Using meta-analysis to inform clinical trial design

Julian Higgins
MRC Biostatistics Unit, Cambridge

With particular thanks to Verena Roloff
Drawing on numerous papers by Alex Sutton, Nicola Cooper and colleagues

MRC | Biostatistics Unit

THE COCHRANE COLLABORATION®

Outline

- Meta-analysis and the decision whether to do a new trial
- Ways in which meta-analysis can be used in the design of a trial
- Cumulative and sequential meta-analysis
- Powering a trial using meta-analysis

May 1986

- “One of the sources of controversy of the breast cancer overviews was concern that such overviews would undermine the ability and motivation to conduct major multi-centre clinical trials.”

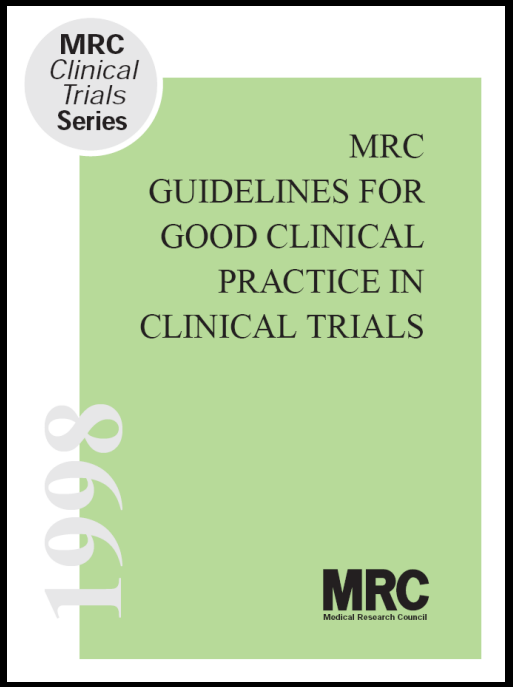
Richard Simon
Workshop on Methodologic Issues in Overviews of Randomized Trials

The importance of building on existing evidence

"New research should not be designed or implemented without first assessing systematically what is known from existing research. The failure to conduct that assessment represents a lack of scientific self-discipline that results in an inexcusable waste of public resources. In applied fields like health care, failure to prepare scientifically defensible reviews of relevant animal and human data results not only in wasted resources but also in unnecessary suffering and premature death."

Iain Chalmers. The scandalous failure of science to cumulate evidence scientifically. *Clinical Trials* 2005: **2**; 229-231

Systematic reviews, meta-analyses and clinical trial funding



appendix 2

MRC PROFORMA APPLICATION FORM

MRC CONTROLLED TRIALS 1997 -1998

PROFORMA FOR FULL PROPOSALS

Please structure Annex 1 of your application form using the headings listed below

- Please make an entry under every heading
- Do not exceed 9 sides of A4 (10 point)

1 FULL TITLE OF TRIAL

1.1 ACRONYM (only if applicable – this is not a requirement)

1.2 CONTACT APPLICANT (name, address, tel, fax, e-mail)

2 THE NEED FOR A TRIAL

2.1 WHAT IS THE PROBLEM TO BE ADDRESSED?

2.2 WHAT IS THE HYPOTHESIS TO BE TESTED?

2.3 WHY IS A TRIAL NEEDED NOW?

2.4 HAS A SYSTEMATIC REVIEW BEEN CARRIED OUT AND WHAT WERE THE FINDINGS?

2.5 HOW WILL THE RESULTS OF THIS TRIAL BE USED?
(eg, inform clinical decision making /improve understanding)

2.6 PLEASE DETAIL ANY RISKS TO THE SAFETY OF PARTICIPANTS INVOLVED IN THE TRIAL

Need for evidence

Outline proposals will first be prioritised for NHS need, using the following criteria.

1. *The importance of the health problem to patients and the NHS.* Applicants should describe the burden (frequency and severity) of the health problem in the population and the potential benefit from the technology.
2. *The relevance of study outcomes to patients and the NHS, and the relevance of participants to the case mix treated in the NHS.* Clinically important outcomes that matter to patients and that measure health gain should be used. These will usually be long-term. Widely accepted surrogate markers may be used if they are strongly linked to health outcomes. For primary research, participants should reflect the mix of patients likely to be seen in normal clinical practice.
3. *Justification of proposed research with reference to the current evidence base.* The importance of the research question to the NHS should be explained. Researchers should describe the current level of uncertainty and how their research will reduce it. This should include an account of the existing evidence, and any relevant research being undertaken in the HTA programme and elsewhere. The applicants should consider evidence in related technologies, diseases or patient groups when justifying their proposal. A systematic review should normally have been undertaken before a trial is considered.
4. *The technology assessment is relevant to the NHS.* There should be an adequate description of the technology and its possible effectiveness range. It must be one that is used in the NHS, or could be adopted into the NHS following the study. The study should usually assess cost-effectiveness in the NHS or justify this omission.

Efficacy and Mechanism Evaluation programme

MRC Medical Research Council

NHS National Institute for Health Research

IMPORTANT INFORMATION & GUIDANCE NOTES – PRELIMINARY APPLICATION

There are two different Efficacy and Mechanism Evaluation (EME) applications – PRELIMINARY application and FULL proposal.

The EME programme accepts PRELIMINARY applications in the first instance.

This document contains information and guidance to applicants submitting a PRELIMINARY application and is comprised of three parts;

- **Part One – Information about the EME programme** (pages 2-4)
- **Part Two - Com**
- **Part Three – Su**

Applications which are n

Applications must be r
considered at the June 2

Section F: Project Details and Justification

F1. Remit

Please explain how your proposed research is within the remit of the EME programme. You should include a clear explanation of the main (single) research question phrased in PICO terms (Population; Intervention; Comparator; Outcome). Give a brief explanation of how or in what ways the design constitutes a clinical trial or evaluation study. You are welcome to highlight any other aspects of the design that you would like to bring particular attention to, in order to explain how it is within remit. Please remember that EME research looks at patients or people seeking healthcare; studies using healthy volunteers and animals are not within the remit of the programme.

F2. Background and References

Please provide a clear explanation of the health problems to be addressed, the impact on patients and healthcare, an explanation of the scientific principles of the proposed research and an overview of the potential economic benefits (you are not required to include health economics analysis within your research). You should give reference to any relevant systematic reviews and discuss the need for your trial in light of these. If you believe that no relevant previous trials have been done, give details of your search strategy for existing trials. Please give details of other trials currently underway, both nationally and internationally, which are relevant to the proposed study. Please explain why this trial is needed now. References should be provided in the Vancouver format *(Author(s). Title. Journal. Year; Volume: Start page - End page).*

Using meta-analysis to design trials

- Meta-analyses may contribute:
  - anticipated effect size
    - example: early beta blockade therapy
      - MIAMI trial sized using pilot of 1400 people: 36% reduction in mortality
      - ISIS-1 trial sized using meta-analysis of 16,000 people in 21 trials: 10% reduction in mortality
      - results: reductions in vascular mortality were
        - » 13% (MIAMI) [not significant because too small]
        - » 15% reduction (ISIS-1) [statistically significant]

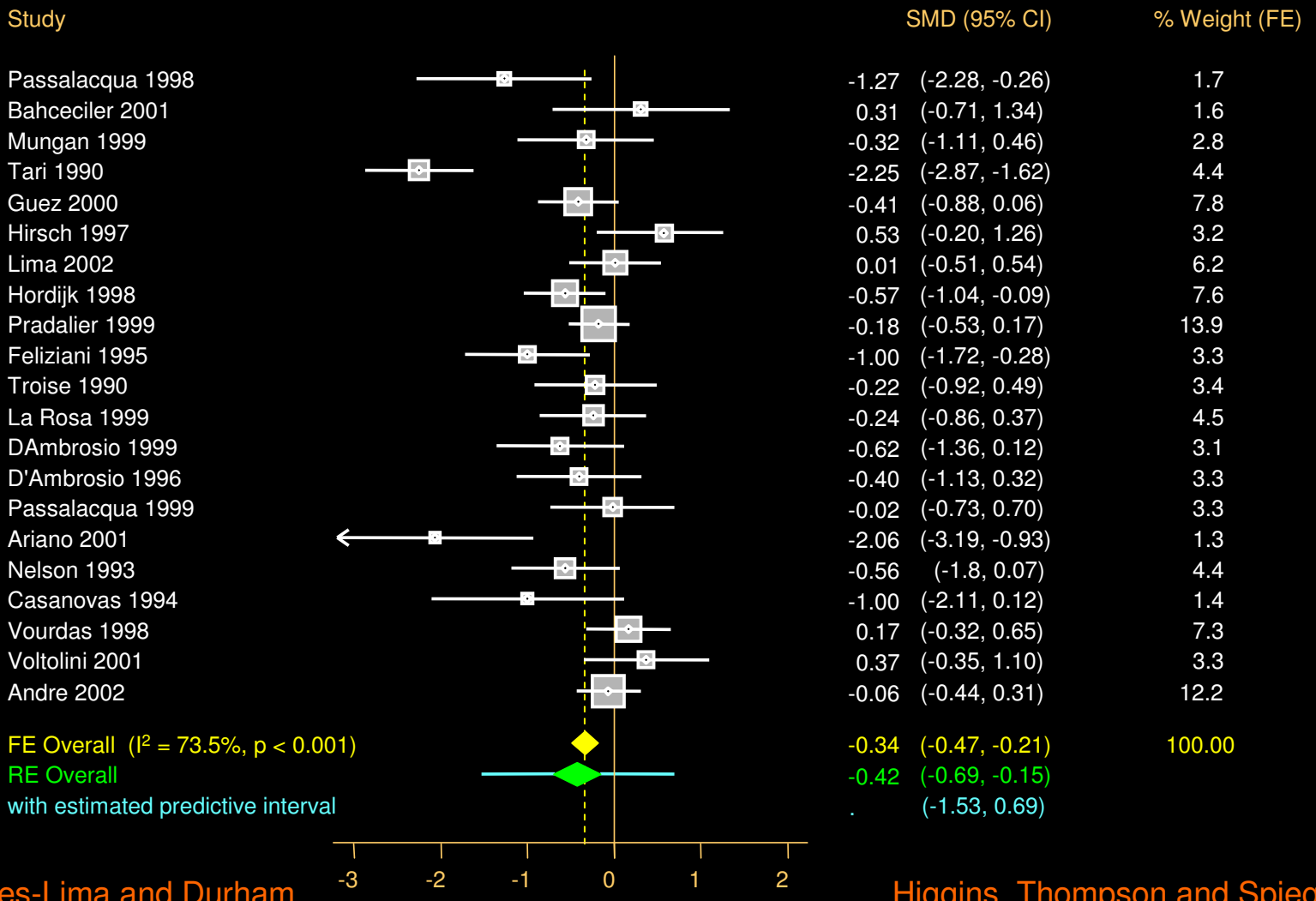
        as described in
        Hennekens, Buring and Hebert. *Stat Med* 1987; **6:** 397-402

Using meta-analysis to design trials

- Meta-analyses may reveal:
  - conflicting evidence
    - conflicts among trials, or between trials and expectations
  - new research questions
    - e.g. promise of an effect in a subgroup; suggestion that a particular mode of administration may be better
  - existing evidence is sufficient (or insufficient)

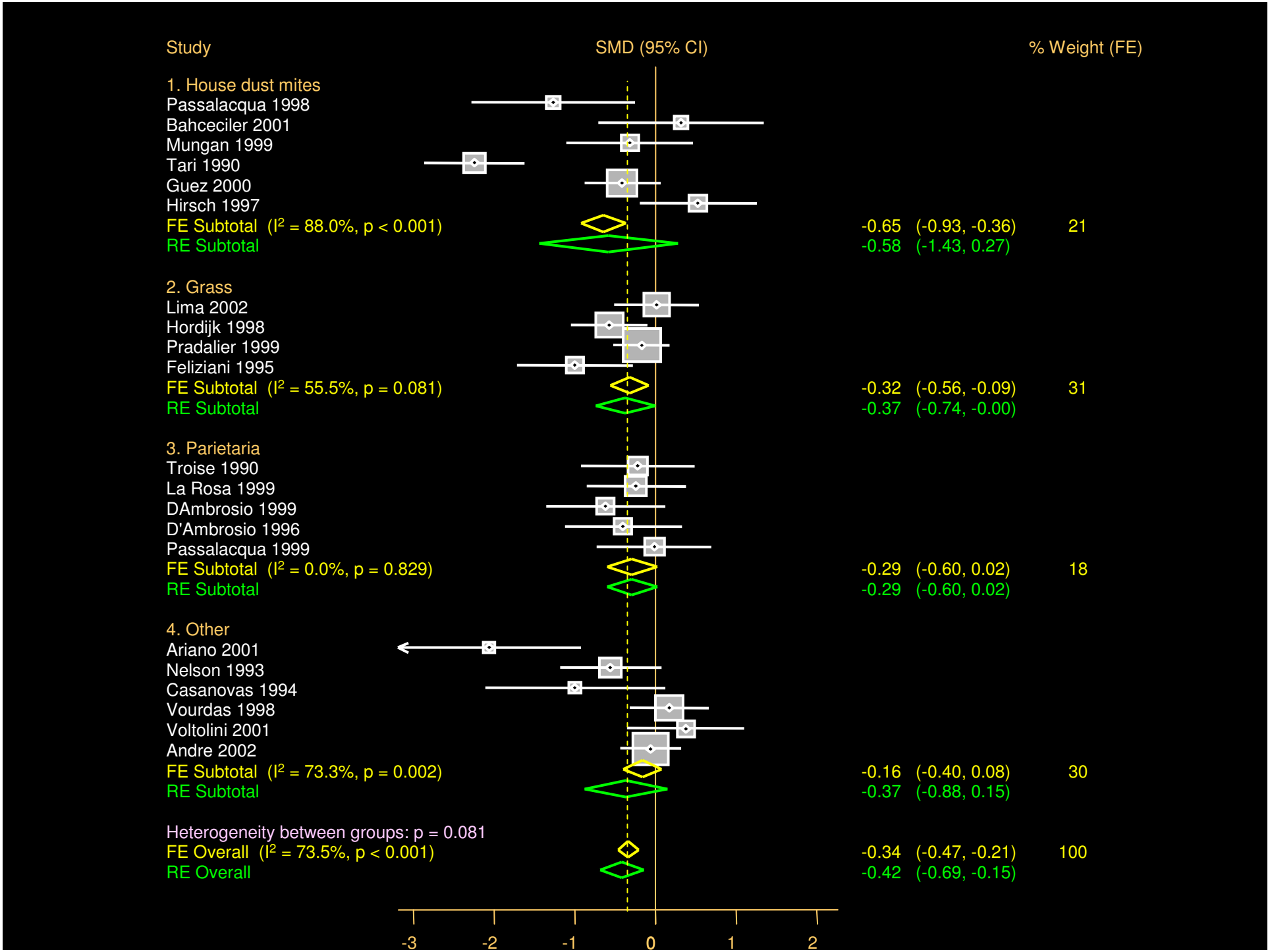
Is the meta-analysis conclusive?

Sublingual immunotherapy for allergic rhinitis: symptom scores



Wilson, Torres-Lima and Durham.
CDSR 2003, Art. No.: CD002893

Higgins, Thompson and Spiegelhalter.
JRSS A 2009; **172**: 137-159


Study
SMD (95% CI)
% Weight (FE)
1. House dust mites
Passalacqua 1998
Bahceciler 2001
Mungan 1999
Tari 1990
Guez 2000
Hirsch 1997
FE Subtotal (I² = 88.0%, p < 0.001)
-0.65 (-0.93, -0.36) 21
RE Subtotal
-0.58 (-1.43, 0.27)
2. Grass
Lima 2002
Hordijk 1998
Pradalier 1999
Feliziani 1995
FE Subtotal (I² = 55.5%, p = 0.081)
-0.32 (-0.56, -0.09) 31
RE Subtotal
-0.37 (-0.74, -0.00)
3. Parietaria
Troise 1990
La Rosa 1999
DAmbrosio 1999
D'Ambrosio 1996
Passalacqua 1999
FE Subtotal (I² = 0.0%, p = 0.829)
-0.29 (-0.60, 0.02) 18
RE Subtotal
-0.29 (-0.60, 0.02)
4. Other
Ariano 2001
Nelson 1993
Casanovas 1994
Vourdas 1998
Voltolini 2001
Andre 2002
FE Subtotal (I² = 73.3%, p = 0.002)
-0.16 (-0.40, 0.08) 30
RE Subtotal
-0.37 (-0.88, 0.15)
Heterogeneity between groups: p = 0.081
FE Overall (I² = 73.5%, p < 0.001)
-0.34 (-0.47, -0.21) 100
RE Overall
-0.42 (-0.69, -0.15)
-3
-2
-1
0
1
2

Beyond the direct randomized evidence

- In the absence of direct evidence, might look at
  - indirect comparison / multiple treatments meta-analysis

No of trials	Streptokinase	Alteplase-	Acclerated alteplase	Streptokinase +alteplase	Reteplase	Tenecteplase	PCTA
Boland et al[15]:							
8	P	P					
1	P		P	P			
1	P			P			
1	P				P		
2			P		P		
1			P			P	
Keeley et al[16]:							
8	P						P
3		P					P
11			P				P

PCTA = primary percutaneous transluminal coronary angioplasty.

Caldwell, Ades and Higgins. *BMJ* 2005; **331**: 897-900

Table 2 Pair-wise odds ratios between seven treatments for acute myocardial infarction obtained by direct and multiple treatment comparisons with fixed effect and random effects analyses*

	Fixed effect		Random effects	
Treatment comparison	**Direct comparisons**	**Multiple comparison**	**Direct comparisons**	**Multiple comparison**
Streptokinase *v*:				
Alteplase	1.00 (0.94 to 1.06)	0.99 (0.94 to 1.06)	0.89 (0.54 to 1.14)	0.96 (0.74 to 1.10)
Accelerated alteplase	0.86 (0.78 to 0.94)	0.86 (0.78 to 0.93)		0.84 (0.68 to 0.99)
Streptokinase+alteplase	0.96 (0.87 to 1.05)	0.96 (0.87 to1.05)		0.97 (0.75 to 1.25)
Reteplase	0.95 (0.79 to 1.12)	0.90 (0.80 to 1.01)		0.88 (0.65 to 1.06)
Tenecteplase		0.86 (0.74 to 1.00)		0.85 (0.57 to 1.17)
PCTA	0.52 (0.36 to 0.73)	0.63 (0.52 to 0.77)	0.49 (0.20 to 0.91)	0.62 (0.47 to 0.77)
Alteplase *v*:				
Accelerated alteplase		0.86 (0.77 to 0.95)		0.88 (0.70 to 1.19)
Streptokinase+alteplase		0.96 (0.86 to 1.07)		1.02 (0.78 to 1.51)
Reteplase		0.90 (0.79 to 1.02)		0.92 (0.70 to 1.24)
Tenecteplase		0.86 (0.73 to 1.01)		0.90 (0.61 to 1.35)
PCTA	0.63 (0.25 to 1.29)	0.64 (0.51 to 0.77)		0.65 (0.49 to 0.86)
Accelerated alteplase *v*:				
Streptokinase+alteplase	1.12 (1.00 to 1.25)	1.12 (1.01 to 1.24)		1.16 (0.91 to 1.55)
Reteplase	1.02 (0.90 to 1.16)	1.05 (0.94 to 1.17)		1.04 (0.81 to 1.28)
Tenecteplase	1.01 (0.88 to 1.14)	1.01 (0.89 to 1.14)		1.01 (0.74 to 1.35)
PCTA	0.81 (0.64 to 1.02)	0.74 (0.61 to 0.89)	0.79 (0.55 to 1.05)	0.73 (0.59 to 0.90)
Streptokinase+alteplase *v*:				
Reteplase		0.94 (0.82 to 1.07)		0.92 (0.62 to 1.19)
Tenecteplase		0.90 (0.76 to 1.05)		0.89 (0.57 to 1.27)
PCTA		0.66 (0.53 to 0.81)		0.64 (0.45 to 0.85)
Reteplase *v*:				
Tenecteplase		0.96 (0.82 to 1.13)		0.98 (0.68 to 1.43)
PCTA		0.71 (0.57 to 0.87)		0.71 (0.53 to 0.94)
Tenecteplase *v* PCTA		0.74 (0.58 to 0.92)		0.74 (0.50 to 1.03)

PCTA= primary percutaneous transluminal coronary angioplasty.

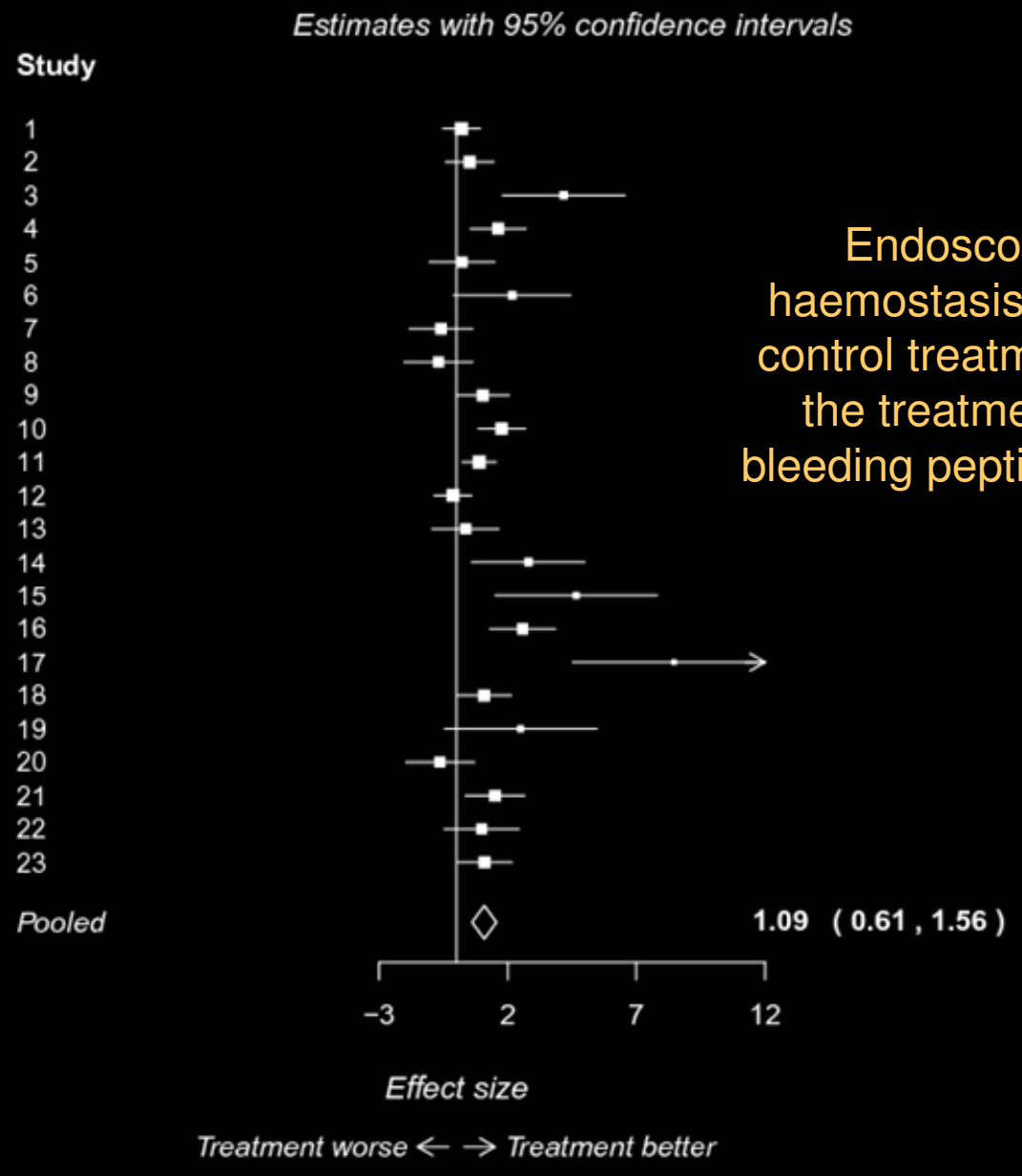
* Empty cells represent pair-wise comparisons that have not been evaluated in trials (fixed effect) or for which there are fewer than three trials (random effects).

Caldwell, Ades and Higgins. *BMJ* 2005; **331**: 897-900

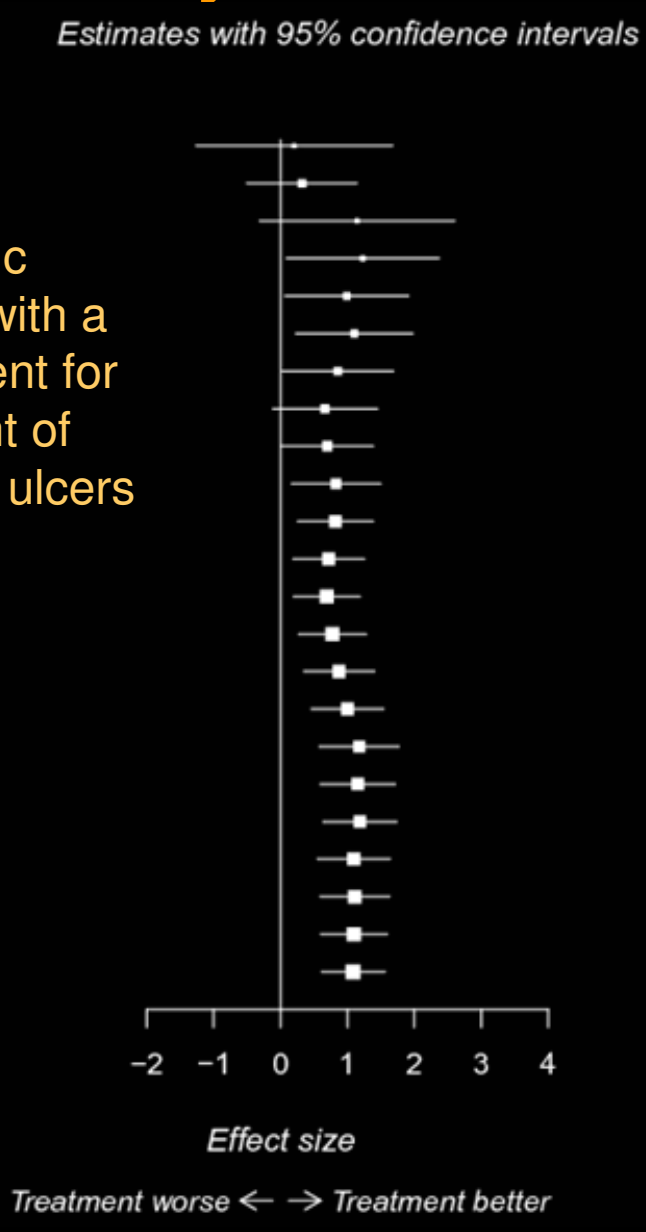
Beyond the direct randomized evidence

- In the absence of direct evidence, might look at
  - indirect comparison / multiple treatments meta-analysis
    - carefully chosen trial designs can add a lot of insight
      - Salanti, Higgins, Ades and Ioannidis. *Stat Meth Med Res* 2008; 17: 279-301
  - non-randomized (or weak randomized) evidence
    - possibly with bias-adjusted analyses
      - Turner, Spiegelhalter, Smith and Thompson. *JRSS A* 2009; **172**: 21-47
      - Welton, Ades, Carlin, Altman and Sterne. *JRSS A* 2009; **172**: 119-136

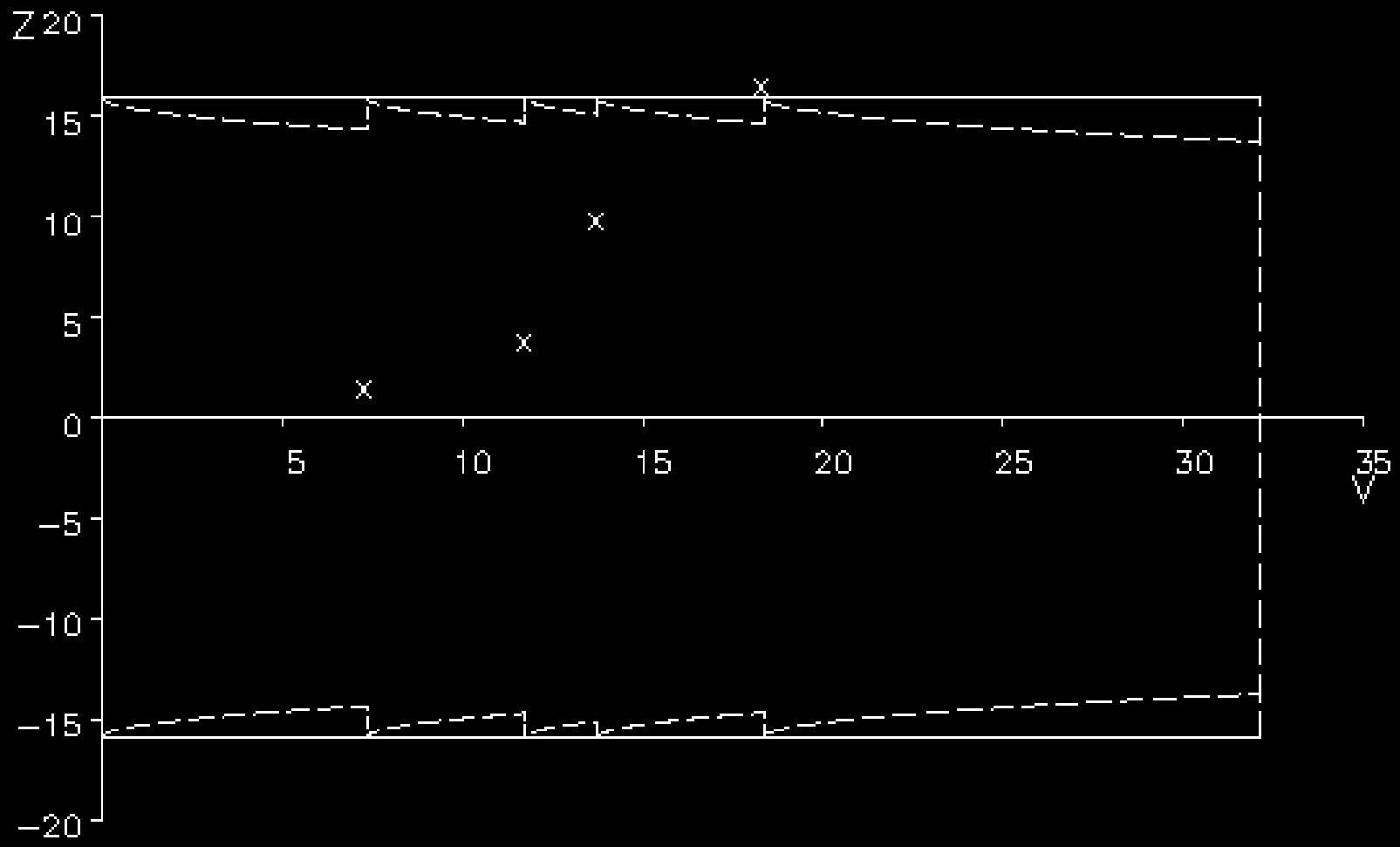
Cumulative meta-analysis



Endoscopic haemostasis with a control treatment for the treatment of bleeding peptic ulcers



Sequential meta-analysis



STOP the study - a boundary has been crossed

Sequential meta-analysis

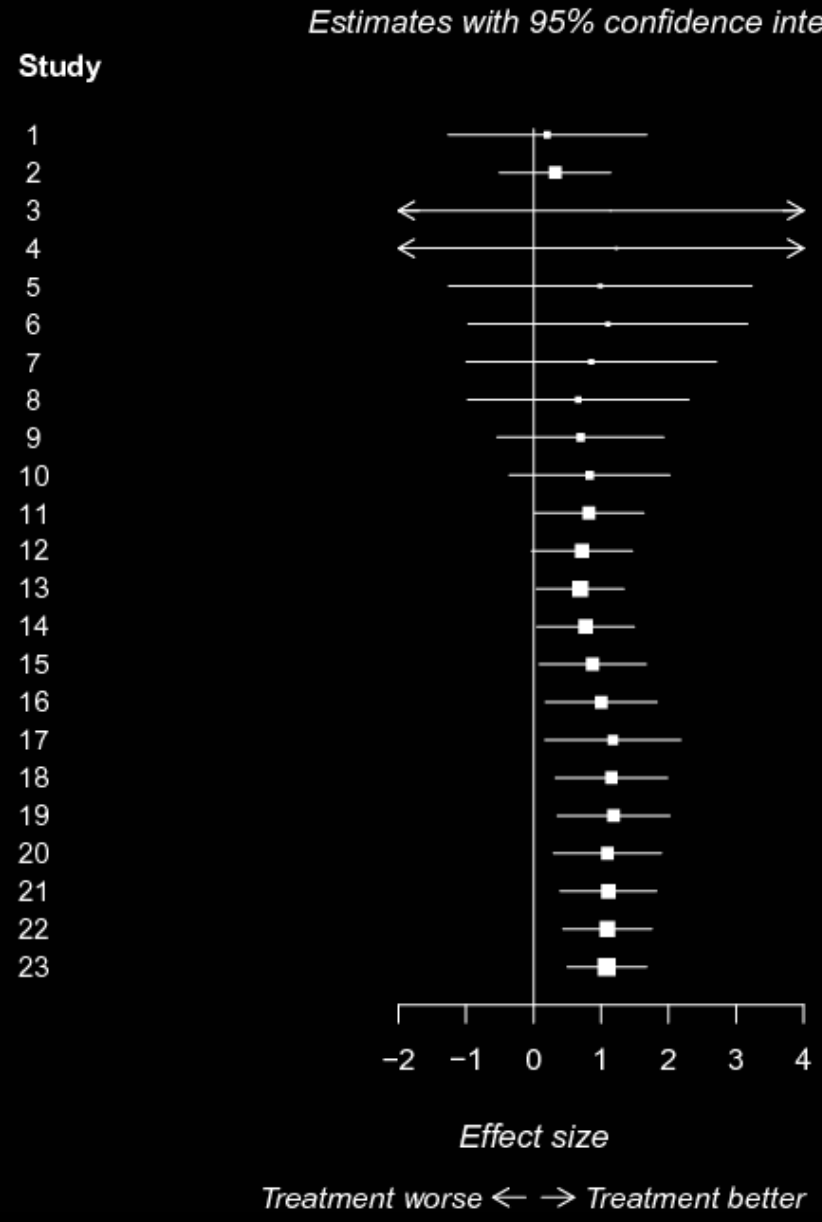


Illustration of repeated confidence intervals from sequential random-effects meta-analysis (O'Brien-Fleming boundary)

Higgins, Whitehead and Simmonds. *Stat Med* (to appear)

Powering a new trial around a meta-analysis

- “In the event that investigators and their advisors or peer reviewers decide that a new study is indicated in spite of the trends indicated by meta-analysis of past studies, there is no reason why the new large study should be sized as if there were no prior information.”
- “... one could estimate various outcomes of a new study necessary to make the confidence intervals avoid 0 or 1, depending on the model.”

Chalmers and Lau, *Stat Med* 1996; **15**: 1263-68

Conference on Meta-analysis in the Design and Monitoring of Clinical Trials, June 1994

Bayesian predictive power

- Usual random-effects meta-analysis model:

$$y_i \mid \theta_i \sim N(\theta_i, v_i)$$

$$\theta_i \sim N(\mu, \tau^2)$$

- Posterior distribution is:

$$\theta_{new} \mid \mu, \tau^2, y_{new}, v_{new} \sim N\left(\frac{v_{new}\mu + \tau^2 y_{new}}{v_{new} + \tau^2}, \frac{v_{new}\tau^2}{v_{new} + \tau^2}\right)$$

- Fully Bayesian predictive power can be derived from this

DerSimonian, *Stat Med* 1996; **15**: 1237-48

Sutton's simulation approach

"the updated meta-analysis will be of central importance and more influential than the results of the new studies on their own" Sutton, Cooper and Jones. *BMC MRM* 2009

1. Use predictive distribution to simulate effect in new trial
2. Generate data for the new trial
3. Repeat meta-analysis with new trial added
4. Test null hypothesis at pre-set significance level
5. Repeat 2-4 many times
6. Estimate power (proportion of simulations rejecting null)

- Iterate until desired power is reached

Sutton, Cooper et al. *Stat Med* 2007; **26**: 2479–2500

Power of a meta-analysis

- From trial *i* have estimate = y_i, variance assumed known

$$\hat{\mu} = \frac{\sum w_i^* y_i}{\sum w_i^*} \qquad \text{var}(\hat{\mu}) = \frac{1}{\sum w_i^*} \qquad w_i^* = \frac{1}{v_i + \hat{\tau}^2}$$

- We can work out the power of a meta-analysis to detect an effect μ

$$\text{Power} = 1 - \Phi\left(c_\alpha - \frac{\mu}{\text{SE}(\mu)}\right) + \Phi\left(-c_\alpha - \frac{\mu}{\text{SE}(\mu)}\right)$$

Hedges and Pigott.
Psychological Methods 2001; **6**: 203–217

Conditional power

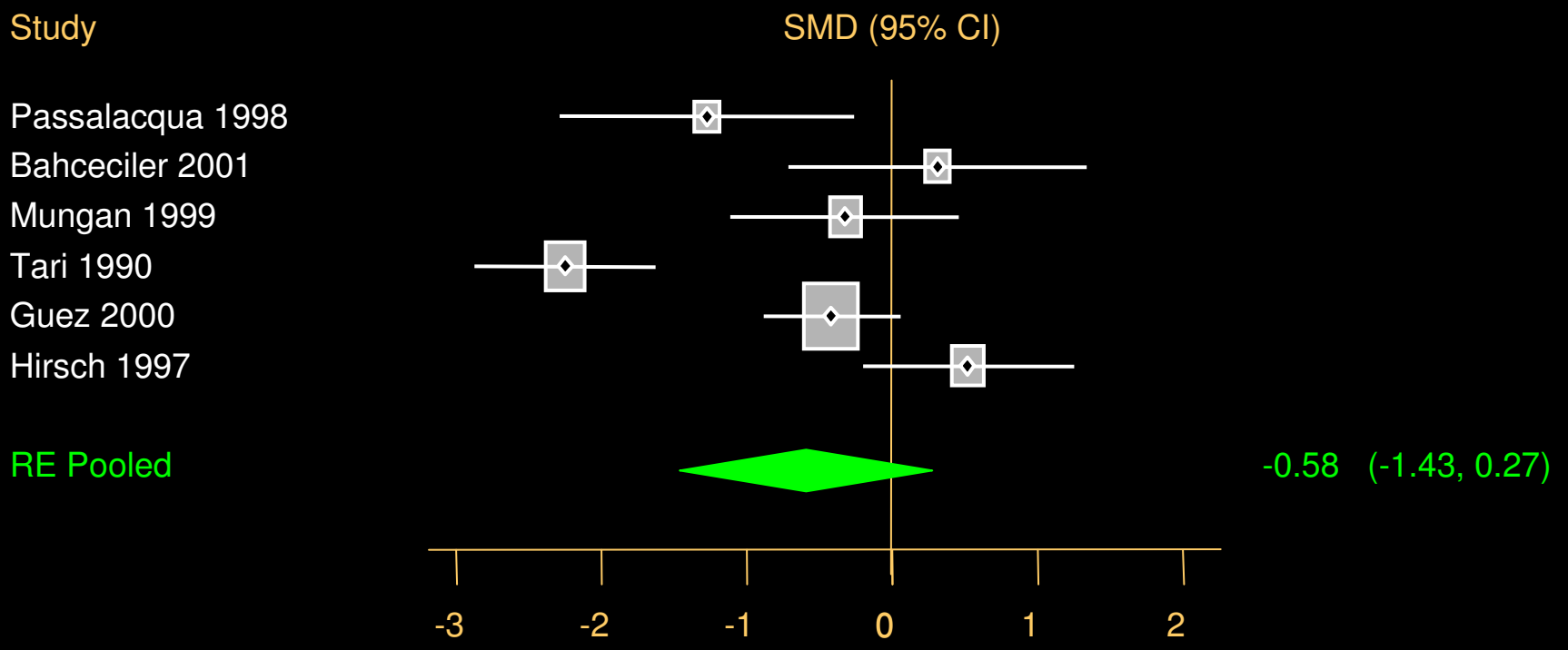
- For future studies, we derive conditional power
  - power to detect overall mean effect μ given the result of the existing meta-analysis
- Suppose there are to be *m* new studies, each with (FE) weight *W*/*m*

$$\text{Power} = \Phi\left(-\sqrt{\frac{m+W\tau^2}{mW}}\left(c_\alpha \sqrt{\sum_{i=1}^{k} w^*_{i,old} + \frac{mW}{m+W\tau^2}} - \sum_{i=1}^{k} w^*_{i,old} y_{i,old} \right) + \frac{m\mu}{\sqrt{\frac{m^2}{W} + m\tau^2}} \right)$$

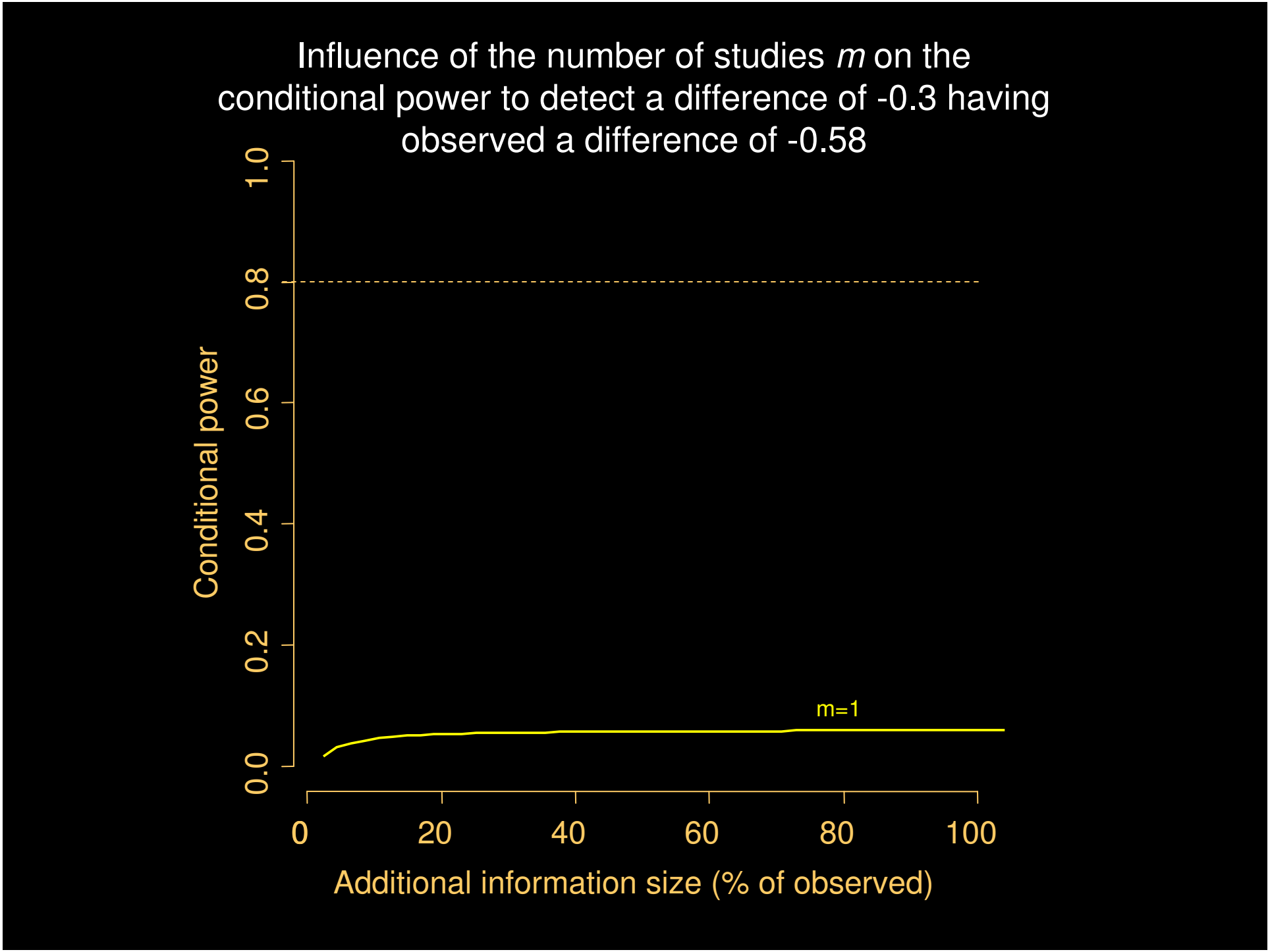
$$+\Phi\left(-\sqrt{\frac{m+W\tau^2}{mW}}\left(c_\alpha \sqrt{\sum_{i=1}^{k} w^*_{i,old} + \frac{mW}{m+W\tau^2}} + \sum_{i=1}^{k} w^*_{i,old} y_{i,old} \right) - \frac{m\mu}{\sqrt{\frac{m^2}{W} + m\tau^2}} \right)$$

- We can partition heterogeneity τ^2 into ‘old’ and ‘new’ bits

Trials of sublingual immunotherapy: house dust mites

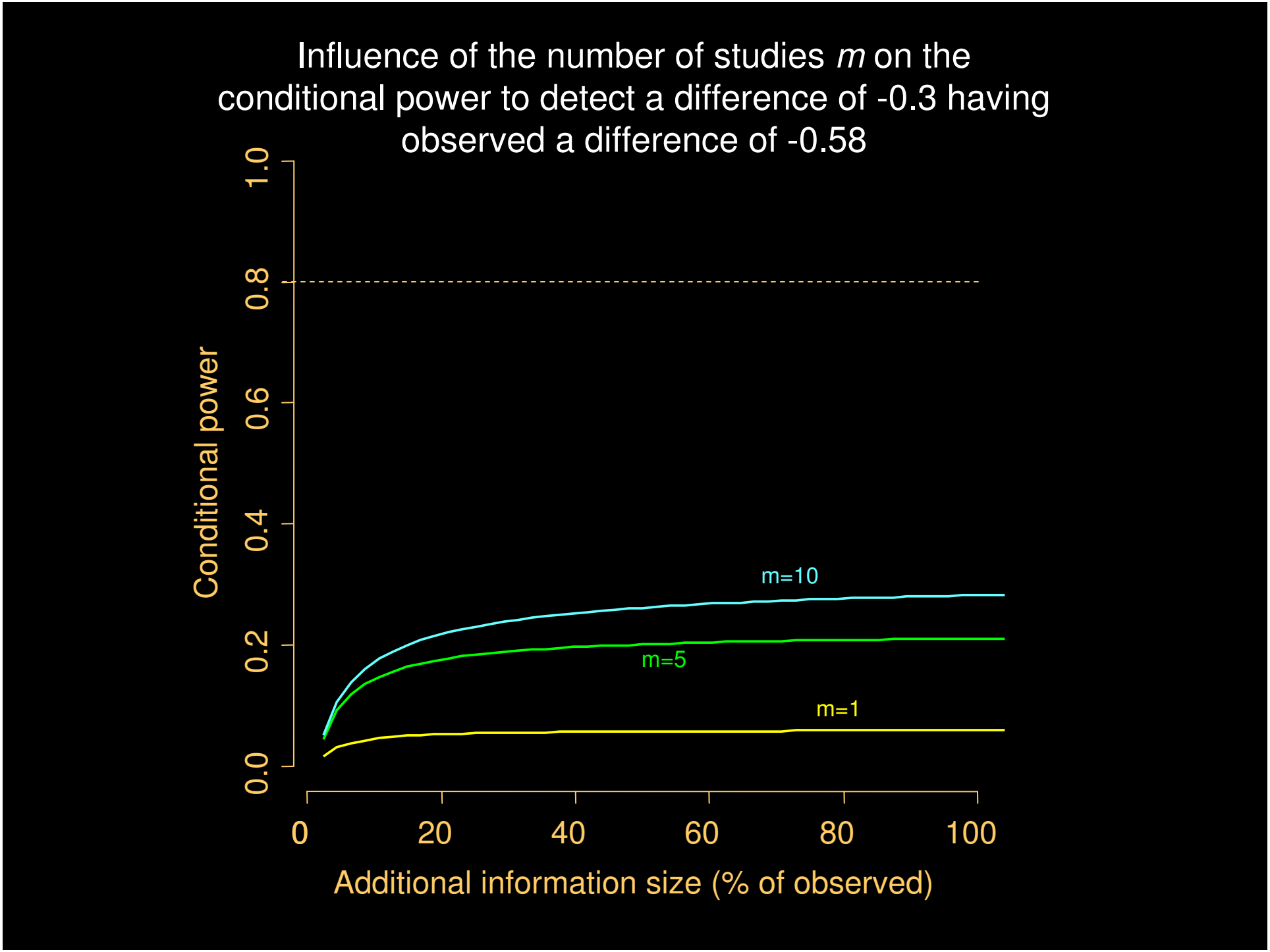


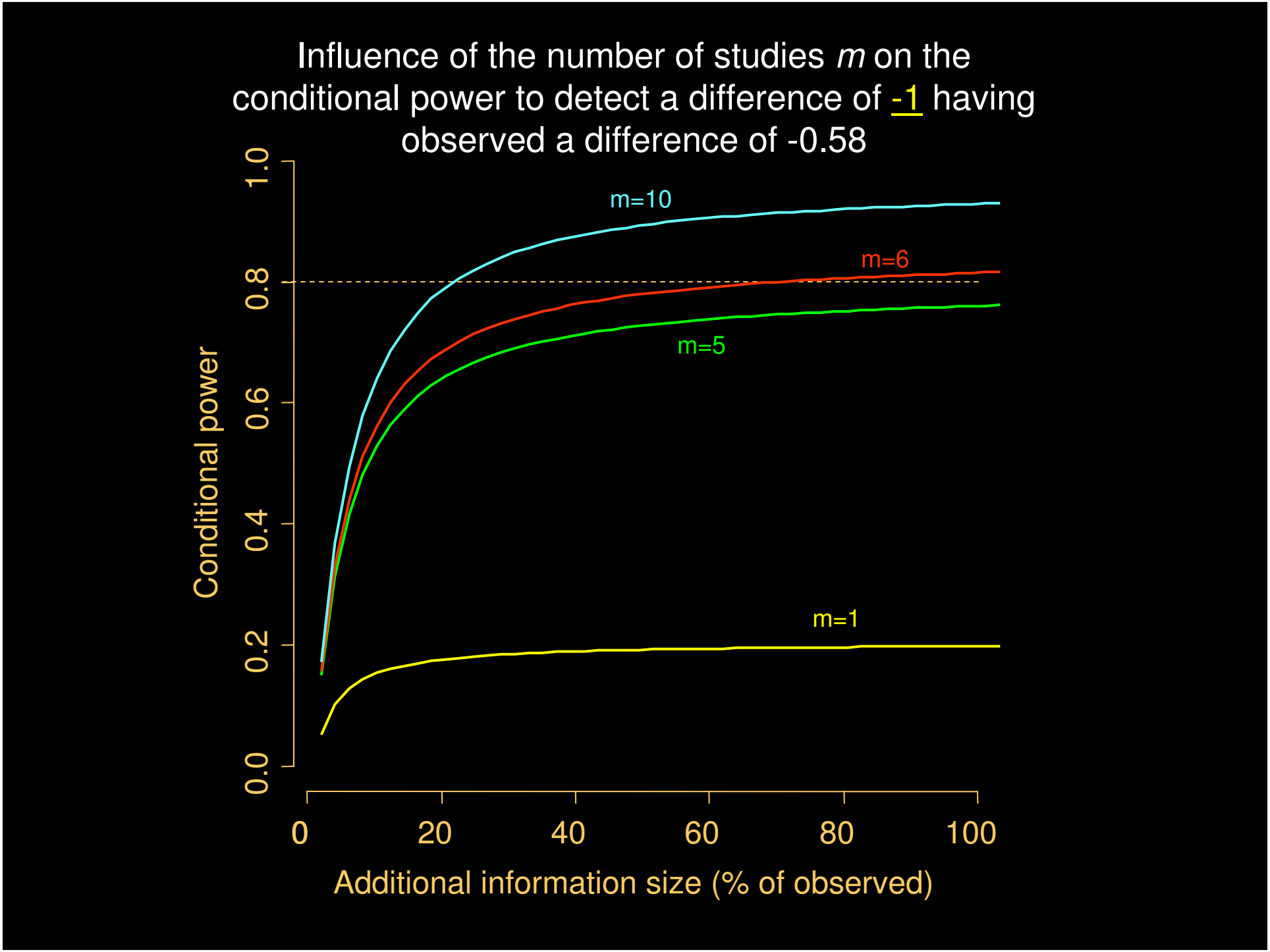
Power of existing data to detect SMD = –0.3 is 11%
to detect SMD = –1.0 is 63%

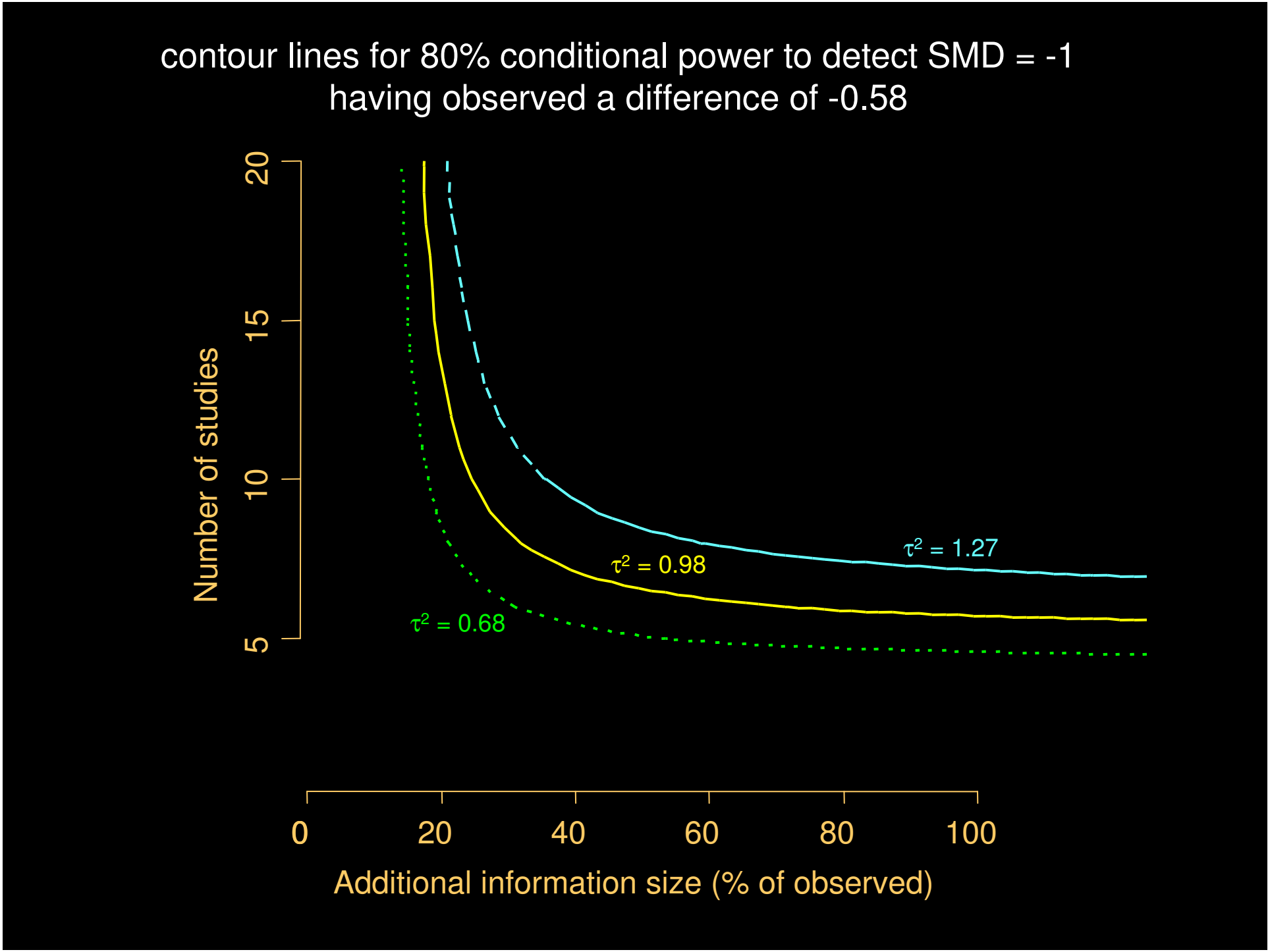

Influence of the number of studies *m* on the conditional power to detect a difference of -0.3 having observed a difference of -0.58
Conditional power
0.0
0.2
0.4
0.6
0.8
1.0
m=1
0
20
40
60
80
100
Additional information size (% of observed)

“one large beautiful trial is not necessarily going to convince the world, because of true between-study variability”

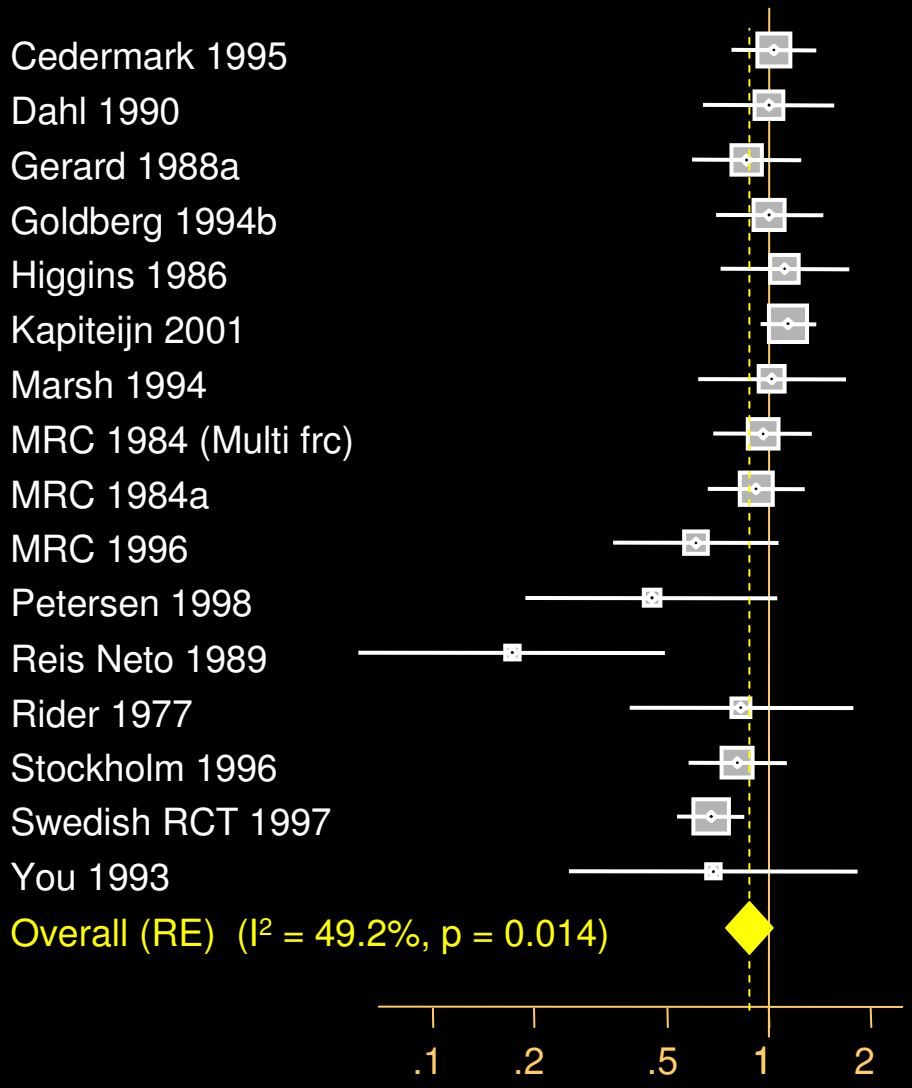
Thomas Louis, *Stat Med* 1996; **15**: 1250
Conference on Meta-analysis in the Design and Monitoring of Clinical Trials, June 1994


Influence of the number of studies *m* on the conditional power to detect a difference of -0.3 having observed a difference of -0.58
Conditional power
0.0
0.2
0.4
0.6
0.8
1.0
m=10
m=5
m=1
0
20
40
60
80
100
Additional information size (% of observed)


Influence of the number of studies *m* on the conditional power to detect a difference of -1 having observed a difference of -0.58
Conditional power
0.0
0.2
0.4
0.6
0.8
1.0
m=10
m=6
m=5
m=1
0
20
40
60
80
100
Additional information size (% of observed)


contour lines for 80% conditional power to detect SMD = -1
having observed a difference of -0.58
Number of studies
5
10
15
20
τ² = 1.27
τ² = 0.98
τ² = 0.68
0
20
40
60
80
100
Additional information size (% of observed)

Pre-operative radiotherapy and curative surgery for the management of localized rectal carcinoma



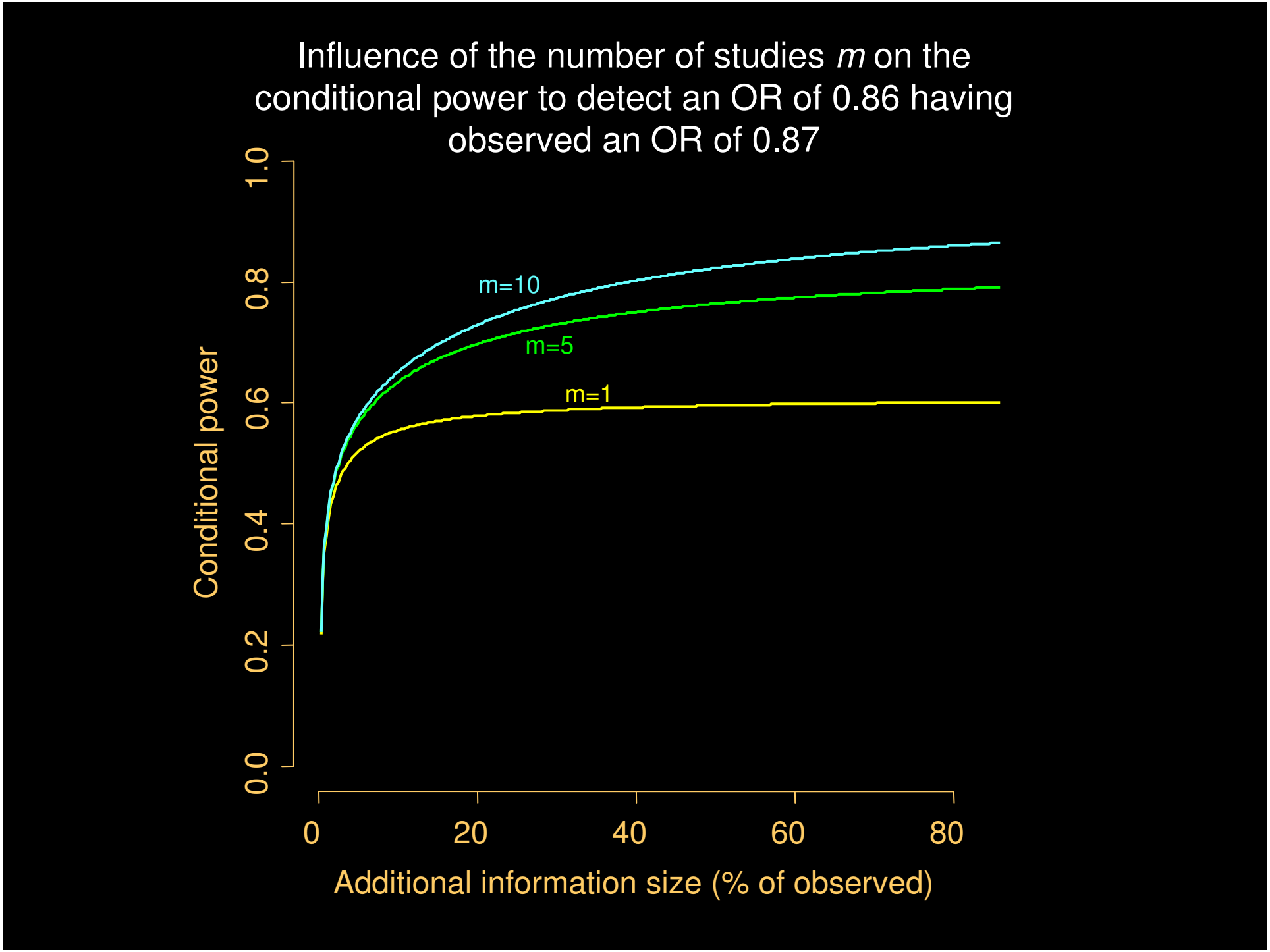
Power to detect OR of…

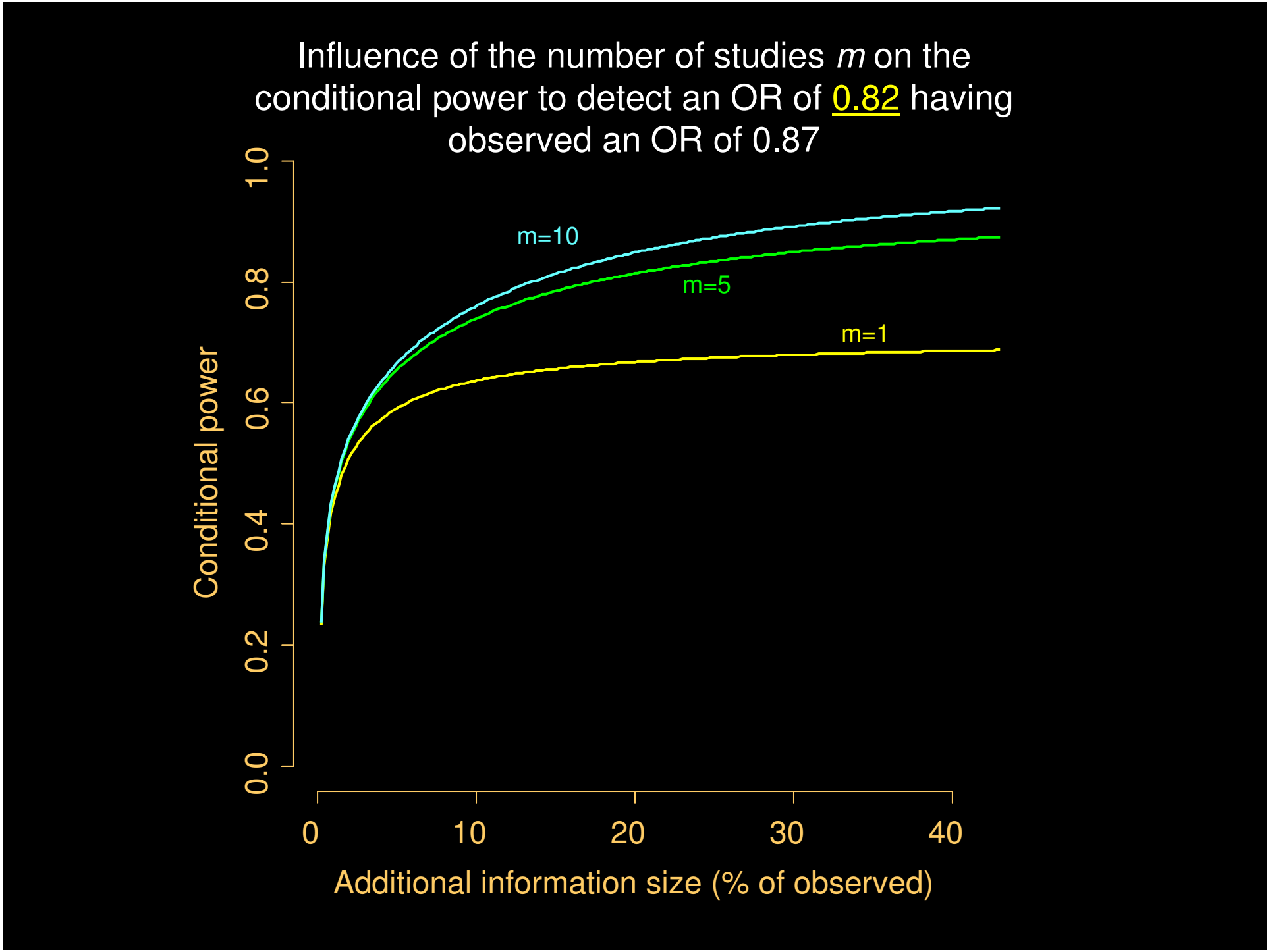
0.86 is 55%

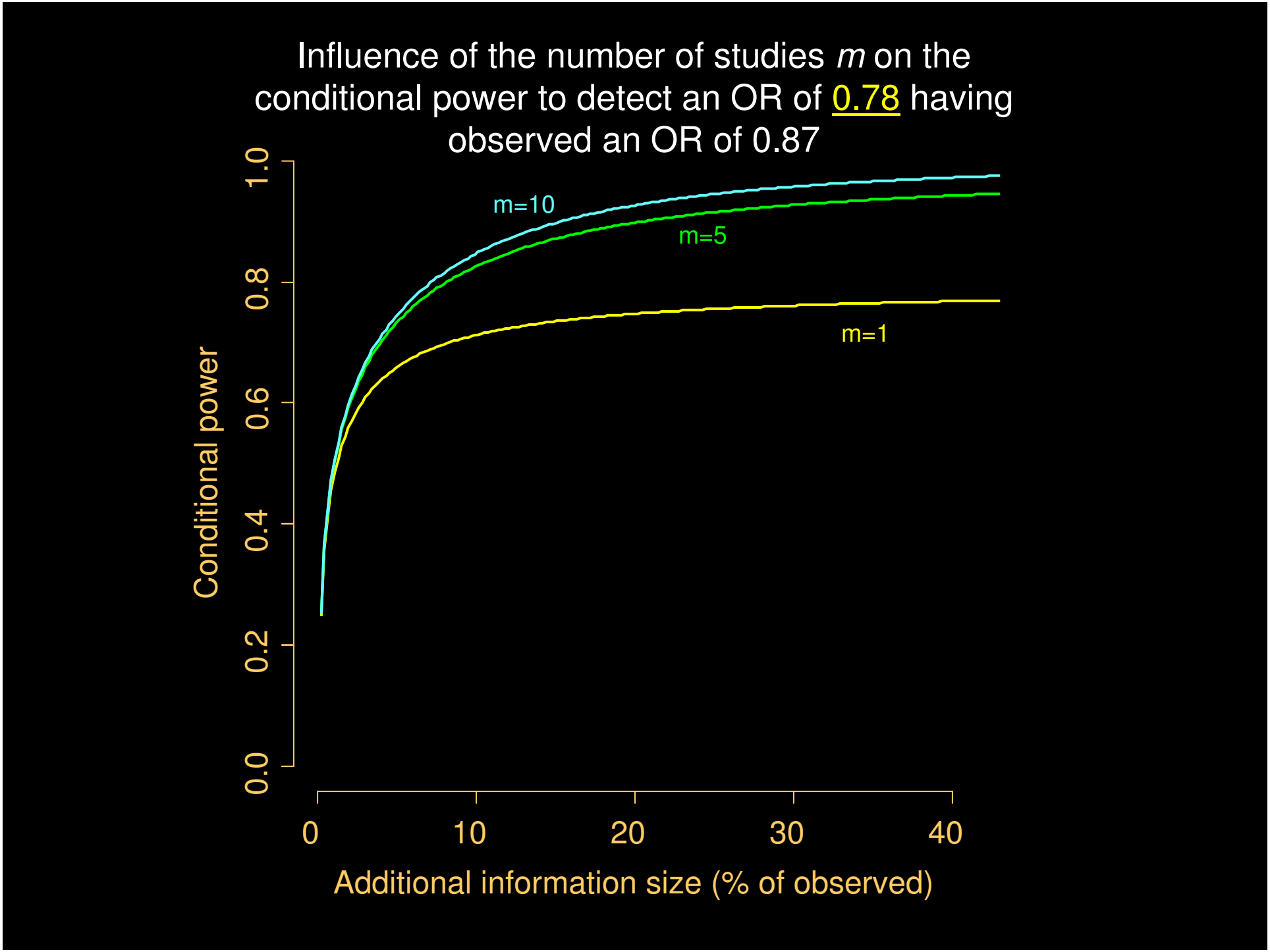
0.82 is 79%

0.78 is 93%

0.74 is 99%


Influence of the number of studies *m* on the conditional power to detect an OR of 0.86 having observed an OR of 0.87
Conditional power
0.0
0.2
0.4
0.6
0.8
1.0
m=10
m=5
m=1
0
20
40
60
80
Additional information size (% of observed)


Influence of the number of studies *m* on the conditional power to detect an OR of 0.82 having observed an OR of 0.87
Conditional power
0.0
0.2
0.4
0.6
0.8
1.0
m=10
m=5
m=1
0
10
20
30
40
Additional information size (% of observed)


Influence of the number of studies *m* on the conditional power to detect an OR of 0.78 having observed an OR of 0.87
m=10
m=5
m=1
Conditional power
0.0
0.2
0.4
0.6
0.8
1.0
0
10
20
30
40
Additional information size (% of observed)

Concluding remarks

- Meta-analyses (or systematic reviews) should be done before a trial is planned
- Sophisticated methods of analysis may add insight
  - understanding heterogeneity
  - adjusting for bias and relevance
  - indirect evidence
  - expected value of information
- Meta-analyses inevitably ask a broader question than an individual trial
- However, policy may be determined primarily using meta-analyses rather than individual trials
- Trials can be powered to tackle the meta-analysis question rather than the individual trial question
- Multi-centre trials, or multiple trials, may be indicated