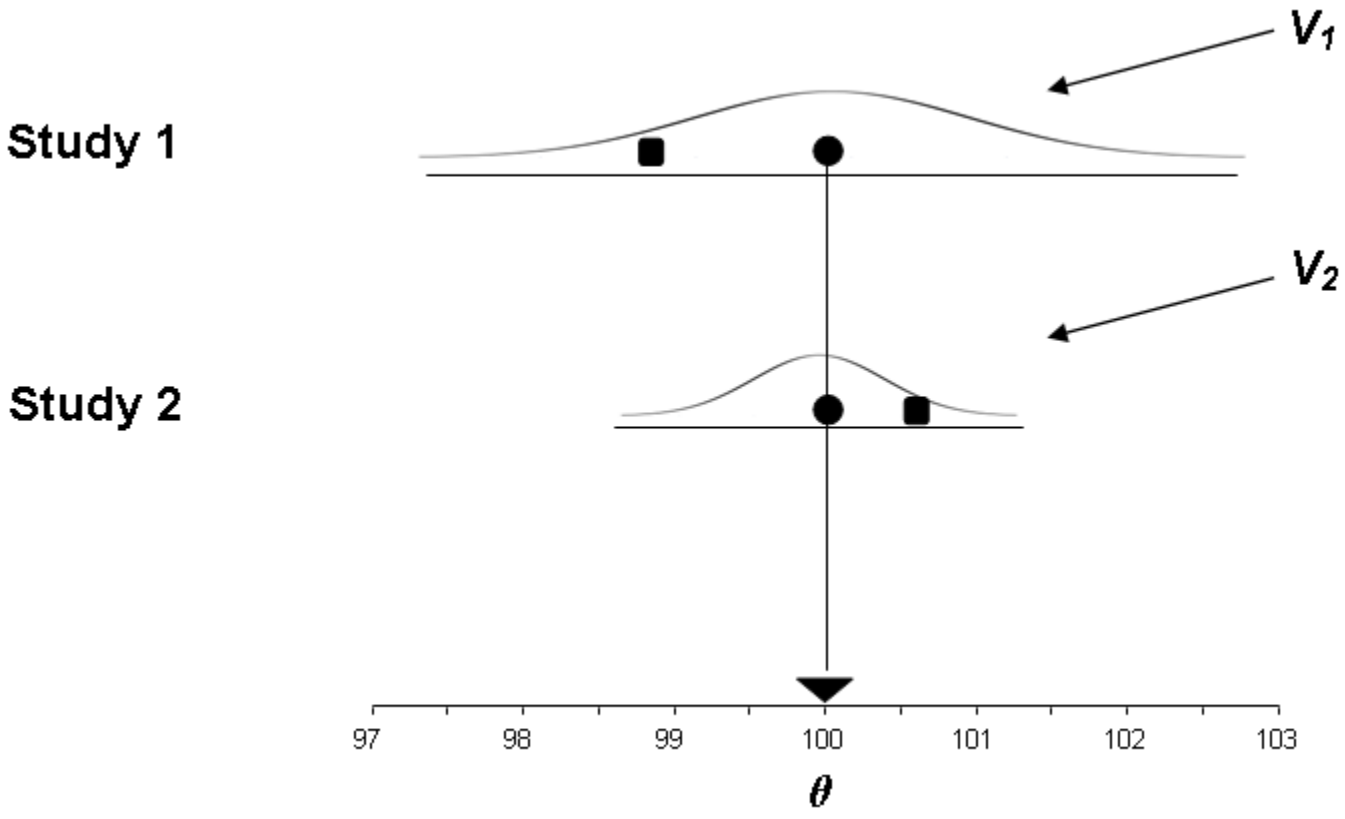
Data types

- Means
- Binary data
- Correlation data
- Survival data
- Rate data

How weights are assigned

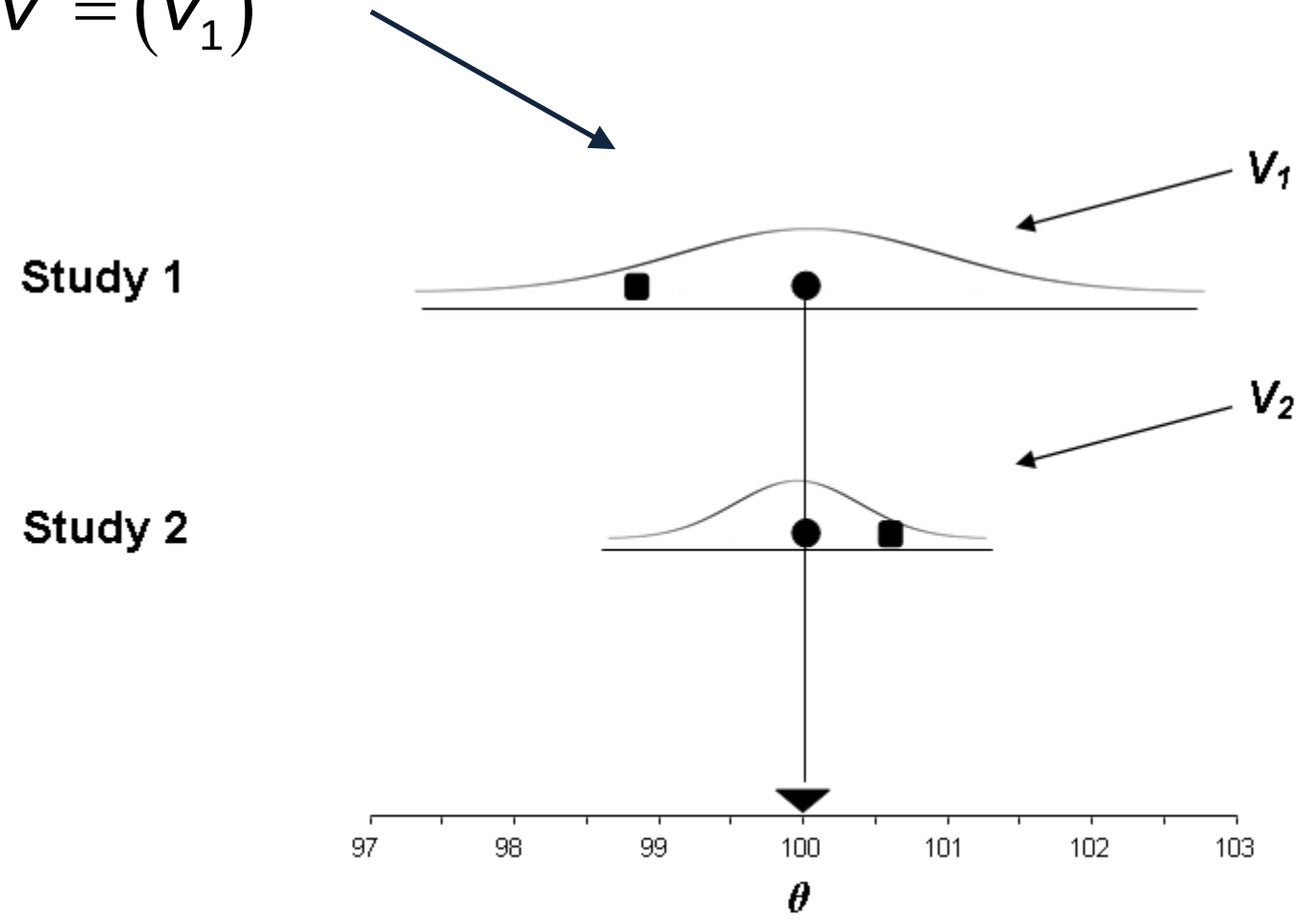
- Fixed-effect model
- Random-effects model

Fixed effect model



Fixed effect model

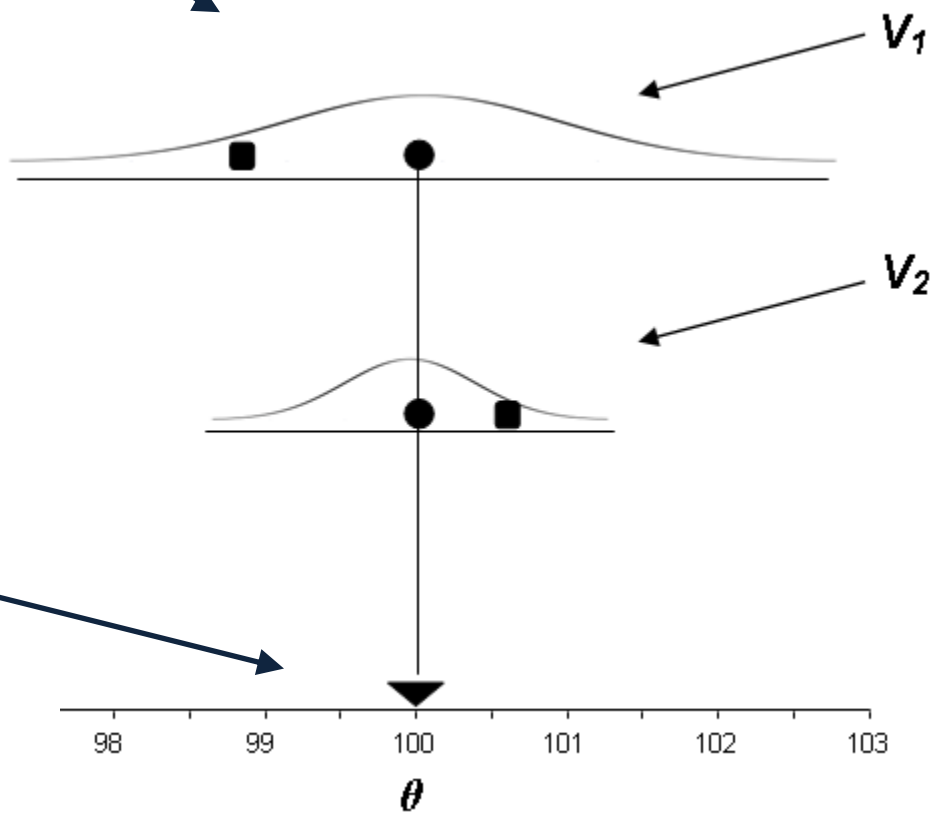
$$V = (V_1)$$



Fixed effect model

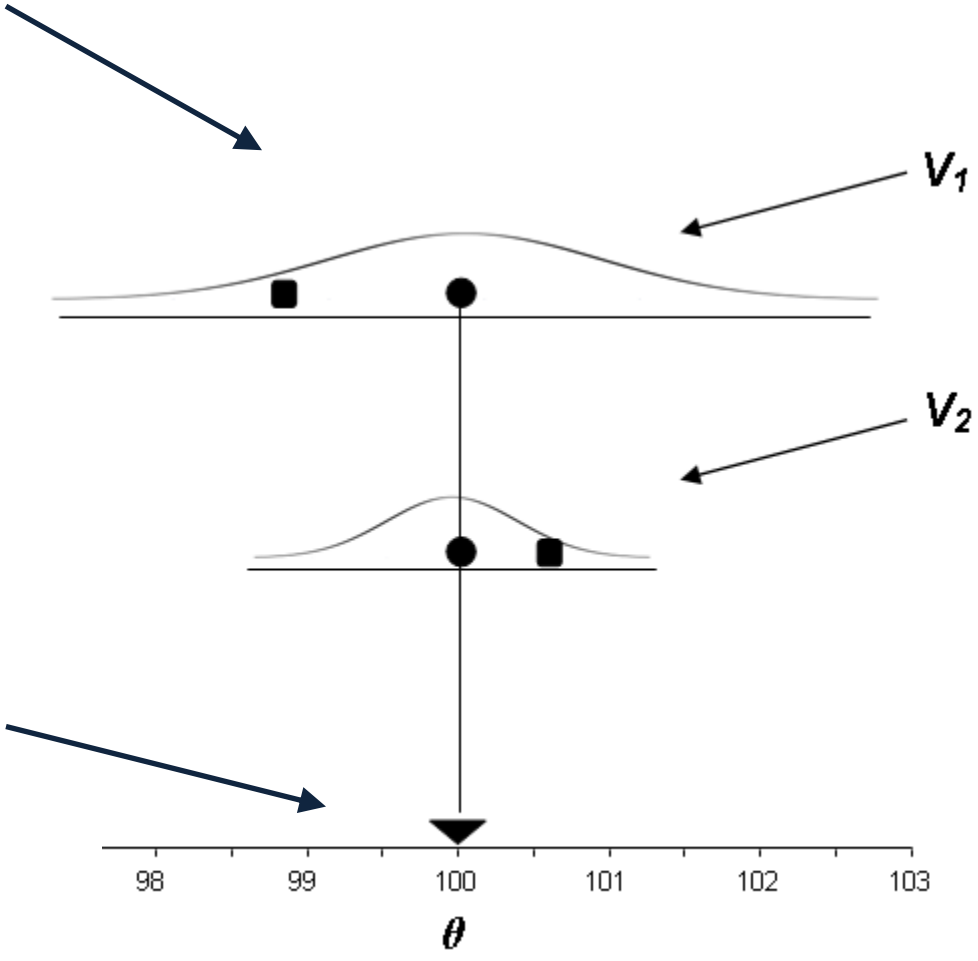
$$W = 1/(V_1)$$

$$M = \frac{\sum_{i=1}^{k} W_i Y_i}{\sum_{i=1}^{k} W_i}$$



Fixed effect model

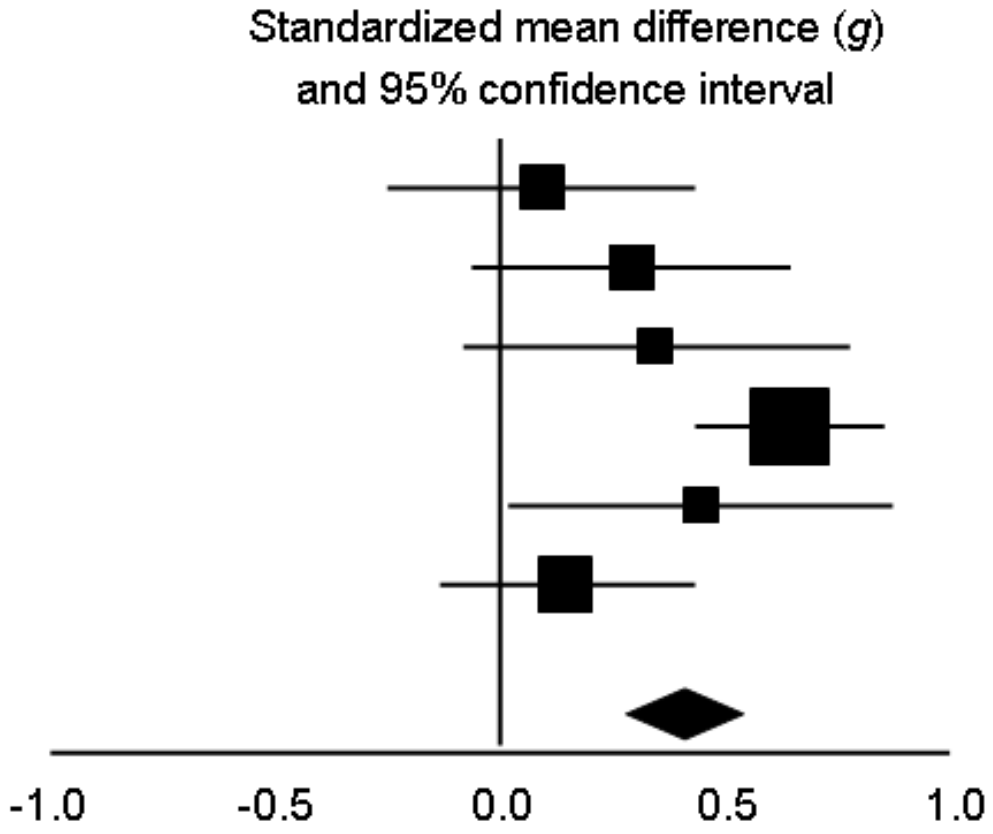
$$W = 1/(V_1)$$



$$V_M = \frac{1}{\sum_{i=1}^{k} W_i}$$

Impact of Intervention (Fixed effect)

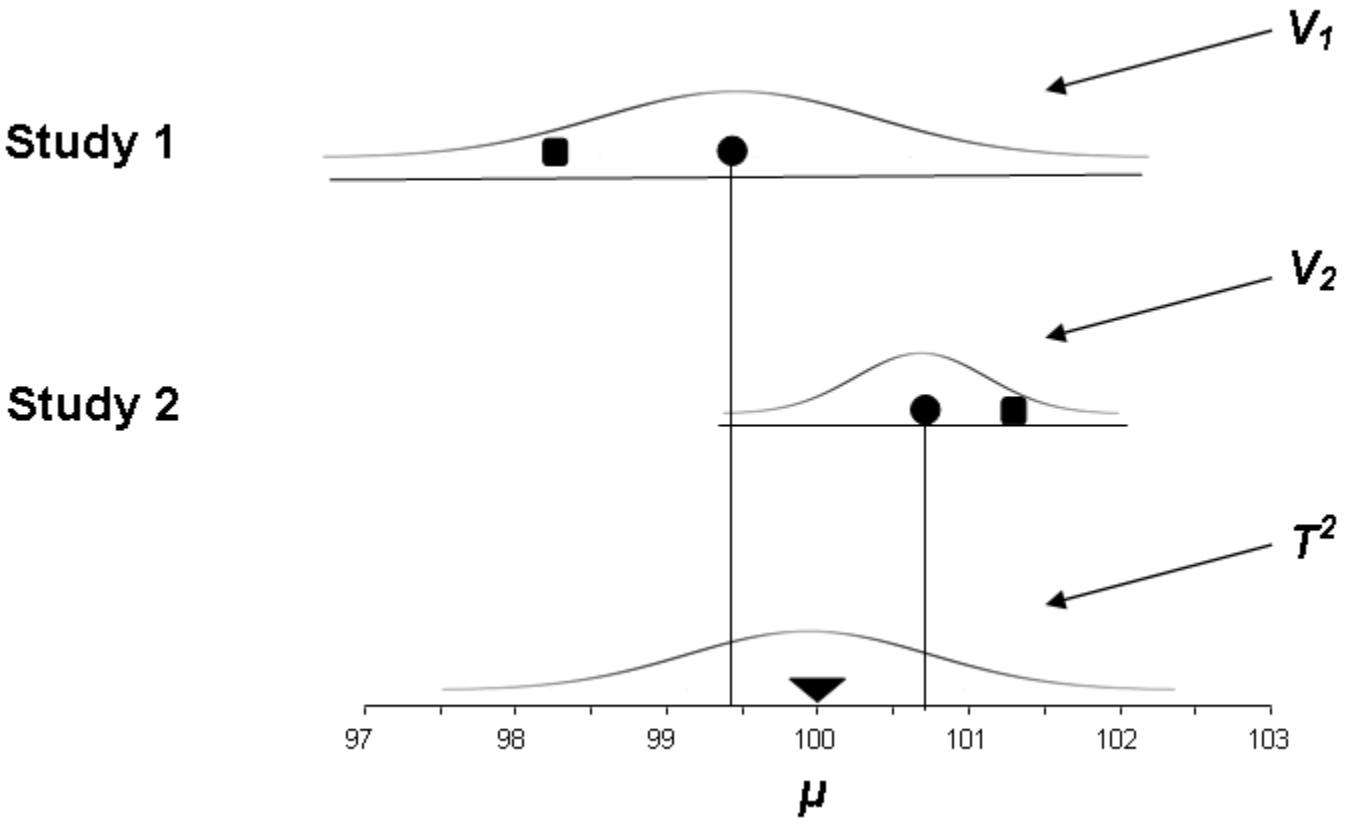
	Std Diff (*g*)	Relative Weight
Carroll	0.10	12%
Grant	0.28	13%
Peck	0.37	8%
Donat	0.66	39%
Stewart	0.46	10%
Young	0.19	18%
Summary	0.41	100%



Fixed effect model

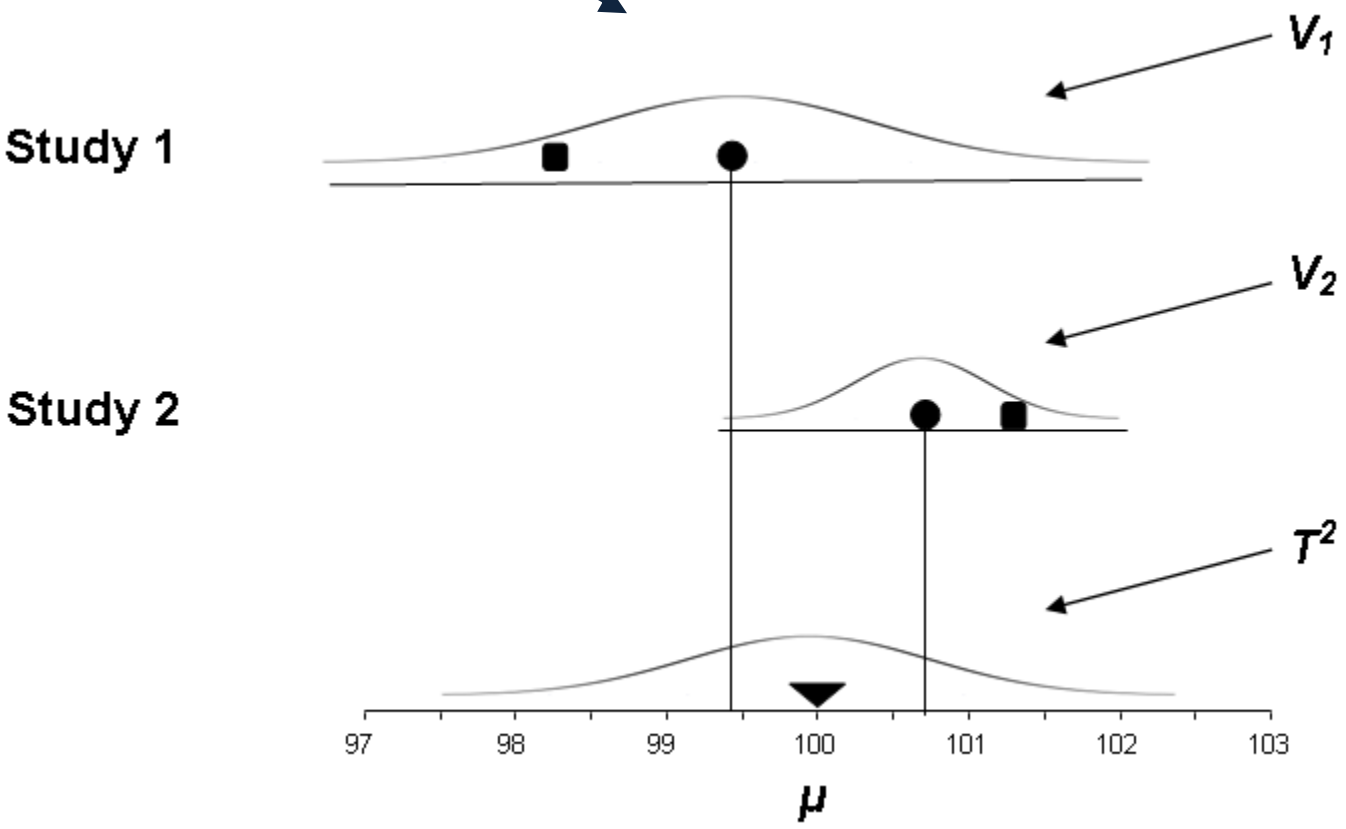
- Drug company example

Random-effects model



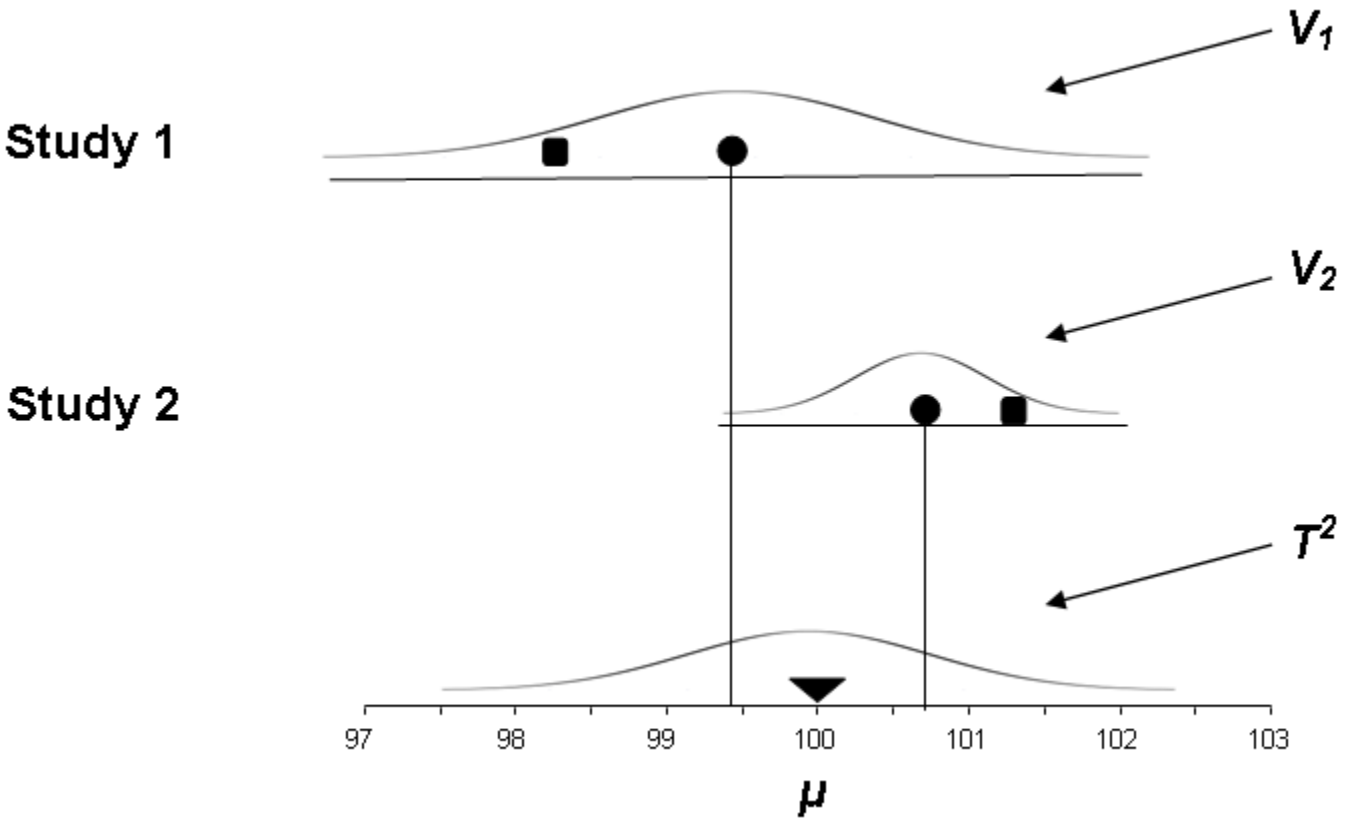
Random-effects model

$$V = \left(V_1 + T^2\right)$$



Random-effects model

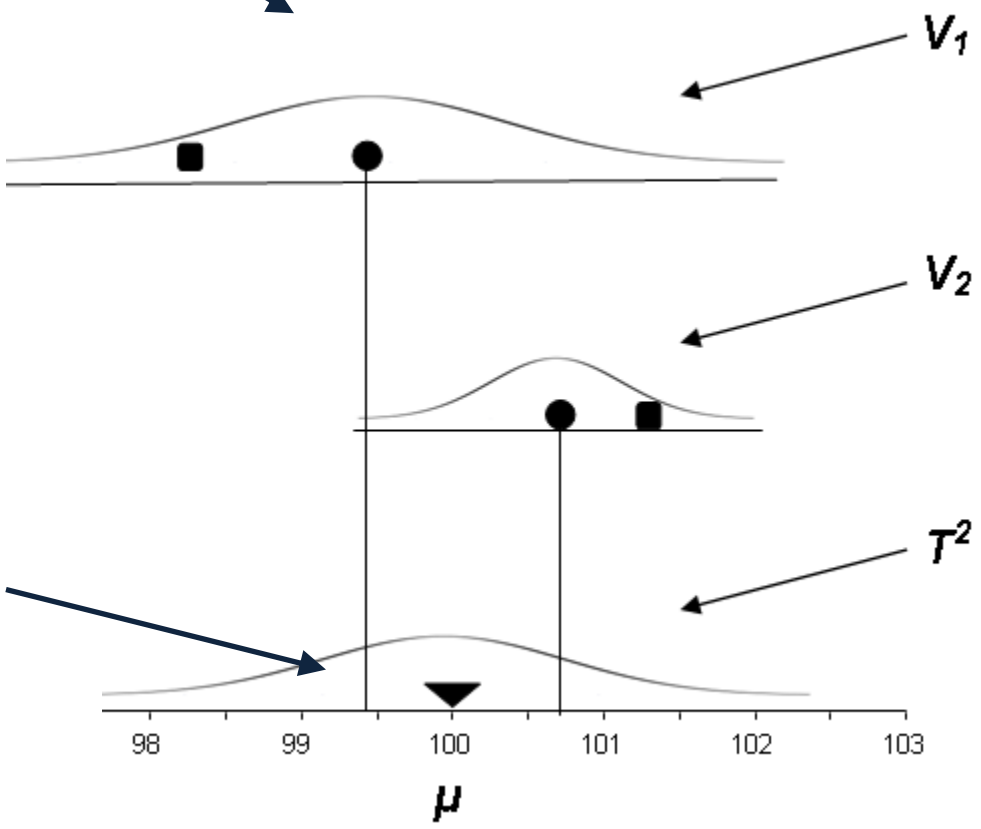
$$W = 1/\left(V_1 + T^2\right)$$



Random-effects model

$$W = 1/\left(V_1 + T^2\right)$$

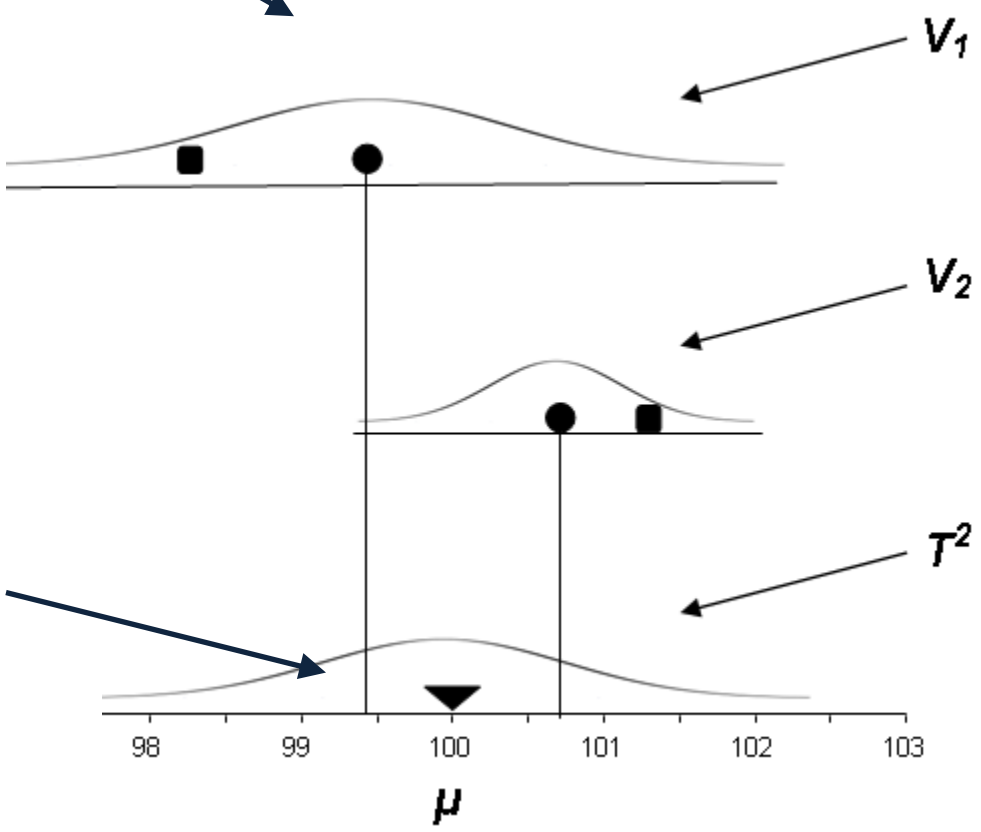
$$M = \frac{\sum_{i=1}^{k} W_i Y_i}{\sum_{i=1}^{k} W_i}$$



Random-effects model

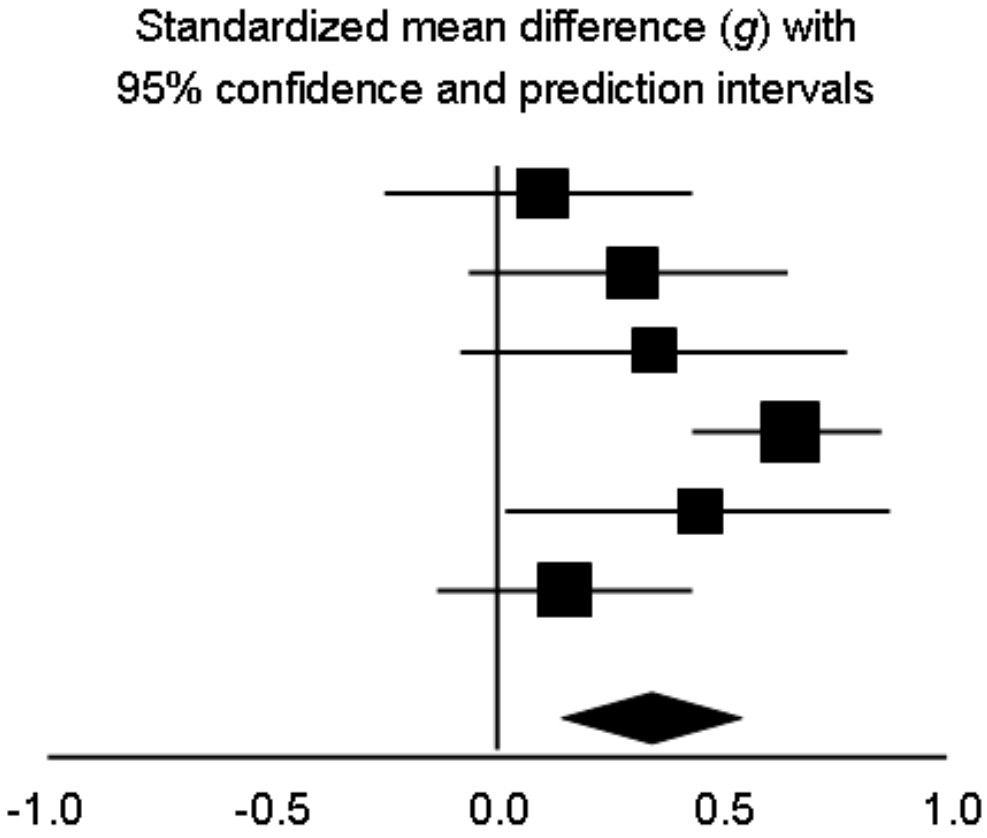
$$W = 1/\left(V_1 + T^2\right)$$

$$V_M = \frac{1}{\sum_{i=1}^{k} W_i}$$



Impact of Intervention (Random effects)

	Std Diff (*g*)	Relative Weight
Carroll	0.10	16%
Grant	0.28	16%
Peck	0.37	13%
Donat	0.66	23%
Stewart	0.46	14%
Young	0.19	18%
Summary	0.36	100%



Fixed-effect model
vs.
Random-effects model

Fixed effect

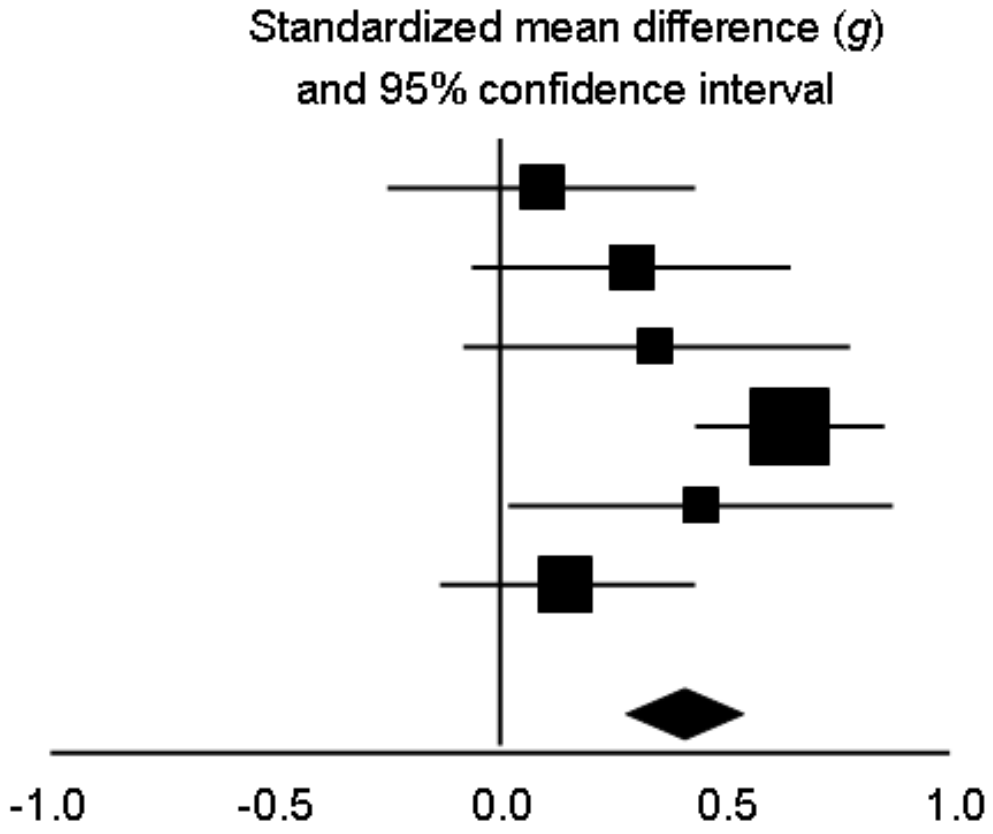
- When there is reason to believe that all the studies are functionally identical
- When our goal is to compute the common effect size, for the studies in the analysis
- Example of drug company has run five studies to assess the effect of a drug.

Random effects

- When not likely that all the studies are functionally equivalent.
- When the goal of this analysis is to generalize to a range of populations.
- Example of studies culled from publications

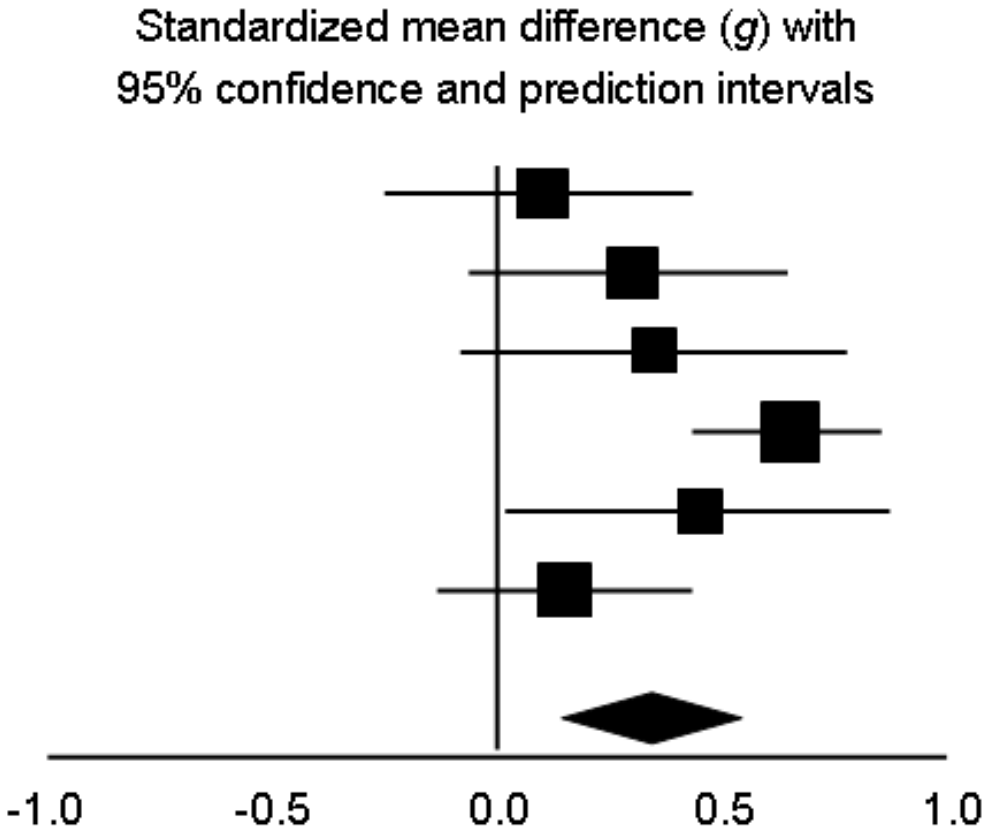
Impact of Intervention (Fixed effect)

	Std Diff (*g*)	Relative Weight
Carroll	0.10	12%
Grant	0.28	13%
Peck	0.37	8%
Donat	0.66	39%
Stewart	0.46	10%
Young	0.19	18%
Summary	0.41	100%



Impact of Intervention (Random effects)

	Std Diff (*g*)	Relative Weight
Carroll	0.10	16%
Grant	0.28	16%
Peck	0.37	13%
Donat	0.66	23%
Stewart	0.46	14%
Young	0.19	18%
Summary	0.36	100%



Using meta-analysis to compare subgroups

What kinds of *subgroups*

- Variants of the intervention
- Various types of patient
- Different times for follow-up
- Can also apply to continuous variables

Second-generation versus first-generation antipsychotic drugs for schizophrenia: a meta-analysis

Stefan Leucht, Caroline Corves, Dieter Arbter, Rolf R Engel, Chunbo Li, John M Davis

Summary

Background Because of the debate about whether second-generation antipsychotic drugs are better than first-generation antipsychotic drugs, we did a meta-analysis of randomised controlled trials to compare the effects of these two types of drugs in patients with schizophrenia.

Methods We compared nine second-generation antipsychotic drugs with first-generation drugs for overall efficacy (main outcome), positive, negative and depressive symptoms, relapse, quality of life, extrapyramidal side-effects, weight gain, and sedation.

Findings We included 150 double-blind, mostly short-term, studies, with 21 533 participants. We excluded open studies because they systematically favoured second-generation drugs. Four of these drugs were better than first-generation antipsychotic drugs for overall efficacy, with small to medium effect sizes (amisulpride −0·31 [95% CI −0·44 to −0·19, p<0·0001], clozapine −0·52 [−0·75 to −0·29, p<0·0001], olanzapine −0·28 [−0·38 to −0·18, p<0·0001], and risperidone −0·13 [−0·22 to −0·05, p=0·002]). The other second-generation drugs were not more efficacious than the first-generation drugs, even for negative symptoms. Therefore efficacy on negative symptoms cannot be a core component of atypicality. Second-generation antipsychotic drugs induced fewer extrapyramidal side-effects than did

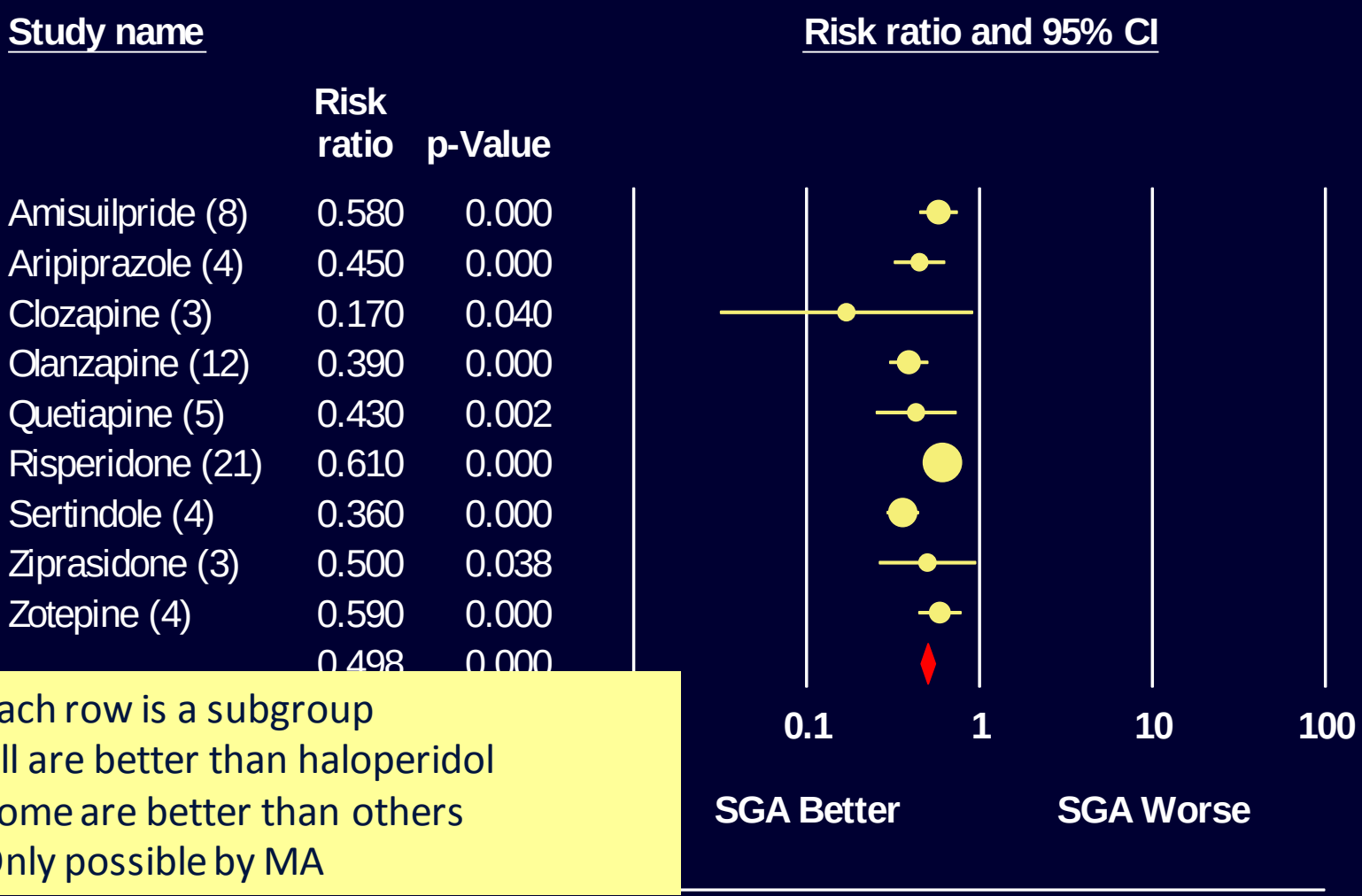
Disclaimer

- Please consult the original paper for context

Meta-analysis *provides a context*

- One SGA shows an advantage but –
- Is this true for all SGA?
- Is it true for all outcomes?
- Is it true for all populations?
- Metaphor of elephant and blind men

Antiparkisonian (vs. Haloperidol) $p = 0.004$

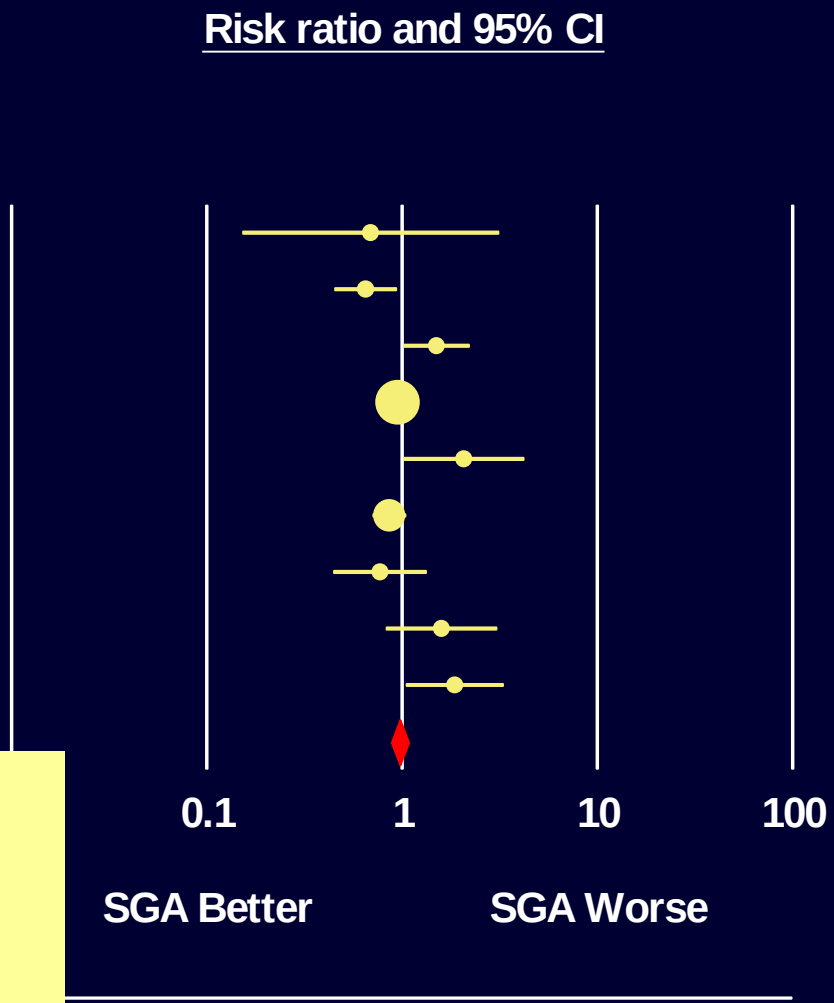


Each row is a subgroup
All are better than haloperidol
Some are better than others
Only possible by MA

Adapted from Leucht et al (2009)

Sedation (vs. Haloperidol) $p = 0.004$

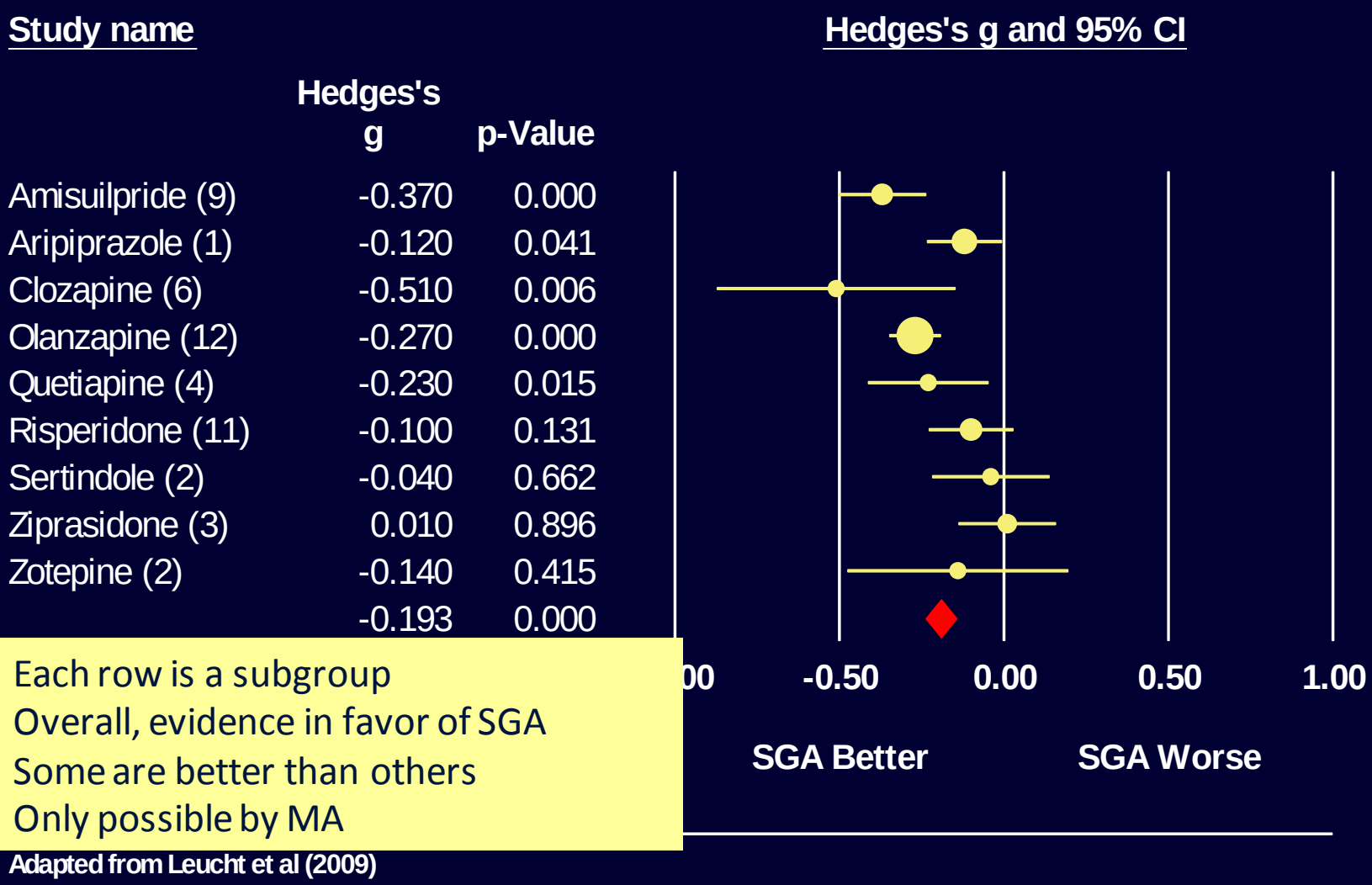
Study name	Risk ratio	p-Value
Amisuilpride (4)	0.690	0.632
Aripiprazole (2)	0.650	0.024
Clozapine (6)	1.500	0.045
Olanzapine (6)	0.950	0.494
Quetiapine (4)	2.070	0.048
Risperidone (15)	0.860	0.145
Sertindole (3)	0.770	0.358
Ziprasidone (1)	1.590	0.170
Zotepine (3)	1.860	0.037
	0.966	0.512



Each row is a subgroup
Overall, is no evidence of difference
Some are better than others
Only possible by MA

Adapted from Leucht et al (2009)

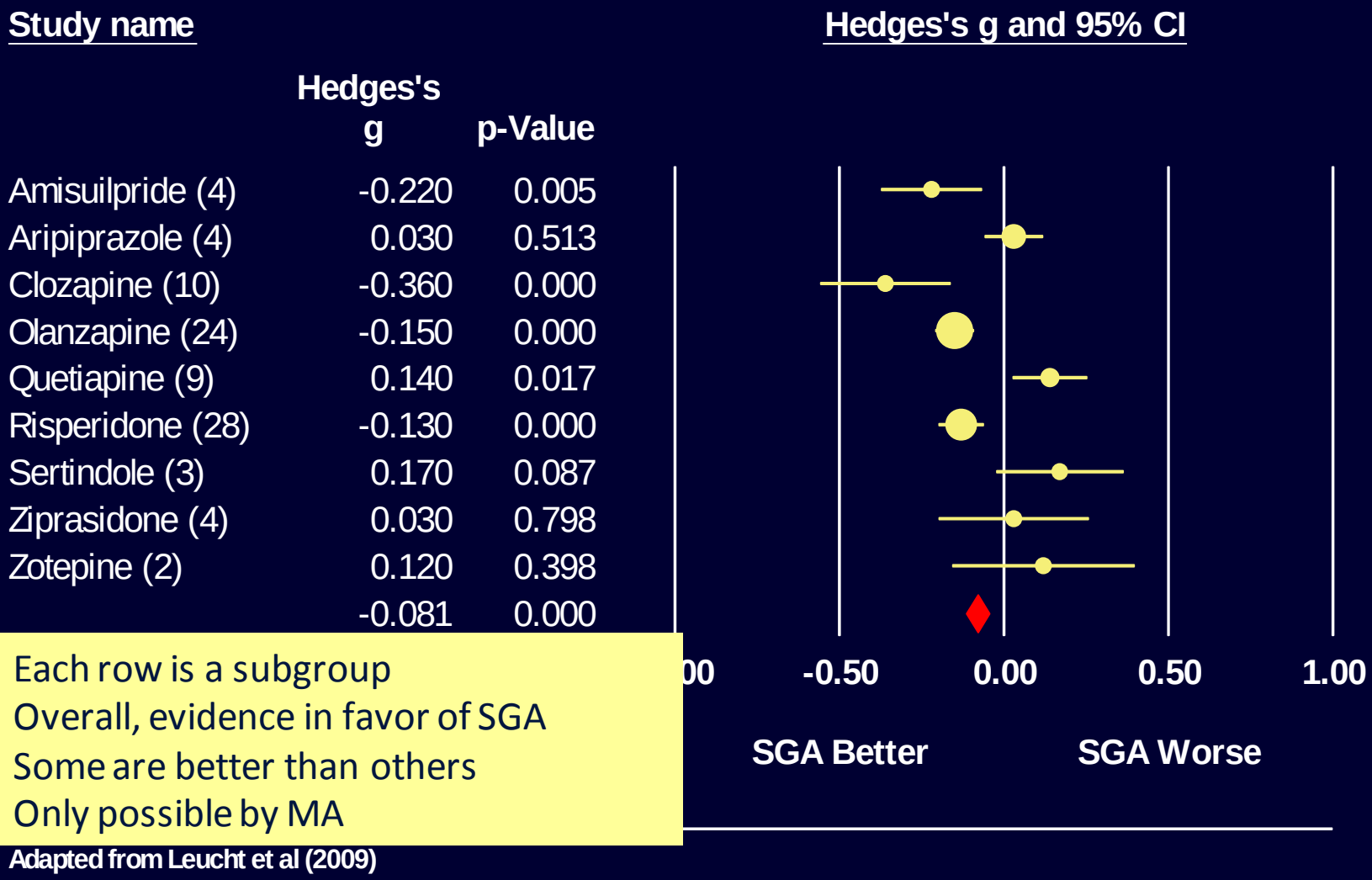
Depression (vs. FGA) $p = 0.001$



Each row is a subgroup
Overall, evidence in favor of SGA
Some are better than others
Only possible by MA

Adapted from Leucht et al (2009)

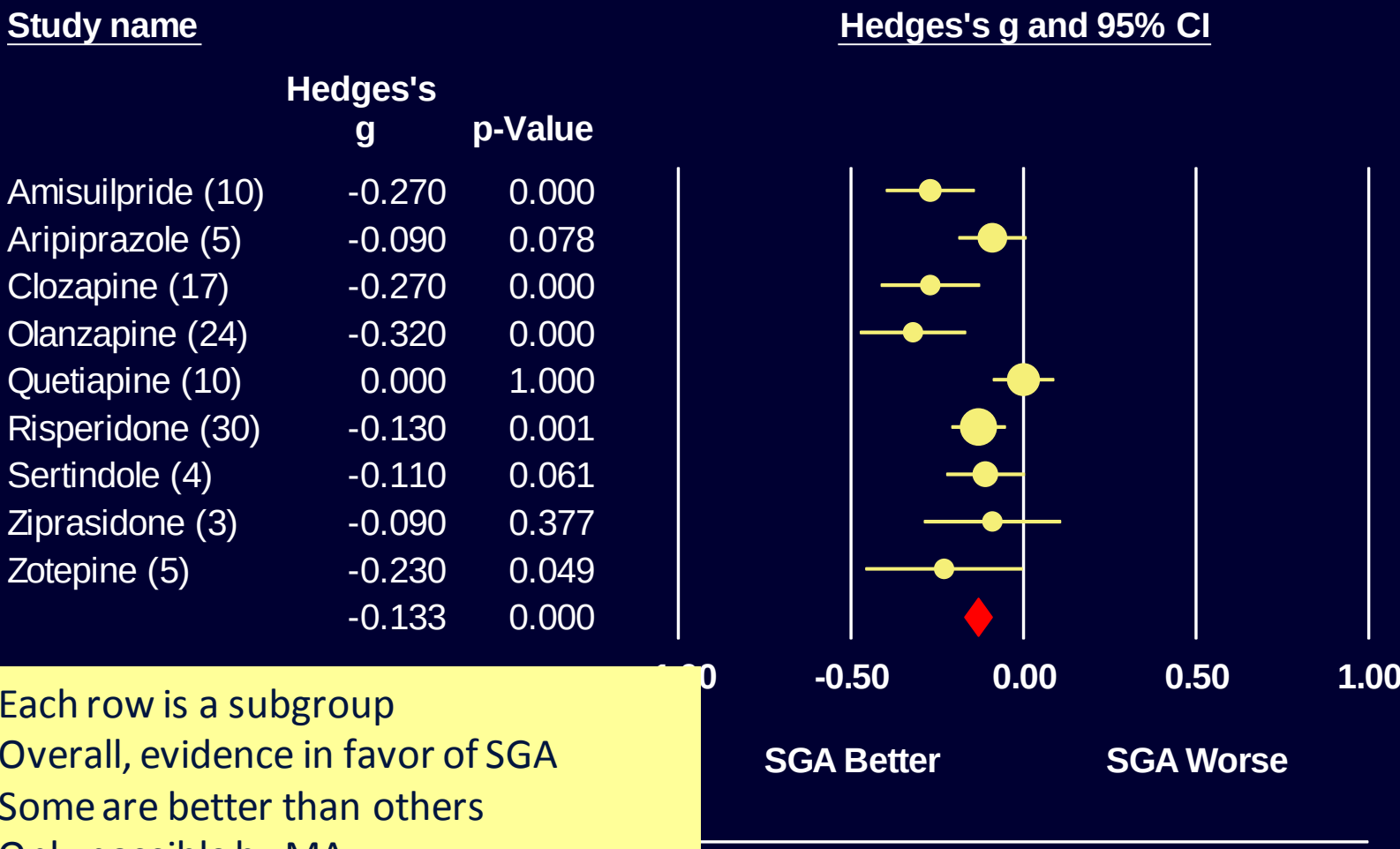
Positive Symptoms (vs. FGA) $p < 0.001$



Each row is a subgroup
Overall, evidence in favor of SGA
Some are better than others
Only possible by MA

Adapted from Leucht et al (2009)

Negative Symptoms (vs. FGA) $p = 0.003$



Each row is a subgroup
Overall, evidence in favor of SGA
Some are better than others
Only possible by MA

Sample size

Drug	Number Studies	Meta-Analysis	Single Study
Amisuilpride	13	1017	78
Aripiprazole	5	2049	410
Clozapine	23	1997	87
Olanzapine	28	4966	177
Quetiapine	11	2412	219
Risperidone	34	4173	123
Sertindole	4	1344	336
Ziprasidone	5	980	196
Zotepine	15	1125	75

Available online at www.sciencedirect.com
ScienceDirect
Clinical Psychology Review 28 (2008) 1021–1037
CLINICAL PSYCHOLOGY REVIEW

Psychological approaches in the treatment of specific phobias: A meta-analysis ☆

Kate B. Wolitzky-Taylor, Jonathan D. Horowitz, Mark B. Powers [1], Michael J. Telch *

Laboratory for the Study of Anxiety Disorders, Department of Psychology, The University of Texas at Austin, United States

Received 21 August 2007; received in revised form 11 February 2008; accepted 27 February 2008

Abstract

Data from 33 randomized treatment studies were subjected to a meta-analysis to address questions surrounding the efficacy of

Fear of enclosed spaces

Fear of heights

1. Introduction

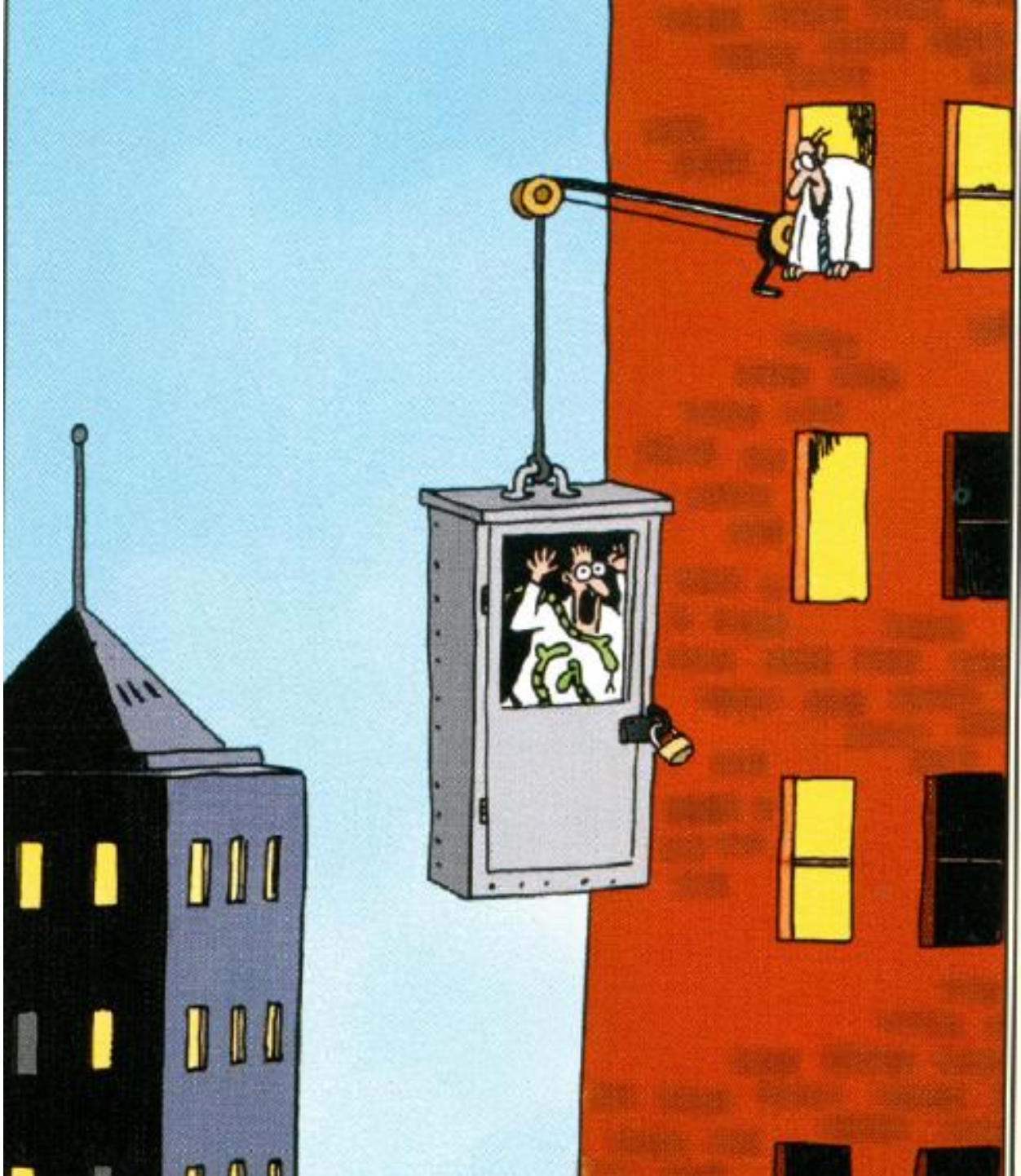
Specific phobia is characterized by a marked and persistent fear of a specific object or situation that causes significant life interference or distress (APA, 1994). With a lifetime prevalence of 12.5% (Kessler, Berglund, & Demler, 2005) specific phobia ranks as the most common anxiety disorder. Specific phobias are currently divided into four subtypes: situational (e.g., fears of enclosed spaces, flying), natural environment (e.g., fears of heights, storms, water), animal (e.g., fears of snakes, spiders, dogs), and blood/injection/injury (e.g. fears of dental or medical procedures, injections, seeing blood), with the animal and natural environment subtypes being more prevalent (Curtis, Magee, Eaton, Wittchen, & Kessler, 1998).

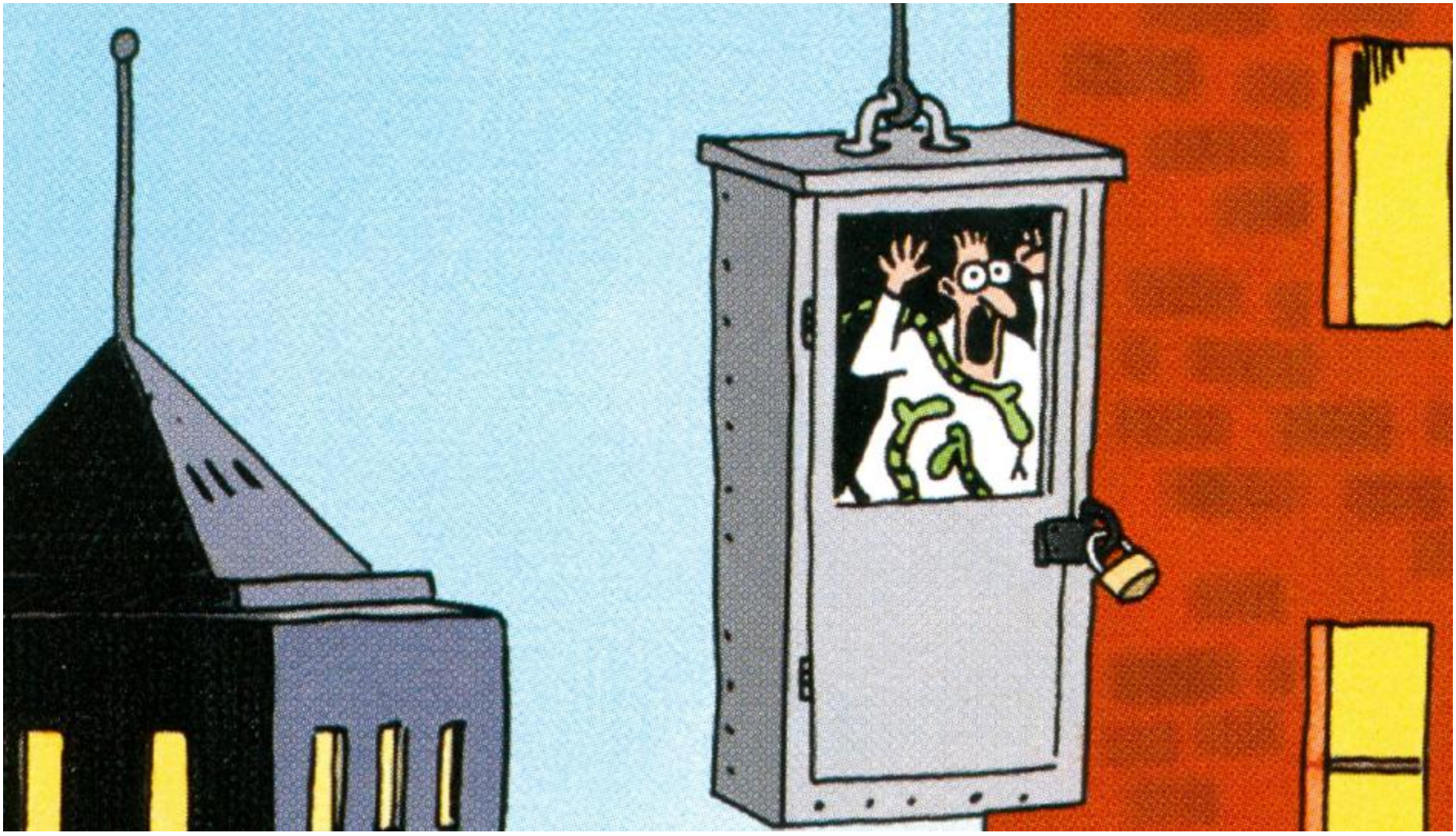
As with most anxiety disorders, specific phobias show a chronic course with low rates of spontaneous remission (Wittchen, 1988). Despite their circumscribed nature, specific phobia is associated with significant impairment. Wittchen, Nelosn, and Lachner (1998) found that young adults with a diagnosis of specific phobia reported severe impairment in their routine activities during the worst episode of their disorder. Specific phobias also represent a significant challenge to the medical field. For example, almost one-third of patients undergoing recumbent MRI (fMRI) are not able to complete the procedure due to severe claustrophobic reactions (Quirk,

Fear of snakes

Professor Gallagher and his controversial technique of simultaneously confronting the fear of enclosed spaces, heights, and snakes.

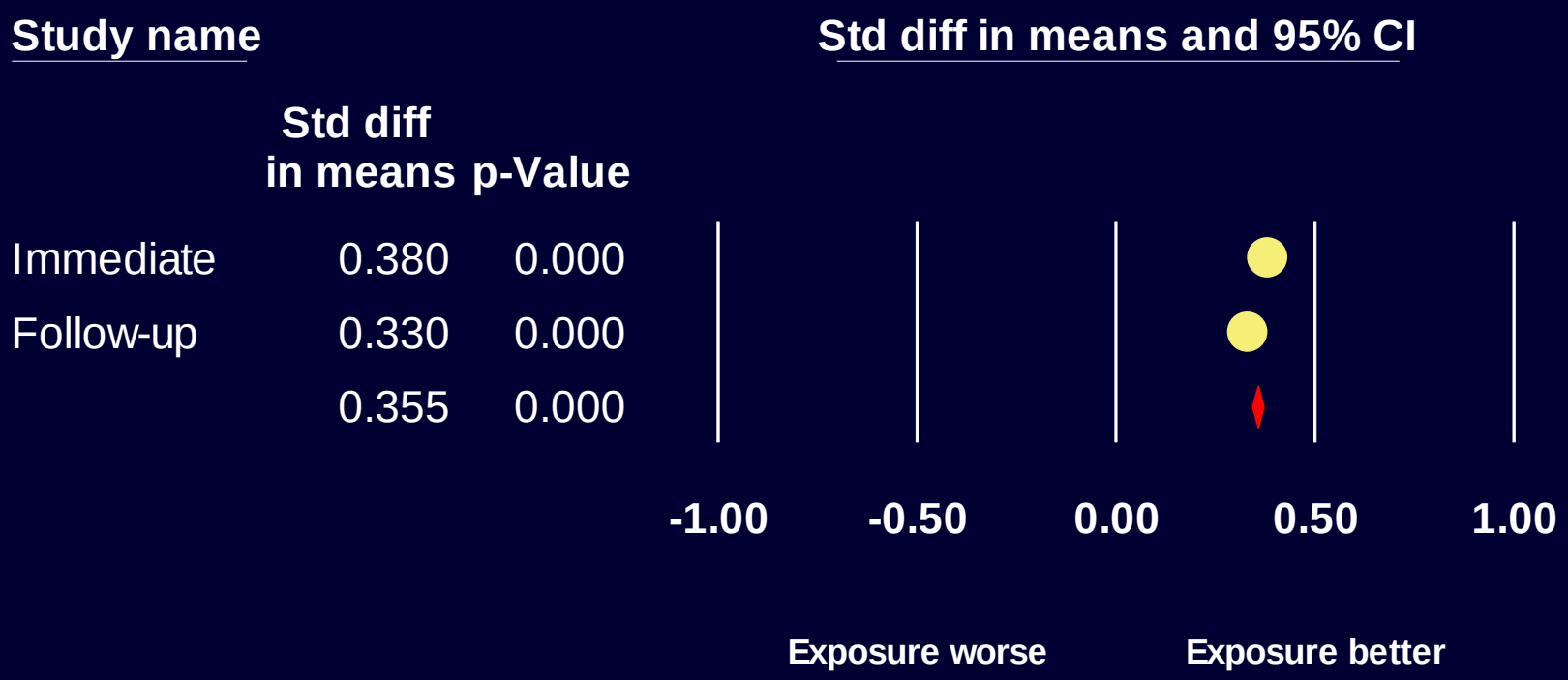
Professor Gallagher and his controversial technique of simultaneously confronting the fear of enclosed spaces, heights, and snakes.





with apologies to Gary Larson

Exposure vs. Non-Exposure for Phobias



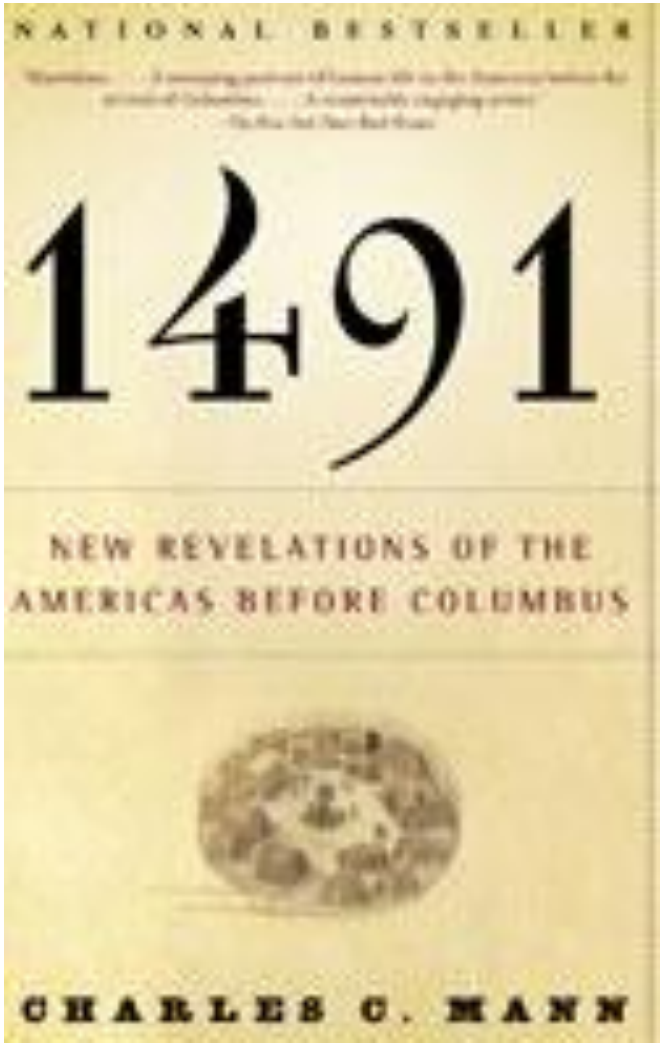
Adapted from Wolitzky-Taylor et al (2009)

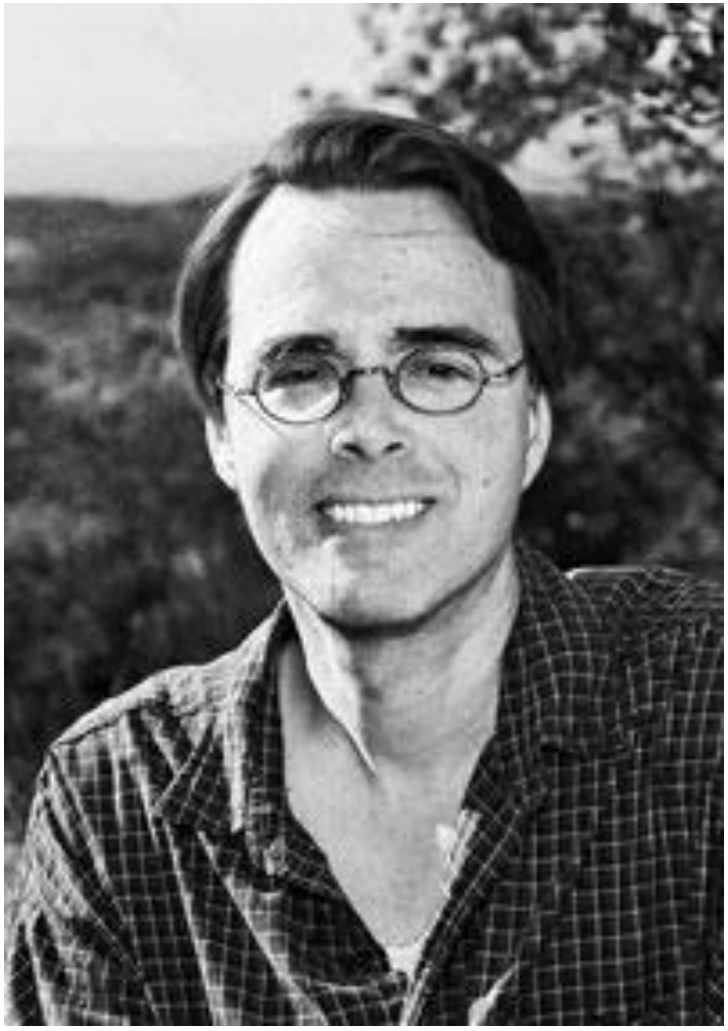
Criticisms of meta-analysis

SCIENCE

3 AUGUST 1990
VOLUME 249
NUMBER 4968

■ CHARLES MANN


NATIONAL BESTSELLER
1491
NEW REVELATIONS OF THE AMERICAS BEFORE COLUMBUS
CHARLES C. MANN



Meta-Analysis in the Breech

A controversial method for grouping results from disparate studies may ultimately revolutionize how research—particularly medical research—is done; for now, the fur is flying

WHEN PRINCESS EUGENIE WAS BORN to Sarah, Duchess of York, last 23 March, she arrived in the world feetfirst. Like many obstetricians, the royal *accoucheur* treated the princess's breech presentation by cesarean section—a practice long attacked by feminists, who claim that the ancient midwife's practice of externally "turning" the baby at term is safe, effective, painless, and less costly. For years, no definitive clinical study had been able to put an end to the controversy. But by the time of Eugenie's birth, a new element had entered the fray: a controversial statistical technique known as "meta-analysis."

Put briefly, meta-analysis is the use of formal statistical techniques to sum up a body of separate (but similar) experiments. It is like an ordinary scientific review of research, except that ordinary reviews provide a qualitative—and often subjective—assessment of a few studies; meta-analysis, on the other hand, promises a *quantitative* synthesis of all available data. "It's a boon for policy-makers who find themselves faced with a mountain of conflicting studies," says

tive Care in Pregnancy and Childbirth is the most extensive collection of meta-analyses yet compiled. The two-volume, 1516-page, $400 work reviews more than 3000 randomized controlled clinical trials in perinatal medicine. Along the way it bluntly rejects such procedures as routine episiotomy (cutting the tissue between the vagina and anus

pages." And one doctor called its authors "an obstetrical Baader-Meinhof gang."

The denunciations don't bother the moving spirits behind the report. "Some [obstetricians] hate it," says Iain Chalmers (no relation to Thomas), director of the National Perinatal Epidemiology Unit at Oxford and one of the report's three editors. "Of

Some Meta-analyses

Subject and year	Meta-analysts	No. of studies	Findings
Effects of desegregation on academic performance of black students (1982)	T. Cook, D. Armor, R. Crain, N. Miller, W. Stephan, H. Walberg, and P. Wortman	157	Desegregation has a tiny positive effect on reading scores and no effect on math scores. More important, formal analysis revealed glaring methodological weaknesses in all but 19 of the studies, suggesting that great effort had not succeeded in providing much of a database.
Measurement of gravitational constant (1930)	P. R. Heyl	3	Tests using different materials gave different results for the gravitational constant. An early meta-analysis, using simpler techniques than those of today, provided a more precise answer than any of the three studies taken alone.
Use of diagnostic nuclear magnetic resonance imaging (1988)	L. S. Cooper, T. C. Chalmers, M. McCally, J. Berrier, H. S. Sacks	54	At a time when nuclear magnetic resonance imaging was widely promoted as superior to computerized tomography, no good evidence existed for this belief.
Effect of coaching on SAT scores (1983)	R. DerSimonian and N. M. Laird	36	Coaching is only slightly effective. Interestingly, the probability that an experiment will find coaching to

Meta-Analysis in the Breech

A controversial method for grouping results from disparate studies may ultimately revolutionize how research—particularly medical research—is done; for now, the fur is flying

When princess Eugenie was born to Sarah, Duchess of York, last 23 March, she arrived in the world feet-first.

Meta-Analysis in the Breech

A controversial method for grouping results from disparate studies may ultimately revolutionize how research—particularly medical research—is done; for now, the fur is flying

Like many obstetricians, the royal *accoucheur* treated the princess's breech presentation by caesarian section – a practice long attacked by feminists, who claim that the ancient midwife's practice of externally turning the baby at term is safe, effective, painless, and less costly.

Meta-Analysis in the Breech

A controversial method for grouping results from disparate studies may ultimately revolutionize how research—particularly medical research—is done; for now, the fur is flying

A remarkable report, published at the end of last year in the United Kingdom, offers a window into the possibilities raised by meta-analysis.

Meta-Analysis in the Breech

A controversial method for grouping results from disparate studies may ultimately revolutionize how research—particularly medical research—is done; for now, the fur is flying

Effective care in pregnancy and childbirth is the most extensive collection of meta-analyses yet compiled.

The two-volume, 1516 page, $400 work reviews more than 3000 randomized controlled clinical trials in perinatal medicine.

Meta-Analysis in the Breech

A controversial method for grouping results from disparate studies may ultimately revolutionize how research—particularly medical research—is done; for now, the fur is flying

The book has triggered an extraordinary range of reactions in the medical profession.

Meta-Analysis in the Breech

A controversial method for grouping results from disparate studies may ultimately revolutionize how research—particularly medical research—is done; for now, the fur is flying

"Meta-Analysis is the wave of the future" says Thomas Chalmers, a former president of Mount Sinai Hospital who is now at the Harvard School of Public Health.

Meta-Analysis in the Breech

A controversial method for grouping results from disparate studies may ultimately revolutionize how research—particularly medical research—is done; for now, the fur is flying

"The days of the expert supposedly putting the state of the field into a review article are numbered."

Meta-Analysis in the Breech

A controversial method for grouping results from disparate studies may ultimately revolutionize how research—particularly medical research—is done; for now, the fur is flying

The *Medical Journal of Australia* described it as "arguably the most important publication in obstetrics since William Smellie wrote *A Treatise on the Theory and Practice of Midwifery* in 1752".

Meta-Analysis in the Breech

A controversial method for grouping results from disparate studies may ultimately revolutionize how research—particularly medical research—is done; for now, the fur is flying

But it was denounced by the editor of the Journal of Obstetrics and Gynaecology :

Meta-Analysis in the Breech

A controversial method for grouping results from disparate studies may ultimately revolutionize how research—particularly medical research—is done; for now, the fur is flying

"The price of £225 should protect aspiring registrars [residents] from acquiring too many confused ideas from its pages."

Meta-Analysis in the Breech

A controversial method for grouping results from disparate studies may ultimately revolutionize how research—particularly medical research—is done; for now, the fur is flying

And one doctor called its authors “an obstetrical Baader-Meinhof gang”.

META-ANALYSIS: STATISTICAL ALCHEMY FOR THE 21st CENTURY

ALVAN R. FEINSTEIN*
Yale University School of Medicine, Clinical Epidemiology Unit, New Haven, CT 06510, U.S.A.

"Statistical Alchemy for the 21st Century"

Enormous
devoted t
acceptable evidence in modern science:

(1) The substances under investigation should be identified in a reproducible manner.
(2) Each substance should receive a precise characterization, which indicates its specific distinctiveness. To achieve this goal, scientists search for substances that have been suitably purified, and try to work with mixtures that are homogeneous, not

alysis, as
moval or
ents that
nd estab-
lished during the 19th and 20th centuries. In the mixtures formed for most statistical meta-analyses, we lose or eliminate the elemental scientific requirements for reproducibility and precision, for suitable extrapolation, and even sometimes for fair comparison.

ANALOGY TO ALCHEMY

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 the reader to study two papers of mine, one of which I have no recollection of writing, and which the reader will look in vain for in the journal app
ences—ind
number giv
agree! T
"astute di
myth" (p.
tioning that the Bergin myth has in turn been astutely dismantled by Rachman (1971). No discussion of the issue can be regarded as meaningful which accepts the quite erroneous and—indeed, in places—absurd arguments of Bergin and pays no attention to the serious criticisms brought forward by Rachman. Indeed, the latter is not even mentioned in the bibliography, although his book *The Effects of Psychotherapy* is a classic

complex statistics offered. Smith and Glass also do not mention the problem of selection, so well discussed by Rachman; patients for psychoanalysis are much more highly selected (for high intelligence, emotional resources, ego strength, etc.) than are patients for behavior therapy, and hence much more likely to improve spontaneously.

The most surprising feature of

and indifferent—are fed into the computer in the hope that people will cease caring about the quality of the material on which the conclusions are based. If their abandonment of scholarship were to be taken seriously, a daunting but improbable likelihood, it would mark the beginning of a passage into the dark age of scientific psychology.

The notion that one can distill scientific knowledge from a compilation

for each method in order to obtain evidence on the therapist variance. I would suggest that there is no single study in existence which does not show serious weaknesses, and until these are overcome I must regretfully restate my conclusion of 1952, namely that there still is no acceptable evidence for the efficacy of psychotherapy.

B., Yalom,
tcome in
indepen-
al of Psy-

psychother-
apy. London: Pergamon Press, 1971.
Smith, M. L., & Glass, G. V. Meta-analysis of psychotherapy outcome studies. *American Psychologist*, 1977, *32*, 752–760.

H. J. EYSENCK
Institute of Psychiatry
London, England

Reply to Eysenck

We have numbered our responses to

"An Exercise in Mega-Silliness"

American Journal of
EPIDEMIOLOGY

Meta-analysis / Shmeta-analysis

other

Meta-analysis/Shmeta-analysis

Samuel Shapiro

Slone Epidemiology Unit 1371 Beacon Street, 3rd Floor, Brookline, MA 02146


©C.P. CORP. D-COL-422-7

Criticisms of meta-analysis

1. Cannot summarize results in one number
2. Apples and oranges
3. Meta-analyses are performed poorly

The NEW ENGLAND JOURNAL of MEDICINE

EDITORIAL

◂ Previous Volume 337:559-561 August 21, 1997 Number 8 Next ▸

The Promise and Problems of Meta-Analysis

Meta-analysis has acquired a substantial following among both statisticians and clinicians. The technique was developed as a way to summarize the results of different research studies of related problems. Meta-analysis may be applied even when the studies are small and there is substantial variation in the specific issues studied, the research methods applied, the

COMMENTARY

- Letters

TOOLS & SERVICES

- Add to Personal Archive
- Add to Citation Manager
- Notify a Friend
- E-mail When Cited

MORE INFORMATION

John C Bailar III

The NEW ENGLAND JOURNAL of MEDICINE

EDITORIAL

"Any attempt to reduce results to a single value, with confidence bounds, is likely to lead to conclusions that are wrong, perhaps seriously so"

Meta-a

among

was de

differen

analysis

small and there is substantial variation in the specific issues studied, the research methods applied, the

RY

ICES

rchive

anager

▸ E-mail When Cited

MORE INFORMATION

Provides Context

- Is the effect robust?
- Does the effect vary?

If consistent

- Report that the effect is robust
- More accurate estimate of the effect
- More powerful test

If varies

- Report that the effect varies
- More accurate estimate of the mean effect
- Estimate of the dispersion
- Possibly more powerful test

Apples and Oranges

Criticism

- The review may include
  - Different populations
  - Different protocols
  - Different study designs
- The summary effect may ignore important differences among these

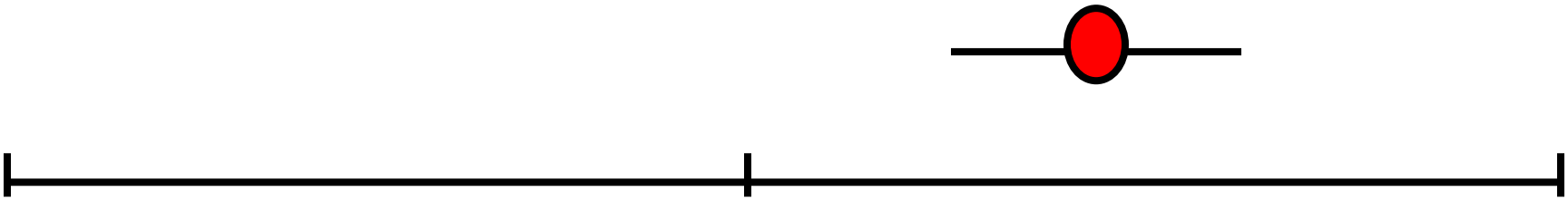
Response

- The analysis may include apples and oranges, and indeed, melons and pears
- But, used properly, this is a strength rather than a weakness, of the analysis

Apples and Oranges

- Ability to work with a range of studies is a key strength rather than a weakness of meta analysis.
- Using meta analysis we can actually assess hypothesis that effect is influenced by population design, etc.
- We cannot do these kinds of tests with a single study.

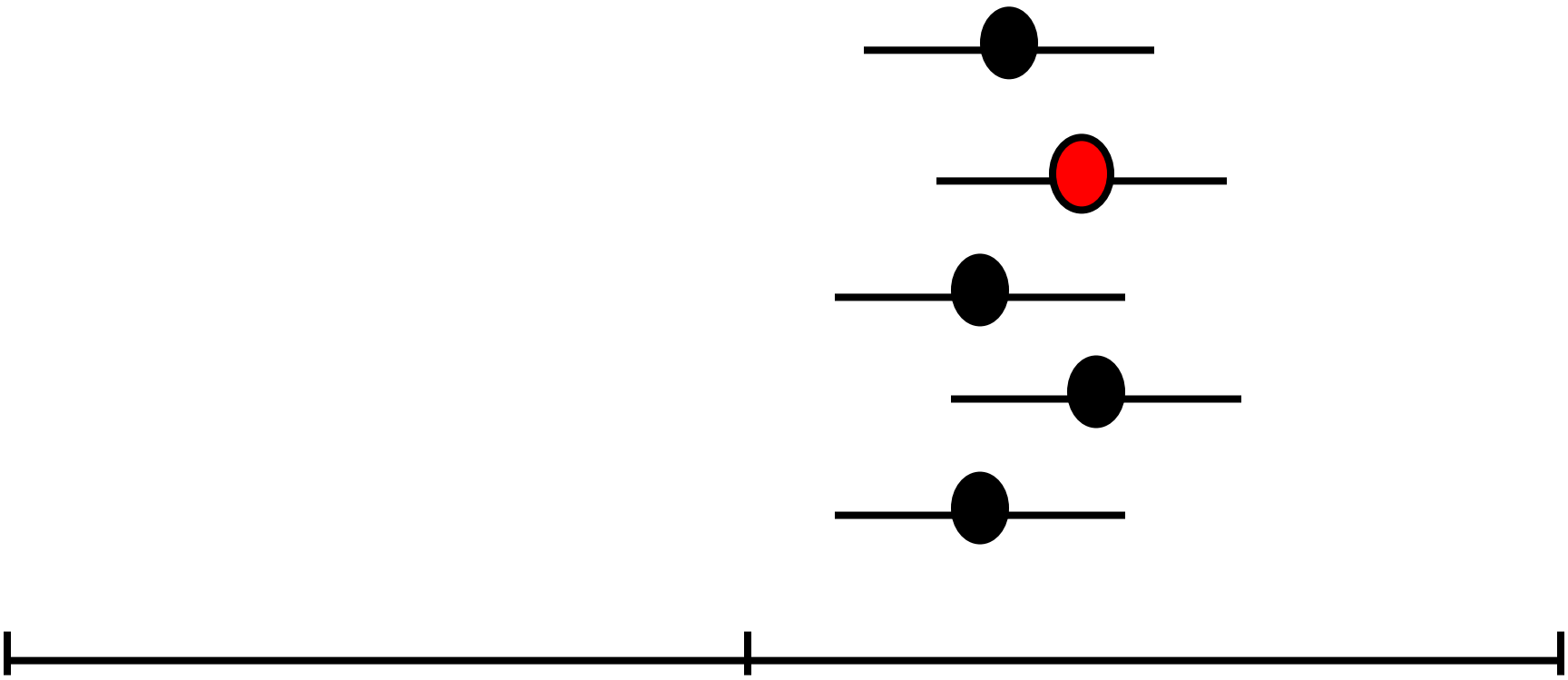
One study



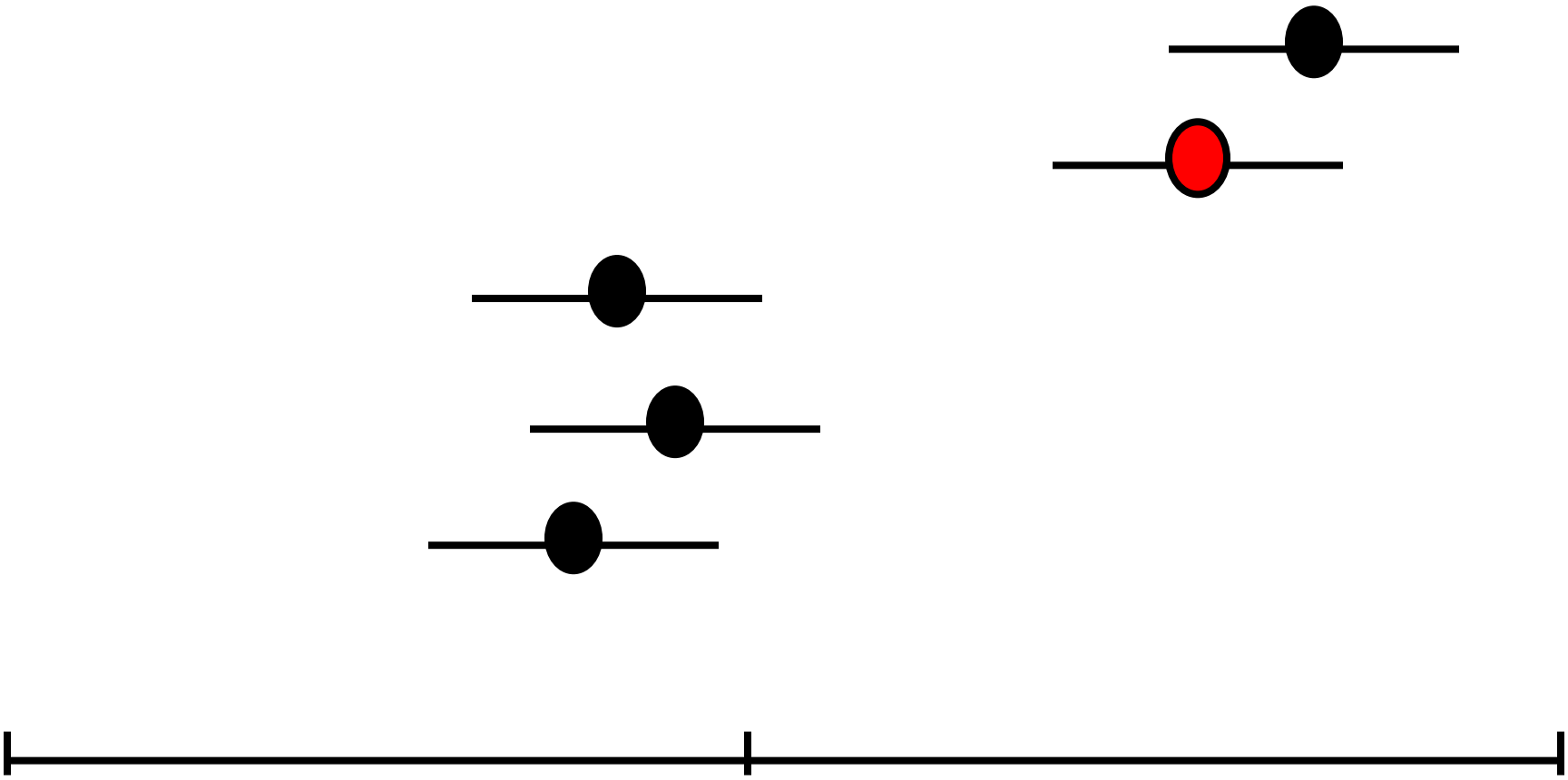
Can only speculate about the effect for

- An alternate intervention
- An alternate comparison group
- An alternate time frame
- And so on

With multiple studies we can report that effect is robust ...



… or that it varies



Analyses are done poorly

The NEW ENGLAND JOURNAL of MEDICINE

EDITORIAL

◂ Previous Volume 337:559-561 August 21, 1997 Number 8 Next ▸

The Promise and Problems of Meta-Analysis

Meta-analysis has acquired a substantial following among both statisticians and clinicians. The technique was developed as a way to summarize the results of different research studies of related problems. Meta-analysis may be applied even when the studies are small and there is substantial variation in the specific issues studied, the research methods applied, the

COMMENTARY

- Letters

TOOLS & SERVICES

- Add to Personal Archive
- Add to Citation Manager
- Notify a Friend
- E-mail When Cited

MORE INFORMATION

Criticism

- Failed to understand basic issues
- Were careless in summarizing data
- Failed to consider covariates
- Overstated strength and precision

The NEW ENGLAND JOURNAL of MEDICINE

EDITORIAL

"I still prefer conventional narrative reviews of the literature, a type of summary familiar to readers of the countless review articles on important medical issues."

Meta-a

among

was de

differen

analysis

small and there is substantial variation in the specific issues studied, the research methods applied, the

RY

ICES

rchive

anager

E-mail When Cited

MORE INFORMATION

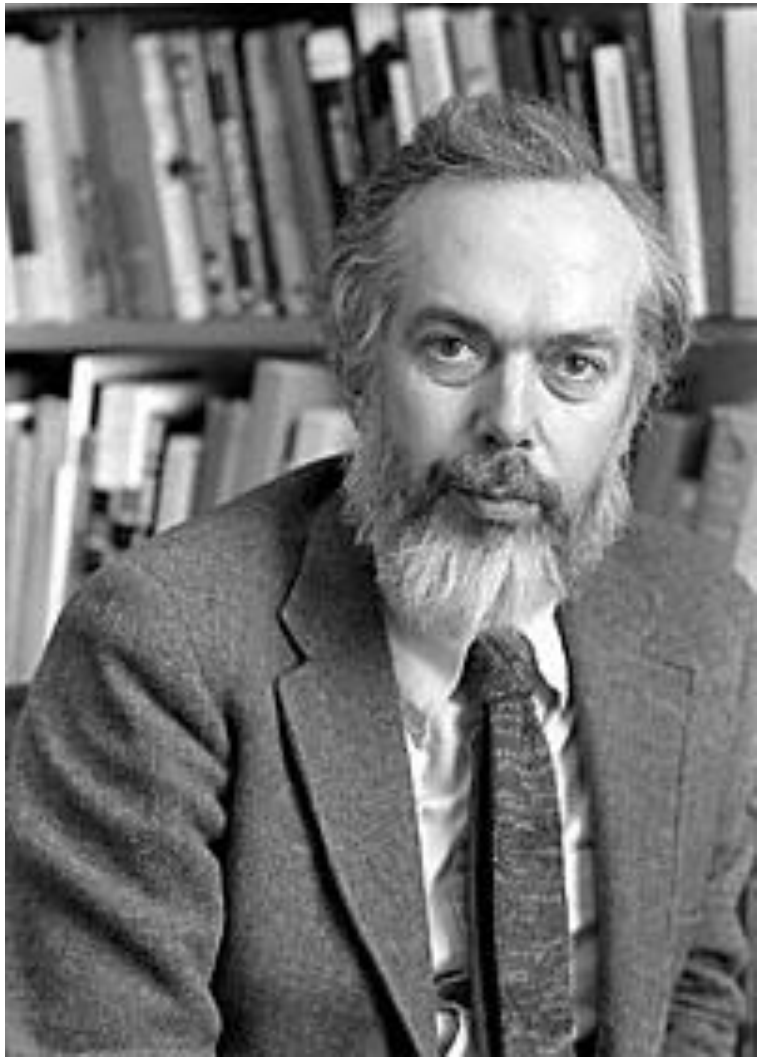
Systematic vs. Narrative Reviews

1. Cannot summarize results in one number
2. Apples and oranges
3. Meta-analyses are performed poorly

And –

- Need to distinguish
  - Problems inherent in the method
  - Problems with the way the method is being used

Bob Abelson



Bob Abelson

- Consider the list of things people abuse

Abelson

- Consider the list of things people abuse
- College educations

Abelson

- Consider the list of things people abuse
- College educations
- Oboes

The Cochrane Library

Evidence for healthcare decision-making

Search Results

Show Results in:

Cochrane Reviews [21] | Other Reviews [6] | Clinical Trials [348] | Methods Studies [4] | Technology Assessments [8] | Economic Evaluations [10] | Cochrane Groups [0]

There are **21** results out of **5821 records** for: "**opioid dependence** in **Title, Abstract or Keywords** in **Cochrane Database of Systematic Reviews**"

View: **1-21**

Export All Results

Record Information — Restrict to: Reviews | Prot

- **Methadone maintenance at different dosages for opioid dependence**
  Fabrizio Faggiano, Federica Vigna-Taglianti, Elisabetta Versino, Patrizia Lemma
  Year: 2003
  Record **Review**

- **Sustained-Release Naltrexone For Opioid Dependence**
  Philipp Lobmaier, Hege Kornor, Nikolaj Kunoe, Arild Bjørndal
  Year: 2008
  Record **Review**

E-Mail
Michael@Meta-Analysis.com