Workshop: Accumulating evidence, sequential synthesis and the challenge of multiplicity

Presenters:

Prof. Alex Sutton, University of Leicester &

Prof. Neil Hawkins, University of Glasgow

With content created by Prof. Nicola Cooper, University of Leicester

All are members of the NIHR Complex Review Support Unit (CRSU)

(http://www.nihrcrsu.org/)

Structure of the workshop:

- Session 1: Synthesising accumulating evidence
- Session 2: The coherent design and analysis of future trials
- Session 3: Sequential analysis and multiplicity
- Session 4: Case-study

- 3 x 25 minute sessions (3 x 12 min lecture + 10 minute exercise + 3 minute discussion) + 20 minute case study to finish.

Aims and objectives of the workshop

- Participants to obtain a deeper appreciation of the pertinent issues surrounding the accumulation of evidence, and, its role in informing future research agendas, by drawing together disparate literatures.

- Facilitate the engagement of participants to discuss the material and explore how it is relevant to their roles in Cochrane and beyond.

- Help participants to develop opinions regarding the methods discussed, appreciating *why* they were developed and their *strengths* and *limitations.*
  - *Acknowledging there are unresolved issues*

- Provide motivation for participants to challenge what they (and others) routinely do with a view to improve current practice(??)

Session 1:
Synthesising accumulating evidence

Why do we **synthesise** (accumulating) evidence?

Cochrane Collaboration aim “We gather and summarize the best evidence from research to help you make informed choices about treatment”

- Underpin the principles of Evidence Based Medicine

- **Meta-analysis** is a statistically rigorous approach to estimate treatment effects
  - Increase power
  - Quantify / explore heterogeneity

The “bad old days”

- Antman EM, Lau J, Kupelnick B, Mosteller F, Chalmers TC. A comparison of results of meta-analyses of randomized control trials and recommendations of clinical experts: Treatments for myocardial infarction. JAMA. 1992;268:240-8
- In some cases effective treatments were not recommended for more than a decade after a meta-analysis of RCTs would have shown them to be effective
- Creation of cumulative meta-analysis (RHS):
  - Studies are added one at a time according to date of publication (or other variable) and the results are summarised as each new study is added
  - Each horizontal line represents the summary of the results as each study is added, rather than the results of a single study.

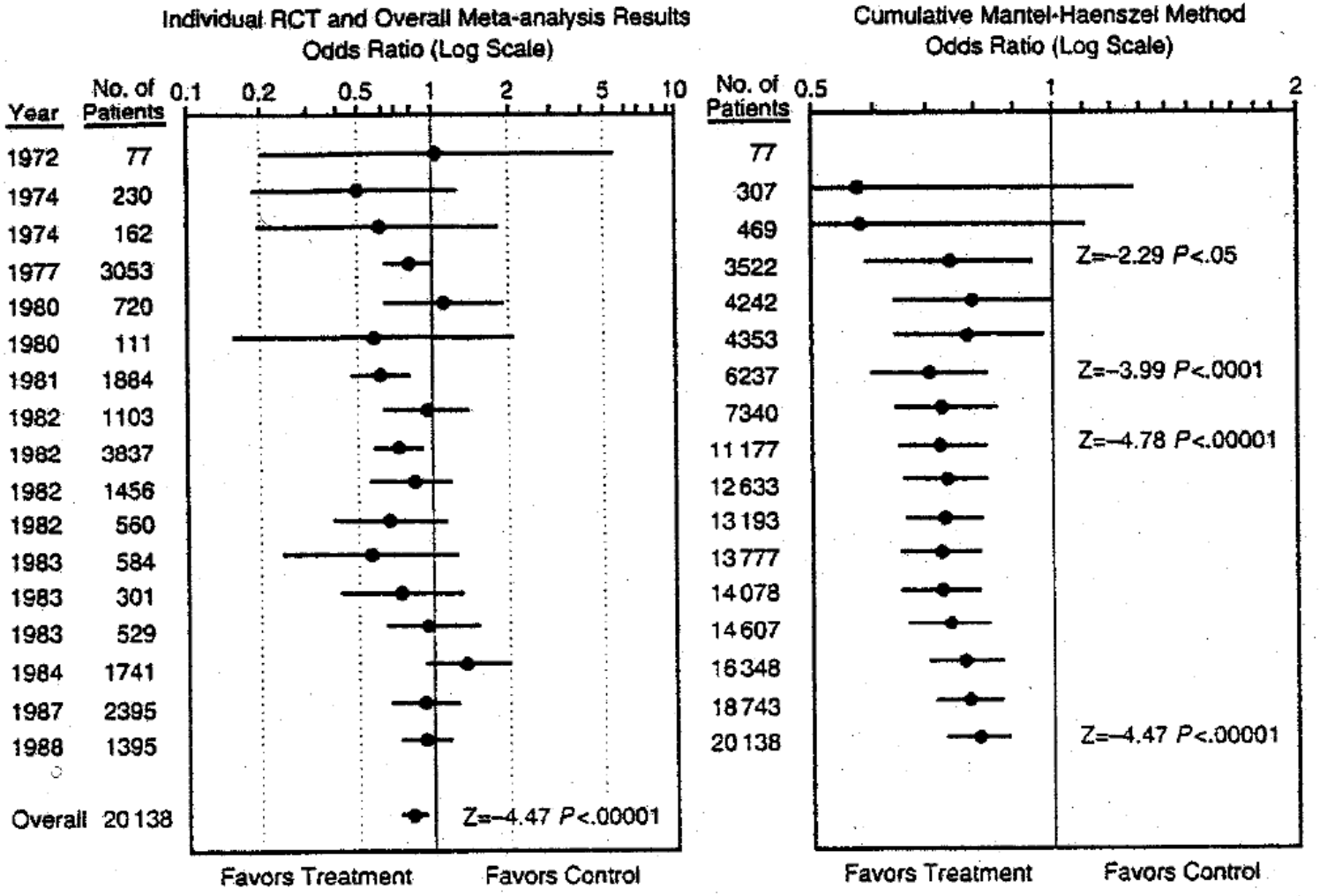


Fig 1.—Results of 17 randomized control trials (RCTs) of the effects of oral β-blockers for secondary prevention of mortality in patients surviving a myocardial infarction presented as two types of meta-analyses. On the left is the traditional one, revealing many trials with nonsignificant results but a highly significant estimate of the pooled results on the bottom of the panel. On the right, the same data are presented as cumulative meta-analyses, illustrating that the updated pooled estimate became statistically significant in 1977 and has remained so up to the present. Note that the scale is changed on the right graph to improve clarity of the confidence intervals.

Cumulative meta-analysis (cont.)

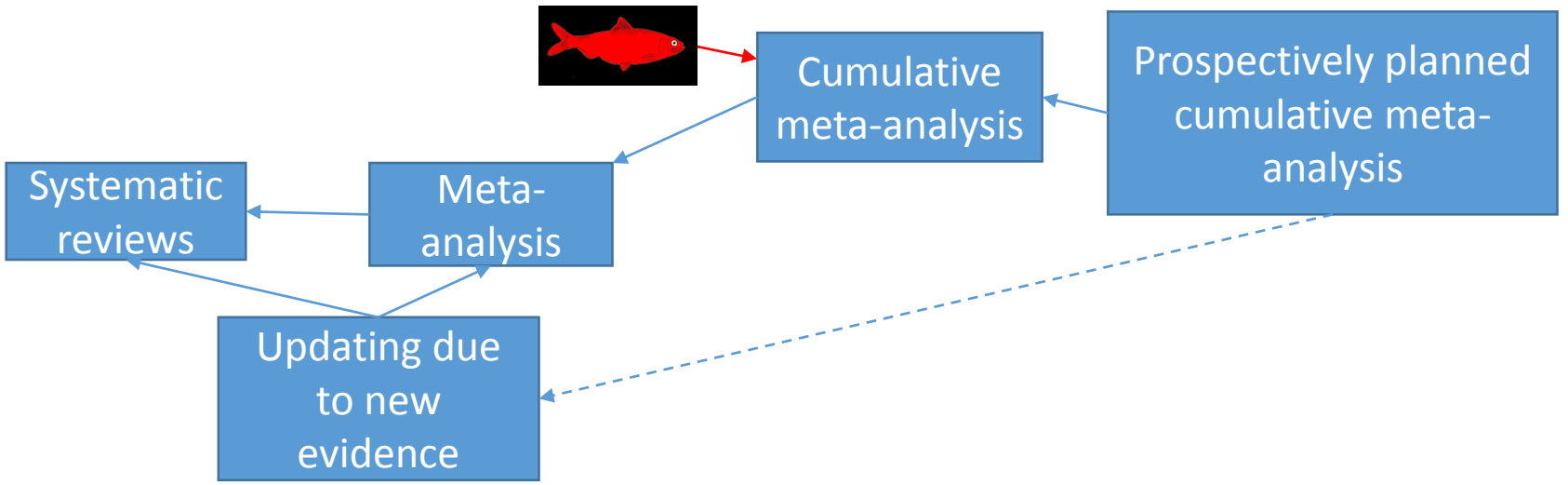
- So why is **cumulative meta-analysis** not in the Cochrane Handbook for systematic reviews of interventions?
  - Other than to illustrate historical issues (such as on previous slide), a **retrospectively focused** research tool (See also Clarke et al Plos One 2014).
- What we most care about is the situation **now** and that in the **future,** if/when new evidence is generated
- **Prospective cumulative meta-analysis**
  - Used to **design multiple future studies,** to ensure they are compatible to be synthesised, e.g. collect correct data, compatible timescales etc.
    - Interesting idea but rarely used(?)
    - More on designing future trials in Session 2

Updating systematic reviews

- Cochrane support updating of existing reviews:
  - Current guidance: “A Cochrane Review should be updated based on need.”
- (In my opinion, the electronic, **updatable**, Cochrane database of systematic reviews was an innovation Cochrane got correct from the start - No such commitment even by the best journals(?))
- So Cochrane **do** perform **cumulative meta-analysis** when updating reviews?
  - Yes, but it is not necessarily done after every new study, and assessing how results change over time is not usually emphasised.

What have we learnt so far?

- Systematic reviews, including **meta-analysis,** can provide clear and up-to-date summaries of the evidence for a given intervention
- **Updating reviews** as new evidence is generated is usually desirable
- **Cumulative meta-analysis** sounds like it should be a framework for updating, but it isn't



So, what's the big deal??

- Message seems straight-forward enough!

  Do good systematic reviews/meta-analyses and keep them up-to-date

- OK, but let’s dig a little deeper. . . . . .

Difficult Question #1:
When and how often should a Cochrane review be updated?

Difficult Question #1:
When and how often should a Cochrane review be updated?

- Often not a quick task
- Limited resources, not all reviews can be kept up-to-date(?)
- Can some reviews be considered “closed” – question already answered (despite new evidence)?
- Can “too many” (however defined) updates potentially contribute to some reviews becoming misleading?
  - By multiple looks & testing of the (overlapping) evidence without acknowledging this in the analysis
  - (But we already have seen why zero updates are not sensible. . . . .)
- Can we devise a more explicit update policy / strategy than current guidance

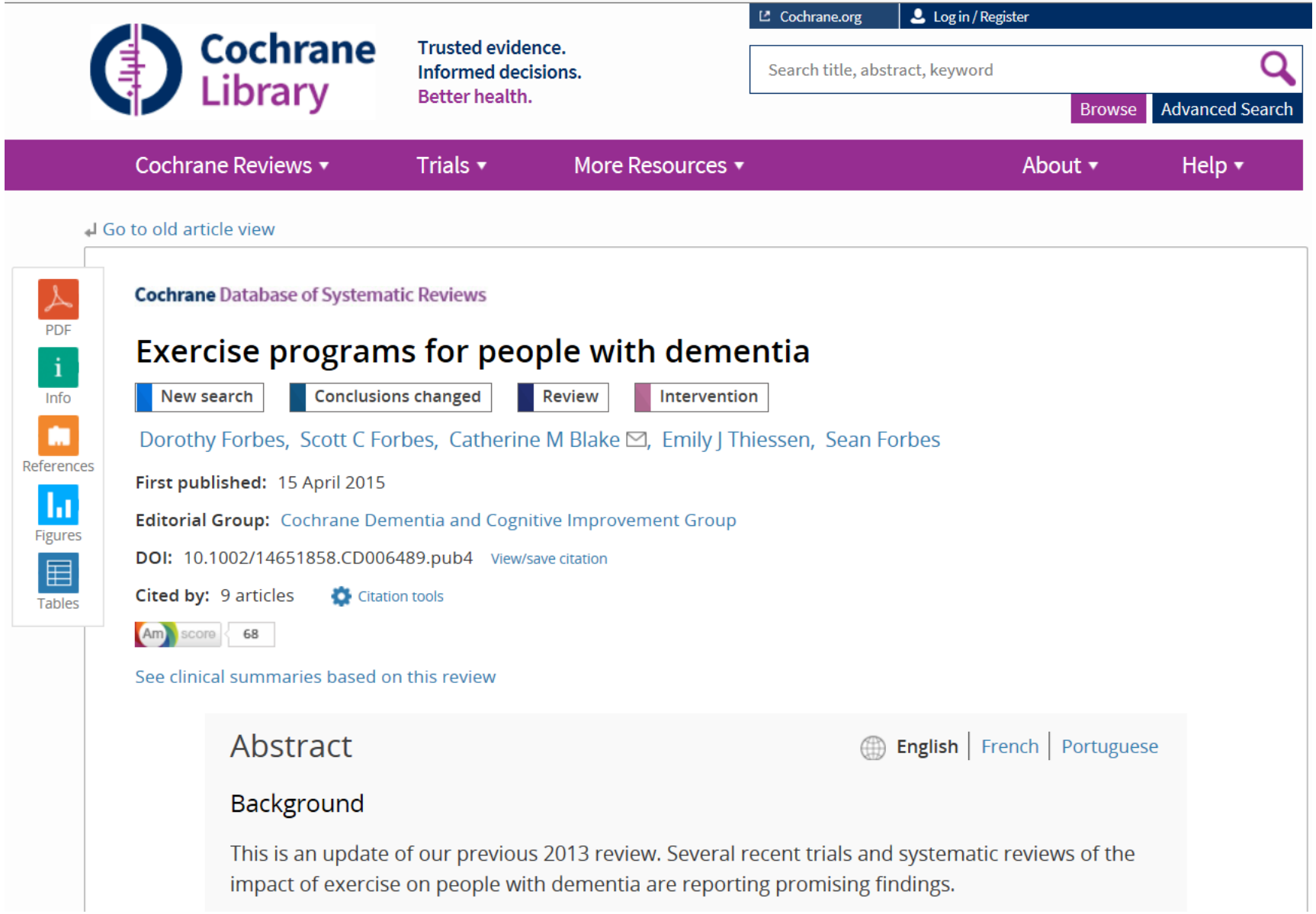
“A Cochrane Review should be updated based on need.”

Difficult Question #2:
How can a systematic reviews / meta-analyses influence and inform the design of future trials (they will potentially include in the future)?

Difficult Question #2:
How can a systematic reviews / meta-analyses influence and inform the design of future trials (they will potentially include in the future)?

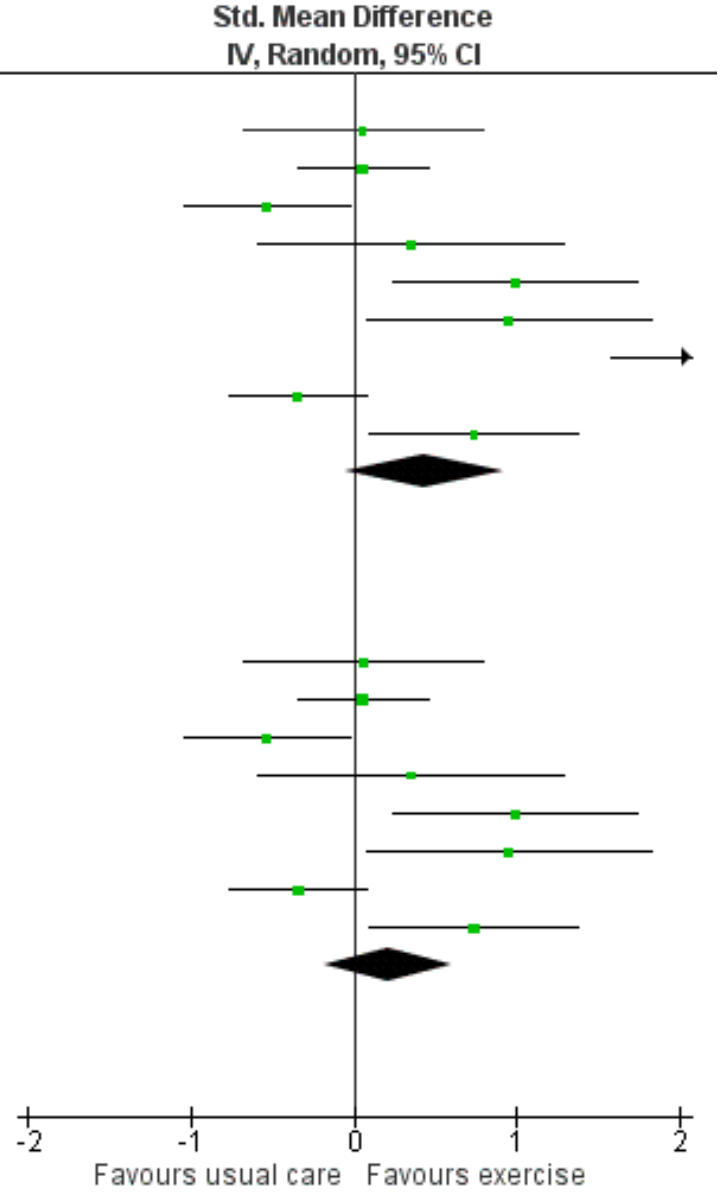
- I would like to think a **systematic review** is **not** "just**" a passive summary document**, but it also helps shape the future
- We all support **evidence based medicine**, so why should we not strive for better **evidence based research**? (See http://ebrnetwork.org/)
  - Use of systematic reviews to *justify* new trials are now commonplace(?)
  - But systematic reviews are not used as much as they could be for *designing* future studies. (Jones et al. 2013)
- Can we provide (more) information in reviews to facilitate their use for informing future research?
  - Hands up all those who have written "more research required" in the recommendations section of a review? (*Clarke et al.. 2007)*
    - Research on which interventions, patient groups, measuring which outcomes, and how big should a new trial be?

Exercise 1: Accumulating evidence



Exercise programs for people with dementia

Study or Subgroup	Exercise Mean	Exercise SD	Exercise Total	Usual care Mean	Usual care SD	Usual care Total	Weight	Std. Mean Difference IV, Random, 95% CI
1.1.1 Cognition: all trials								
Christofoletti 2008	14.9	2.2	12	14.8	1.3	17	11.0%	0.06 [-0.68, 0.80]
Eggermont 2009a	0.24	0.78	51	0.2	0.63	46	13.5%	0.06 [-0.34, 0.45]
Eggermont 2009b	0.07	0.37	30	0.47	0.97	31	12.7%	-0.53 [-1.05, -0.02]
Hwang 2010	28.9	11.86	10	24	14.68	8	9.5%	0.35 [-0.58, 1.29]
Kemoun 2010	30.38	7.66	16	22.23	8.37	15	10.9%	0.99 [0.24, 1.74]
Van de Winckel 2004	15.33	4.44	15	11	4.3	9	10.0%	0.95 [0.08, 1.83]
Venturelli 2011	12	2	11	6	2	10	7.2%	2.88 [1.59, 4.17]
Volkers 2012	0.14	0.5	50	0.38	0.89	38	13.3%	-0.34 [-0.77, 0.08]
Vreugdenhil 2012	23.9	5	20	19	7.7	20	11.8%	0.74 [0.10, 1.38]
Subtotal (95% CI)			**215**			**194**	**100.0%**	**0.43 [-0.05, 0.92]**
Heterogeneity: Tau² = 0.42; Chi² = 40.90, df = 8 (P < 0.00001); I² = 80%								
Test for overall effect: Z = 1.75 (P = 0.08)								
1.1.2 Cognition: excluded moderate-severe dementia								
Christofoletti 2008	14.9	2.2	12	14.8	1.3	17	11.4%	0.06 [-0.68, 0.80]
Eggermont 2009a	0.24	0.78	51	0.2	0.63	46	15.9%	0.06 [-0.34, 0.45]
Eggermont 2009b	0.07	0.37	30	0.47	0.97	31	14.4%	-0.53 [-1.05, -0.02]
Hwang 2010	28.9	11.86	10	24	14.68	8	9.1%	0.35 [-0.58, 1.29]
Kemoun 2010	30.38	7.66	16	22.23	8.37	15	11.2%	0.99 [0.24, 1.74]
Van de Winckel 2004	15.33	4.44	15	11	4.3	9	9.8%	0.95 [0.08, 1.83]
Volkers 2012	0.14	0.5	50	0.38	0.89	38	15.6%	-0.34 [-0.77, 0.08]
Vreugdenhil 2012	23.9	5	20	19	7.7	20	12.6%	0.74 [0.10, 1.38]
Subtotal (95% CI)			**204**			**184**	**100.0%**	**0.21 [-0.18, 0.61]**
Heterogeneity: Tau² = 0.21; Chi² = 23.15, df = 7 (P = 0.002); I² = 70%								
Test for overall effect: Z = 1.07 (P = 0.28)								



Cochrane Database of Systematic Reviews
15 APR 2015 DOI: 10.1002/14651858.CD006489.pub4
http://onlinelibrary.wiley.com/doi/10.1002/14651858.CD006489.pub4/full#CD006489-fig-0004

Exercise 1 (cont.): Questions to discuss in small groups

- What would you conclude from this analysis?
  - Are the results consistent across trials?
- Do you think more trials are worthwhile? (Would you need other information before you could answer this?)
- If you think a further trial(s) are worthwhile, how could the review help inform their design?
- How big would a future trial have to be to have a reasonable chance (i.e. moderate power, say 80%) of producing a “significant” treatment effect?
- If new evidence starts to emerge, when should an updated meta-analysis be conducted?

- The three primary outcomes (**cognition**, ADLs, and depression) were all rated as **very low on quality** of evidence due to **serious risk of bias, inconsistency, indirectness, and imprecision, and potential publication bias** in some or all of these outcomes (see GRADE, Summary of findings for the main comparison).
  - Should the meta-analysis acknowledge/incorporate this information in some way?
- **Additional well designed trials** that are **conducted in the community setting**, which is where most people with dementia live, and that examine **outcomes** of relevance to people with dementia (e.g. **cognition**, ADLs, depression, neuropsychiatric symptoms, quality of life and mortality), family caregiver outcomes (e.g. caregiver burden, quality of life, and mortality) and economic analysis of visits to emergency departments, acute care settings, and cost of residential care **are also needed.**
- Clinical **researchers should** make a practice of ensuring that their trials ….. **provide information** on the **randomization process** (sequence generation and allocation concealment), **blinding of outcome assessors, attrition rates and reasons for drop-outs** from both treatment and control groups, **rate of adherence** to the exercise programs and **reasons for withdrawal**, and adverse events to the exercise programs in published articles, or be willing to share this information with reviewers when contacted. **Providing statistically appropriate data** (e.g. end point means and standard deviations) would also ensure that the trial **results can be incorporated into meta-analysis**.

Garner et al. When and how to update systematic reviews: consensus and checklist. *BMJ 2016*

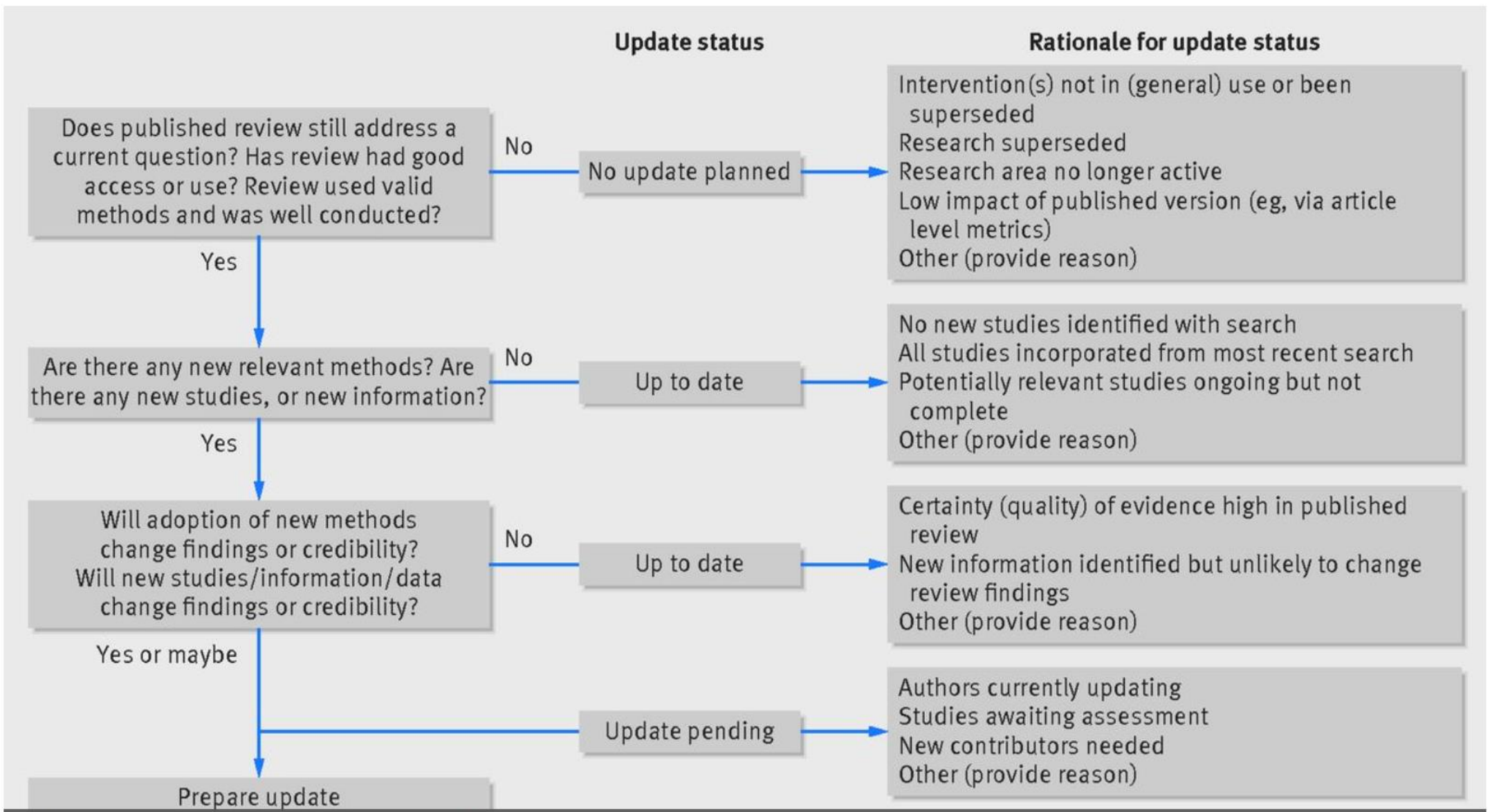
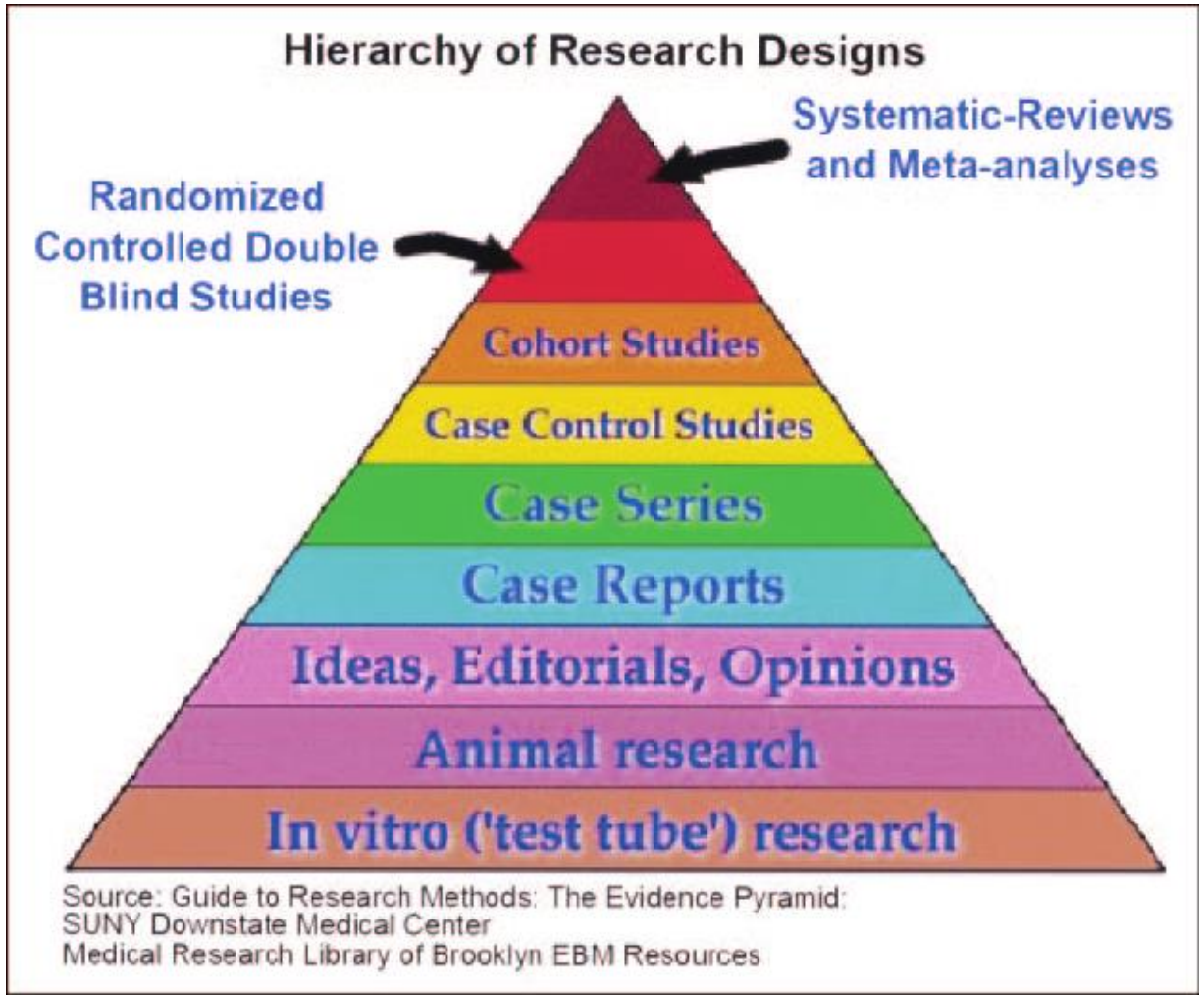


Fig 1 Decision framework to assess systematic reviews for updating, with standard terms to report such decisions

Session 2:
The coherent design and analysis of future trials

The Hierarchy Of Research Evidence

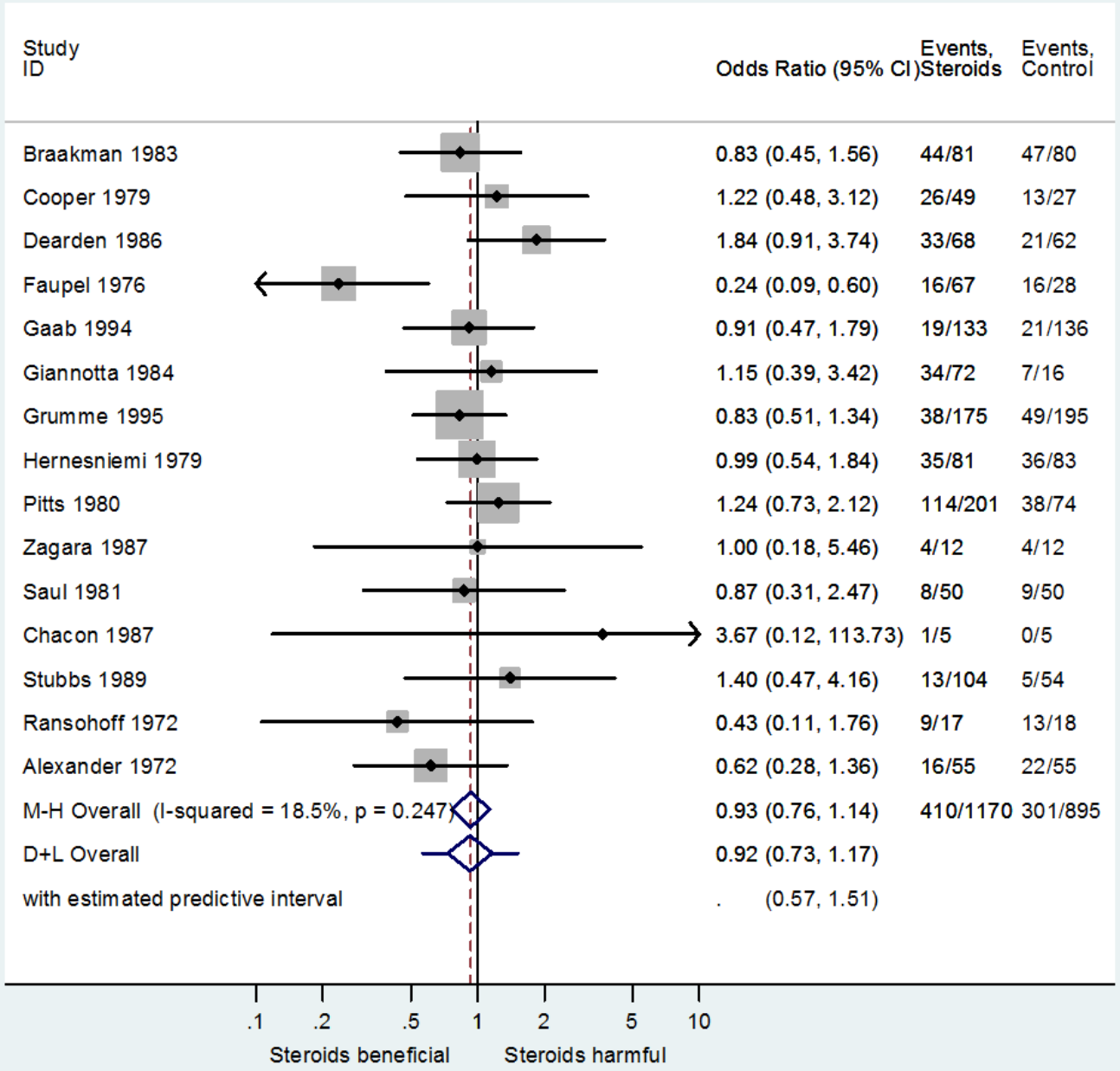


- Commonly accepted that **systematic review** and **meta-analysis** are at the **top of the hierarchy**
  - (Who invented this hierarchy??)

- When **new, primary, research** is carried out, should we consider its likely **impact on an existing meta-analysis?**
  - Seems illogical not to?
  - Given the order of the hierarchy, **updated meta-analysis** will (often) have **more impact than the results of the trial reported separately**(?)
  - However, **sample size** / power considerations usually relate to the **analysis of the trial on its own** & as a whole **not its impact on the existing evidence base**
  - Is this **incoherent**?

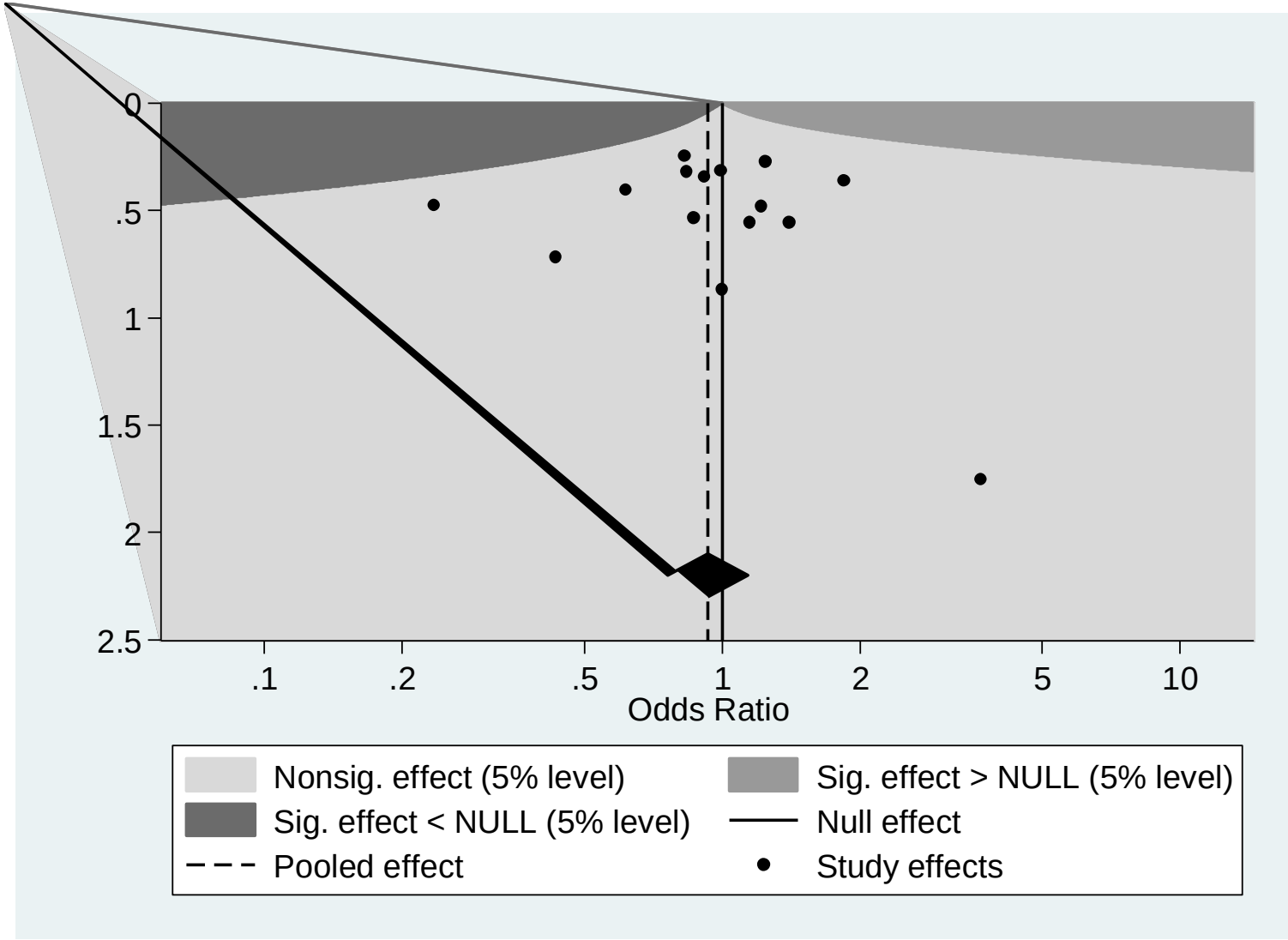
How to estimate trial sample size to power an updated meta-analysis?

- Example: Steroids after serious brain injury
  - (Meta-analysis before the big CRASH trial (randomised 10,008 persons), result: RR 1.15, 95% CI 1.07-1.24; steroids harmful – but that's another story)



What effect size/precision combination would a new trial have to have to make the fixed effect meta-analysis statistically significant?

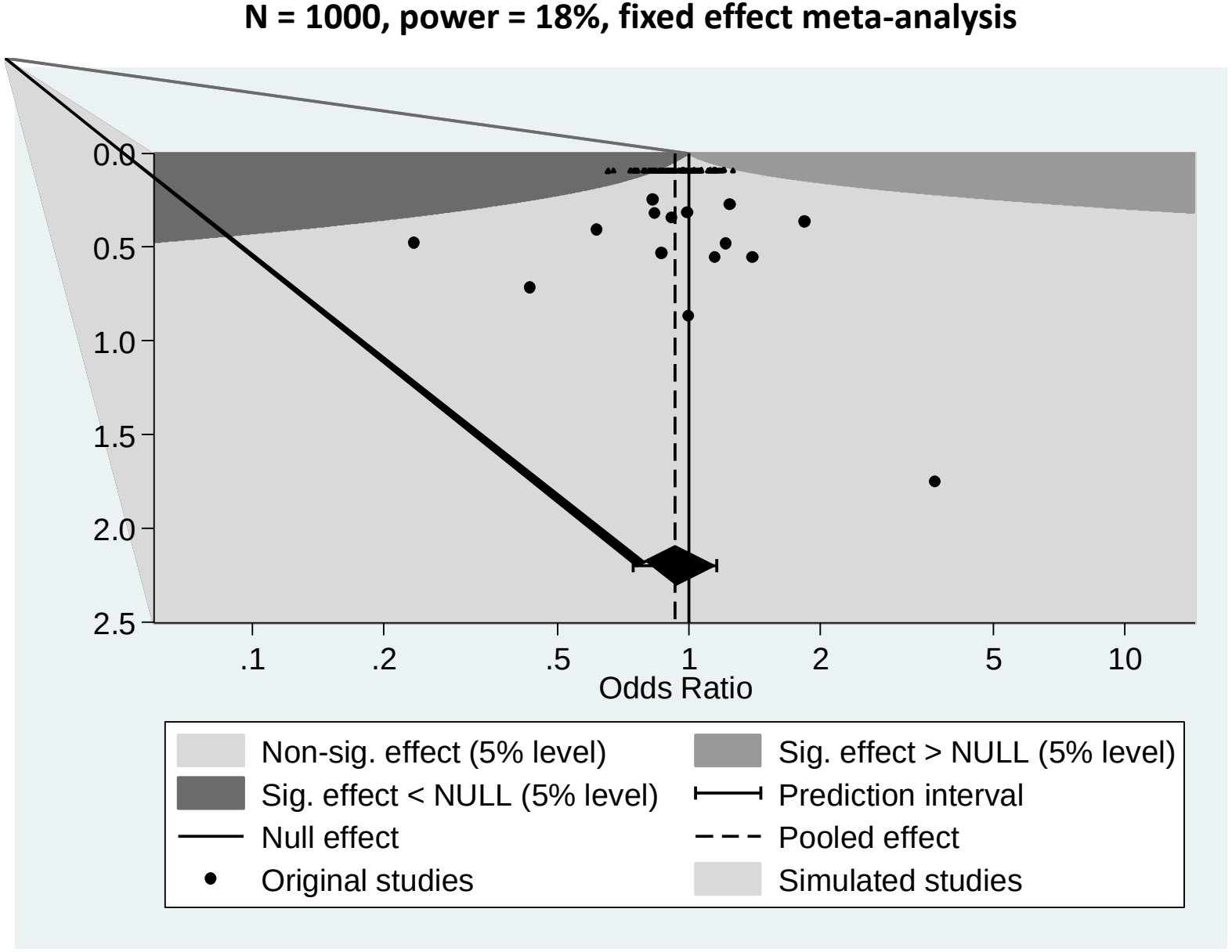
- **Funnel plot** (effect size vs standard error scatter plot)
- Dots existing studies.
- Shaded areas at top where a new study would have to be located i.e. very precise & a large effect size helps)
- Conclusion, only a very large amount of evidence would change conclusions of meta-analysis



Calculating power for a new study: Fixed Effects

- If you **assume** the **effect estimate from the (fixed effect) meta-analysis is** an unbiased estimate of **the true treatment effect**
- Predict effect of new study using meta-analysis model
  - Allowing for **uncertainty in meta-analysis** result & random error
  - **Simulated studies** are **small dots** on figure – **sample size = 1000**.
  - Only **18% of "dots" in dark regions**, power to change conclusions of meta-analysis

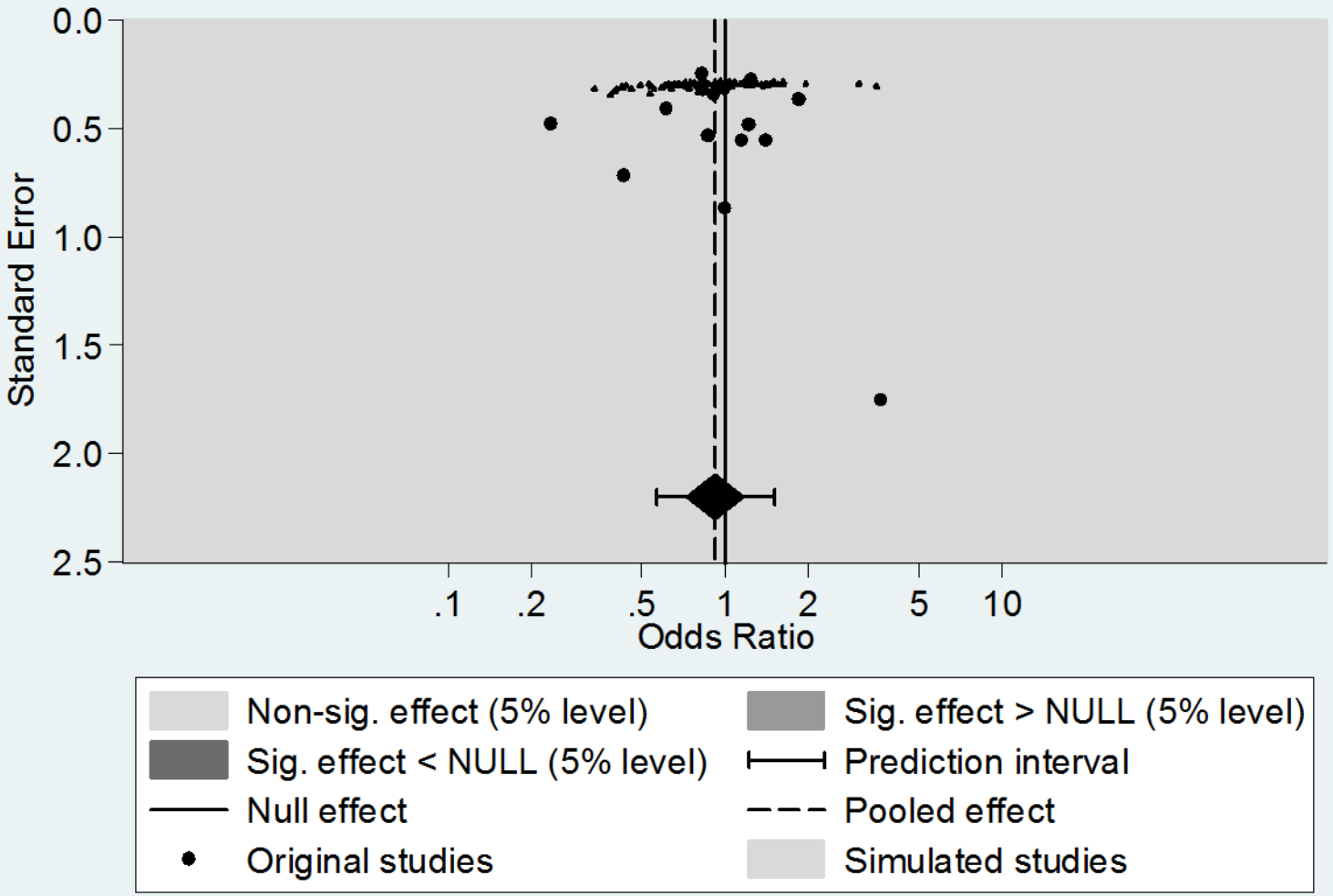
    = 18% (v.low) for n=1000



Calculating power for a new study: Random Effects

- Do the same thing but using a **random effects model**
- Now use **predictive distribution** to estimate new study
  - It is the **line extending from diamond**
  - It **incorporates heterogeneity**
- (Important aside: the **prediction interval** is an **important summary** of **random effect meta-analysis** & under used. If you don't know about it, I recommend reading (IntHout et al., 2016))

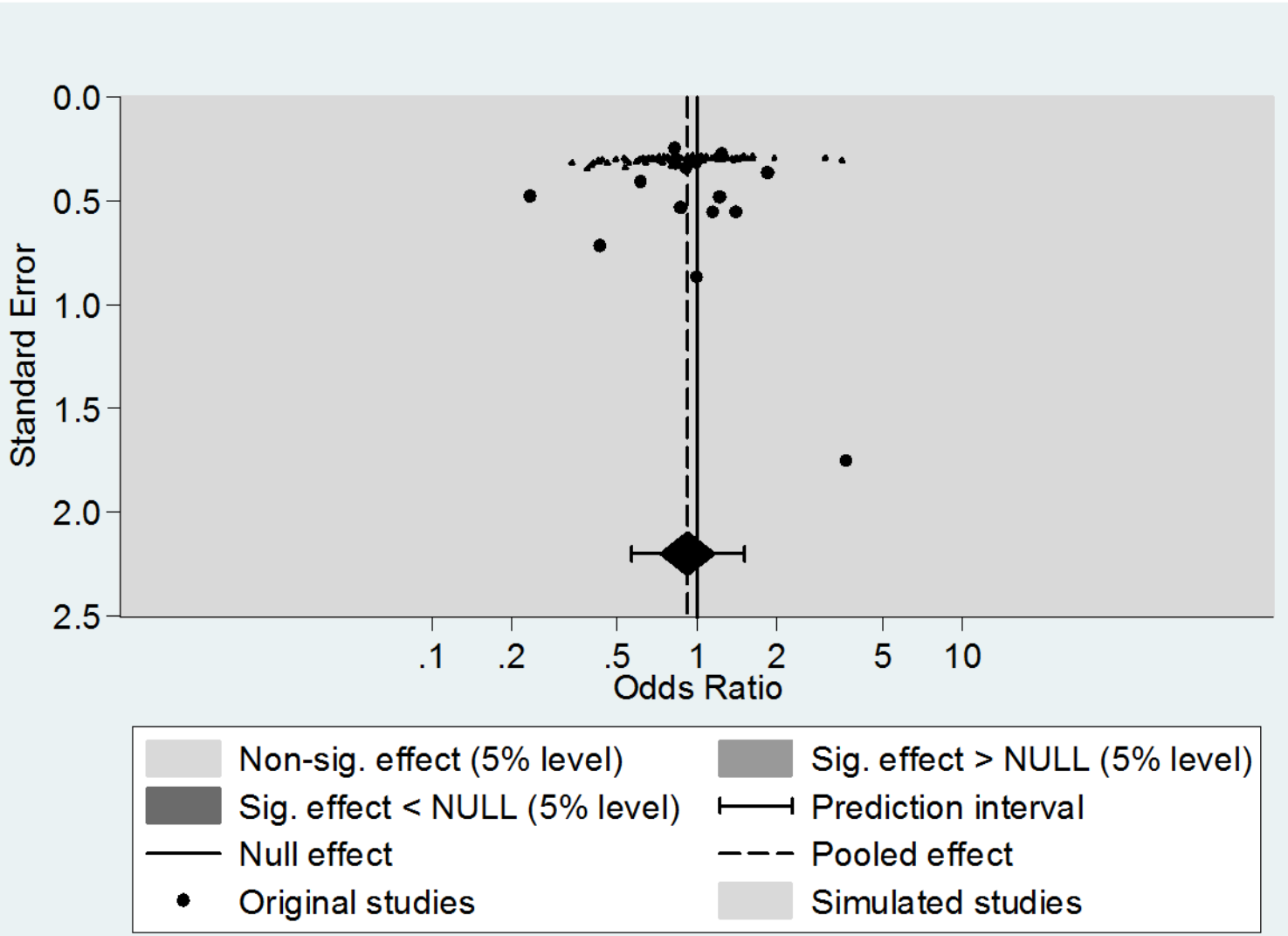
Power = 0%, random effect meta-analysis



Calculating power for a new study: Random Effects

- Do the same thing but using a **random effects model**
- Now use **predictive distribution** to estimate new study
  - It is the **line extending from diamond**
  - It **incorporates heterogeneity**
- Plot shows **no dark shaded regions**
  - **Power = 0% for any sample size!!!**
  - **No one study can change conclusions of this meta-analysis!!**
  - **I think this is useful to know**

Power = 0%, random effect meta-analysis



How can power of a new study to change conclusions of the meta-analysis be 0??

- Weighting of study in random effect model:

$$\frac{1}{variance\ of\ study\ estimate + between\ study\ heterogeneity}$$

- Therefore for a huge study - with essentially 0 variance - maximum weighting:

$$\frac{1}{between\ study\ heterogeneity}$$

- Which may down-weight influence considerably

Updating Systematic Reviews

- We recently got guidance(!)

Garner et al., When and how to update systematic reviews: consensus and checklist *BMJ* 2016; **354** doi: http://dx.doi.org/10.1136/bmj.i3507 (Published 20 July 2016)

- Draws on previous attempts including: (Takwoingi et al. 2013)

Combined a **qualitative** and a **quantitative tool**

- Quantitative tool based on methods for estimating power to change conclusions of a meta-analysis (as explained in the previous slides)
  - Requires (only) sample sizes of any new trials
  - For the given sample sizes it estimates power (rather than fixing power and estimating sample size as before)
  - Which review should we update next? The one with the greatest power to change the conclusions of the primary meta-analysis
  - **CRUCIALLY, if every trial was adequately powered to update the meta-analysis then update decisions probably disappear (i.e. always update)**
    - We are some way off this ideal . . . . . .

The hierarchy of evidence revisited!

M Hassan Murad et al. New Evidence Pyramid. *Evid Based Med* doi:10.1136/ebmed-2016-110401

The proposed new evidence-based medicine pyramid.

- **Meta-analysis is not a type of evidence, but a lens under which to look at the evidence**
  "The systematic review and meta-analysis are tools to consume and apply the evidence by stakeholders."
- If you take meta-analysis off the top of the hierarchy, then the argument for powering trials to inform an updated meta-analysis are weakened

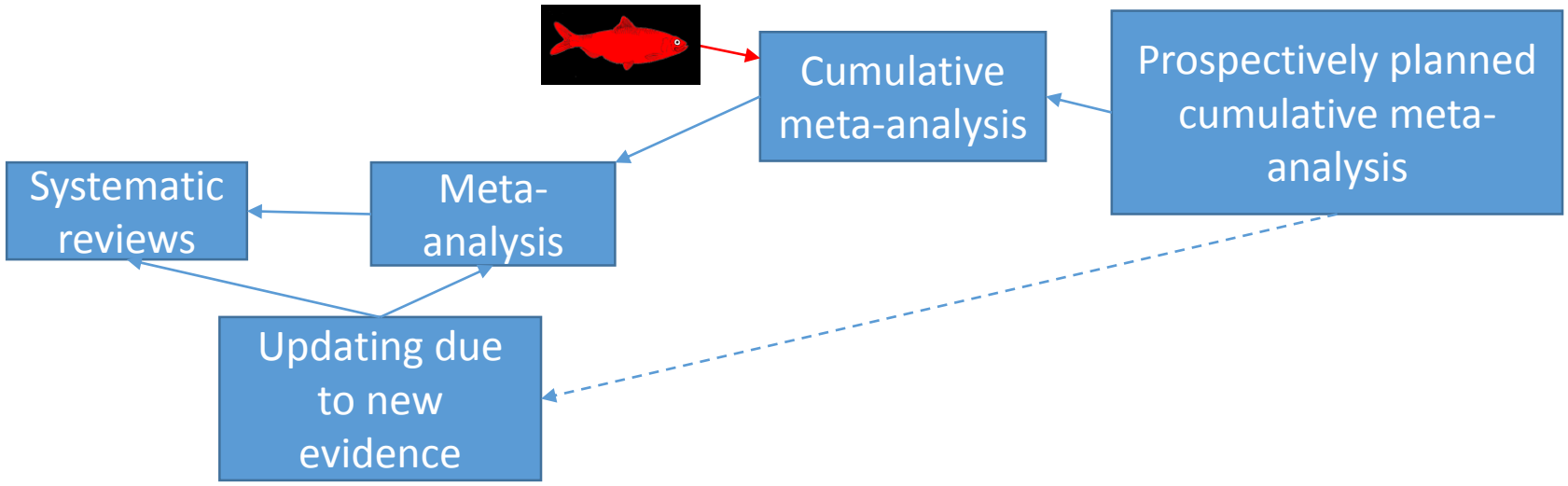
The hierarchy of evidence revisited! *(cont.)*

- Implications are to take the emphasis off the “overall answer” when there is heterogeneity between trials
- I think we were always kidding ourselves that we could interpret and apply the results of a random effects meta-analysis to clinical practice anyway (?)
- Exploring why results between studies differ should be the priority
- If we know what is causing the heterogeneity (differences in patients, interventions, poor study conduct etc), alternative estimates of effect are more relevant (Welton & Ades, 2011)
  - Ultimately leads to different sample size calculations for trials, but could still focus on impact on the systematic review/meta-analysis

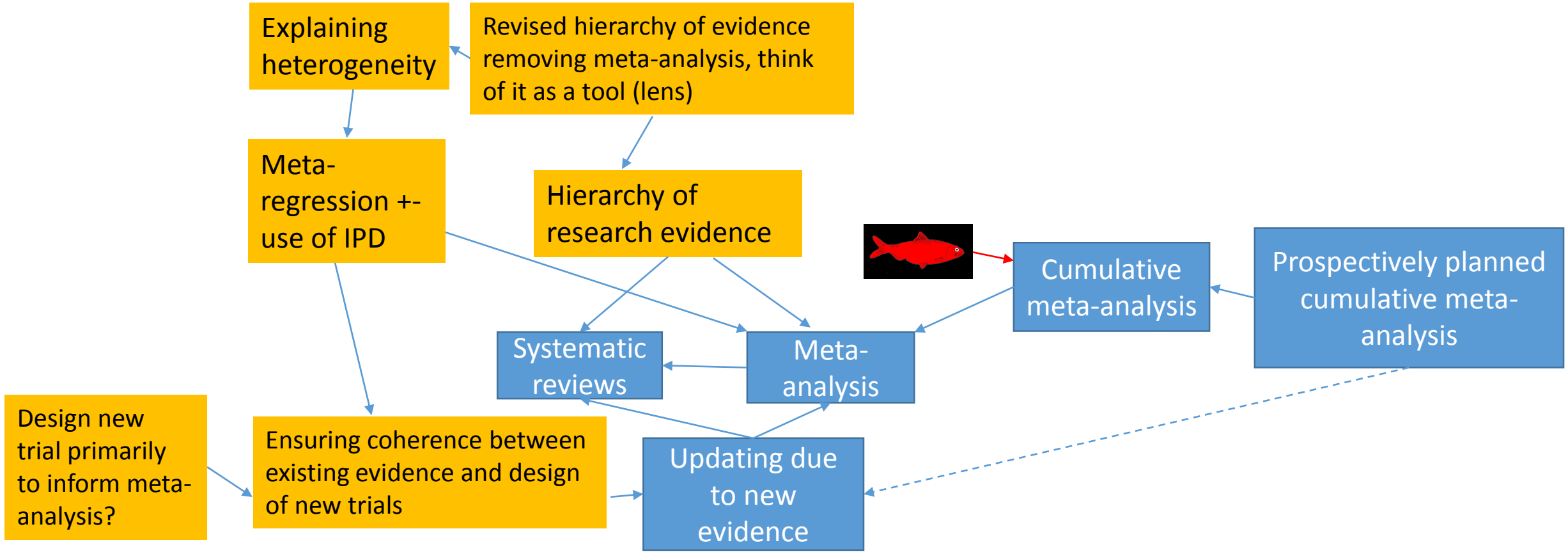
The proposed new evidence-based medicine pyramid.

- ?? If we can’t explain heterogeneity in existing studies, could it be best to “wipe the slate clean” and start again with the gathering of evidence??
  - Existing heterogeneity in a meta-analysis downweights the contribution to future evidence considerably
  - Adding good evidence to a “mess” still results in a “mess”

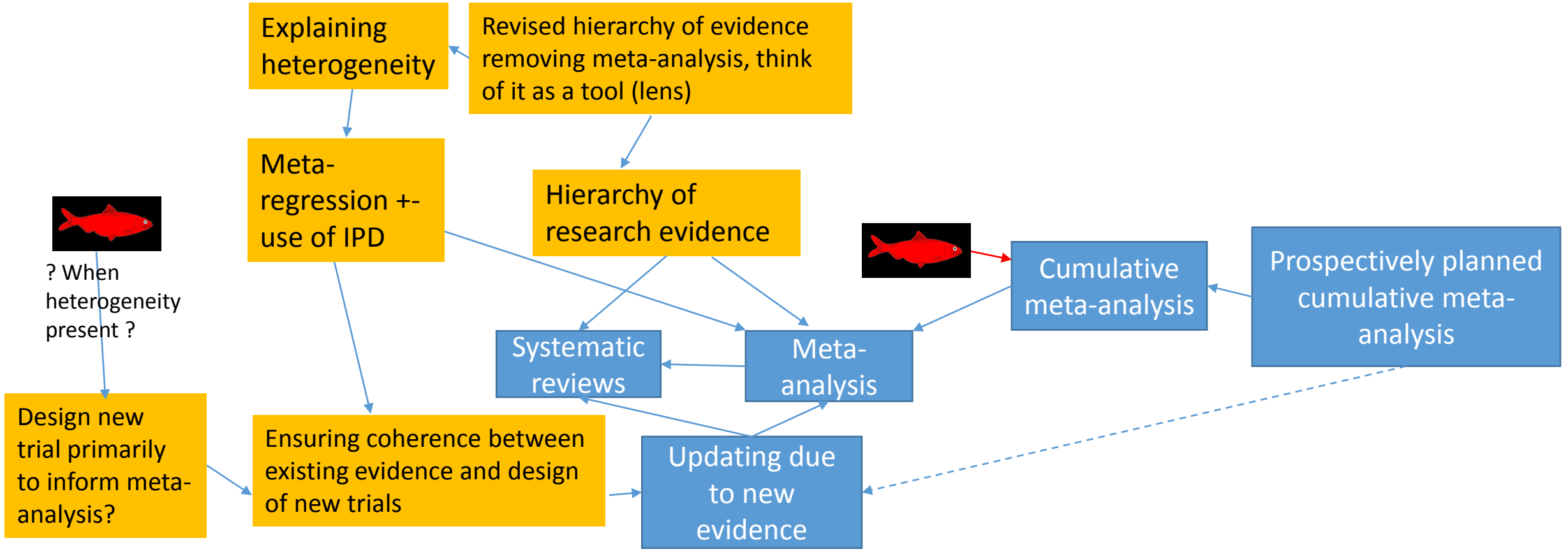
Schematic – Key concepts and approximate relationships



Schematic – Key concepts and approximate relationships



Schematic – Key concepts and approximate relationships



Exercise 2

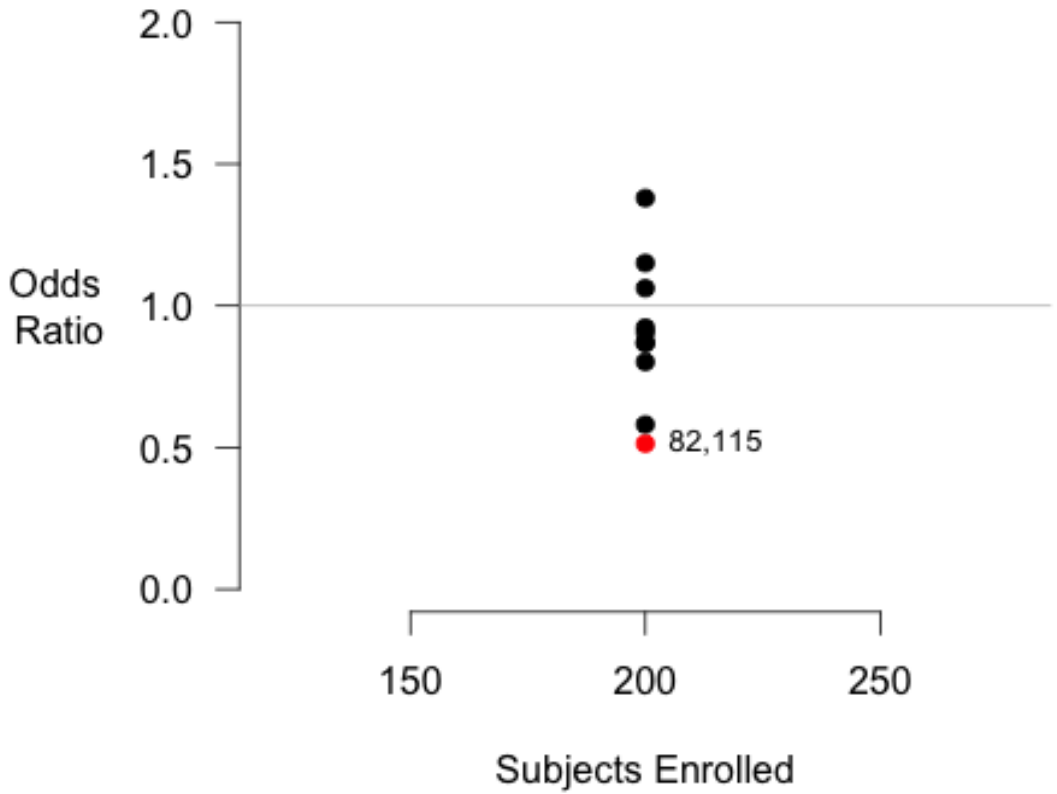
- Think how the **recommendations for (future) research sections** of a **review** could be augmented to give **more specific/explicit information**
  - With respect to recommendations for **updating the review** (when, how much evidence etc)
  - With respect to **designing new trials** (do we need more trials?, if we do, what sort of size will they have to be to have a meaningful impact on the review? Any support for my “start again” recommendation in areas that are a “dog’s dinner”??
- Have you **any concerns / objections** to the **methods/ideas presented**?
- Why do you think they have **not been used more frequently**?
- Do you agree **different stakeholders** may **value different analyses** of the data in Cochrane reviews? (If so think of hypothetical examples)
  - Should Cochrane meta-analyses be made flexible enough to **allow the user to conduct their own analysis** of the data identified?

Session 3:
Sequential analysis and multiplicity

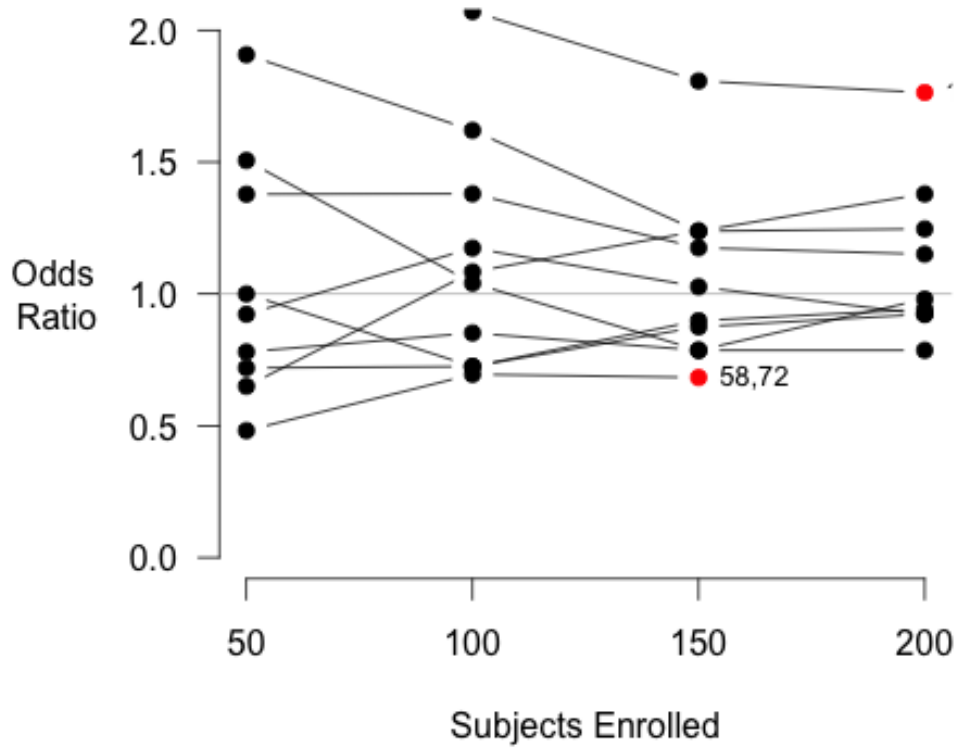
Imagine a (simulated) clinical trial

- 200 patients per arm
- The probability of response is 50% in both arms
- There is no difference in the treatment effect

We may get a statistically significant result by chance



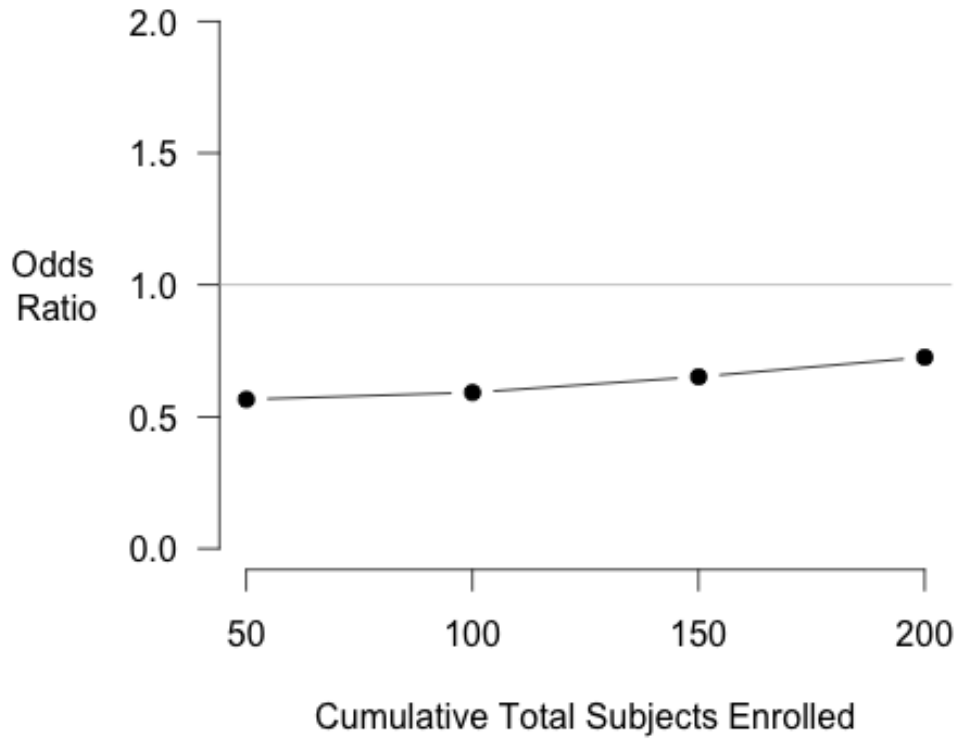
If we conduct interim analyses, the chance increases



Hence, trial analyses are carefully adjusted to take account of interim analysis

- P-Values are adjusted to take account of the effects of multiple ‘looks’ at the data

A series of meta-analyses also represents multiple ‘looks’



Published by Oxford University Press on behalf of the International Epidemiological Association
© The Author 2008; all rights reserved. Advance Access publication 29 September 2008

International Journal of Epidemiology 2009;**38**:287–298
doi:10.1093/ije/dyn188

Apparently conclusive meta-analyses may be inconclusive—Trial sequential analysis adjustment of random error risk due to repetitive testing of accumulating data in apparently conclusive neonatal meta-analyses

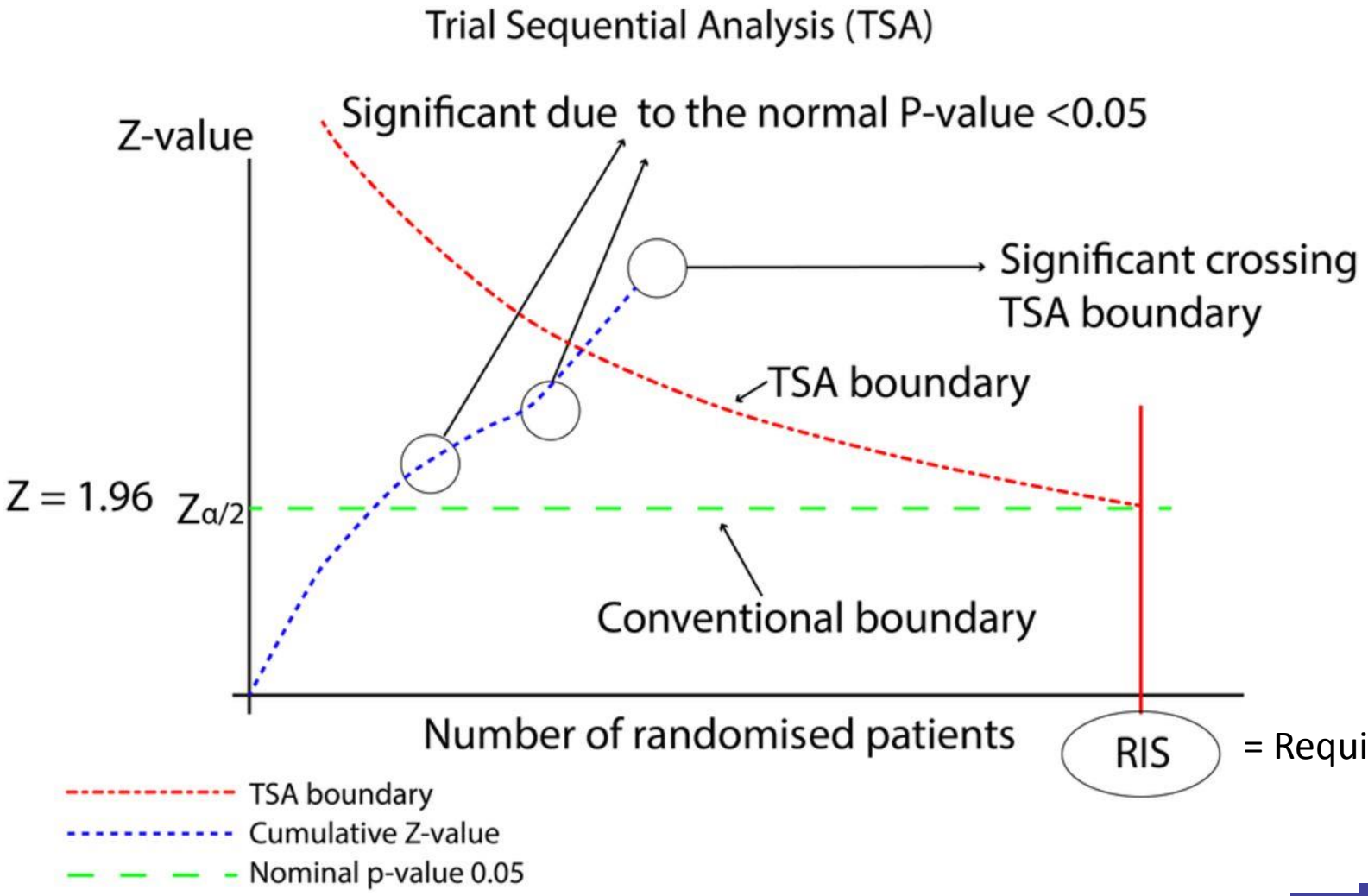
Jesper Brok,* Kristian Thorlund, Jørn Wetterslev and Christian Gluud

Accepted 13 August 2008

Background Random error may cause misleading evidence in meta-analyses. The required number of participants in a meta-analysis (i.e. information size) should be at least as large as an adequately powered single trial. Trial sequential analysis (TSA) may reduce risk of random errors due to repetitive testing of accumulating data by evaluating meta-analyses not reaching the information size with monitoring boundaries. This is analogous to sequential monitoring boundaries in a single trial.

Graphical representation of trial sequential analysis (TSA)

- Adapts ideas used for individual clinical trials when conducting interim analyses for multiple looks at the data



Georgina Imberger et al. BMJ Open 2016;6:e011890

©2016 by British Medical Journal Publishing Group

BMJ Open

Assumptions of Trial Sequential Analysis

- Assumption that data will accumulate until the **required** (sufficient) **information size** is surpassed
- Requires a *heterogeneity adjustment factor:*

". . .equal to the ratio of the total variance in a random-effects model meta-analysis and the total variance in a fixed-effect model meta-analysis."

  - This seems at odds with the power calculations shown in Session 2 where power can be 0 for a huge new single study, which would surpass the required information size (??)

Requires estimates of heterogeneity and effect size which are "correct" when each update is done.

Stop Press !!! New Paper on Trial Sequential Analysis:

Imberger G, Thorlund K, Gluud C, Wetterslev J. **False-positive findings in Cochrane meta-analyses with and without application of trial sequential analysis: an empirical review. (2016)** BMJ Open http://bmjopen.bmj.com/content/6/8/e011890.full

- Used TSA retrospectively
- Took 100 meta-analyses with a binary outcome, a negative result and “sufficient power” (As defined by their calculation of the Required Information Size)
- Re-analysed adding one study at a time using cumulative meta-analysis (session 1)
- Counted how many “false positives” (significant beneficial treatment effect) would have been obtained in one of the updates
  - This happened in 7 of the 100 meta-analyses

- An important subsidiary finding of the study:
  - Overall, 1.8% (95% CI 1.3% to 2.3%) of Cochrane reviews were sufficiently powered(!)

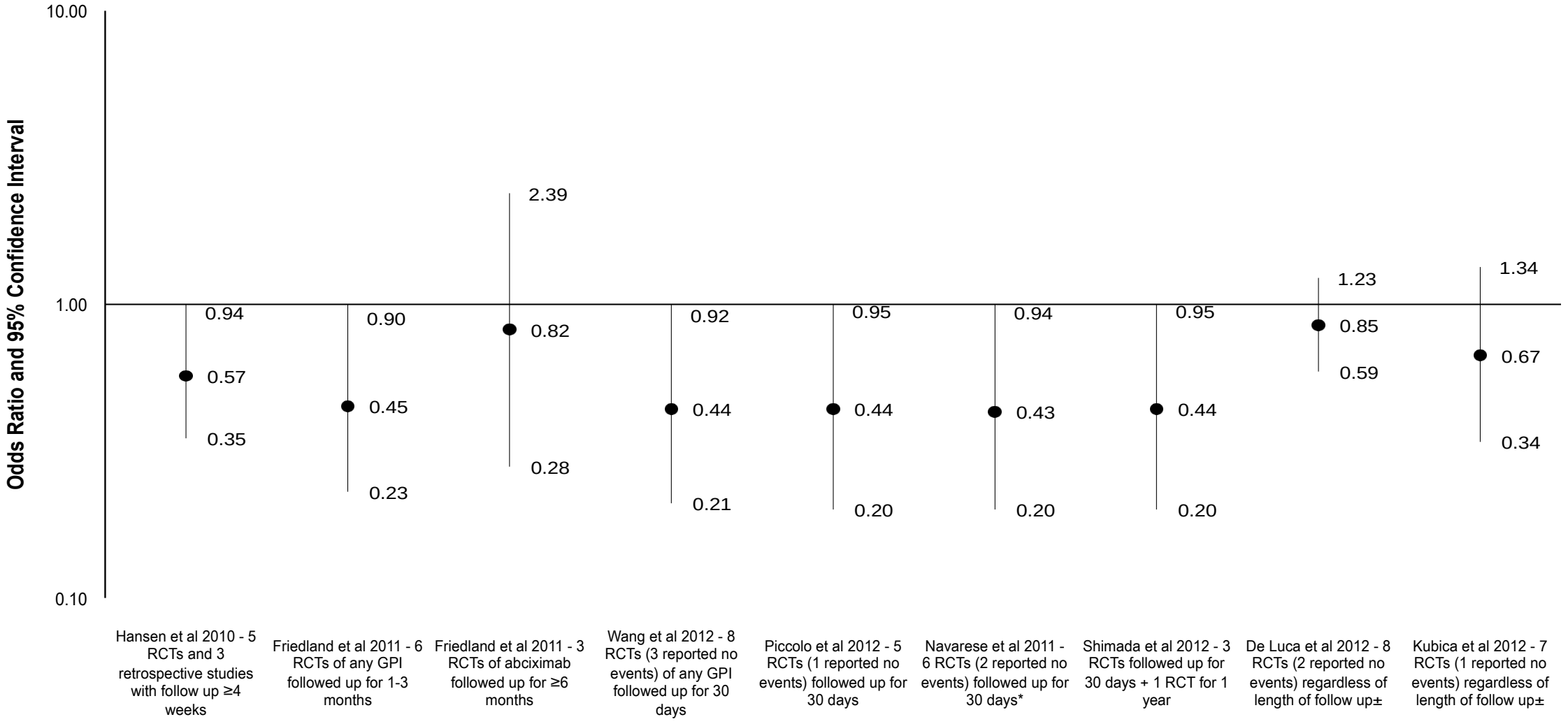
Case Study

Case Study: Intracoronary versus intravenous abciximab in ST-elevation myocardial infarction

Individual Trials

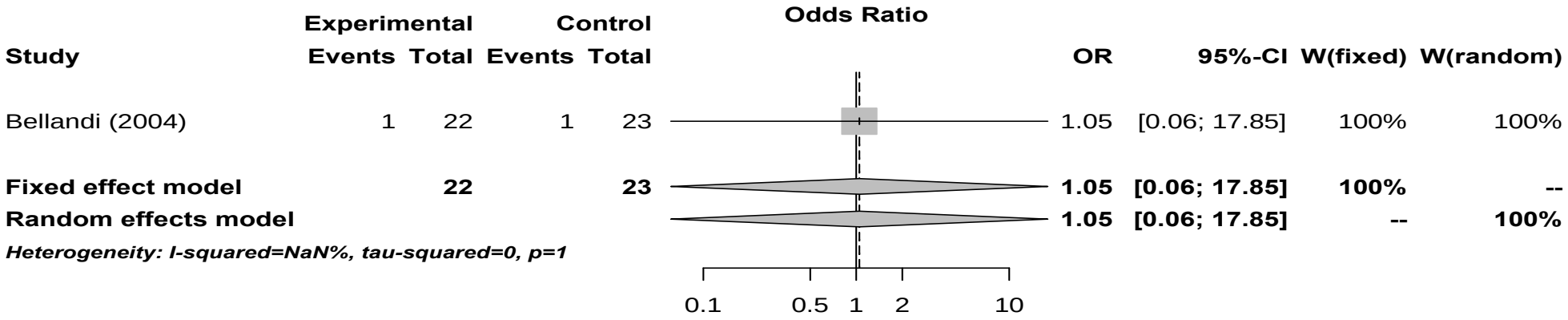
Year	Study	FollowUp
2004	Bellandi (2004)	30 days
2006	Galache Osuna (2006)	>1 year
2008	LIPSIA (2008)	30 days
2009	Dominguez-Rodriguez (2009)	30 days
2010	Iversen (2010)	30 days
2010	EASY-MI (2010)	1 year
2010	CICERO (2010)	30 days
2010	CRYSTAL AMI (2010)	30 days
2011	LIPSIA (2008)	6 months
2011	Iversen (2010)	1 year
2012	AIDA STEMI (2012)	90 days

Alongside these trials, a series of meta-analyses have been published

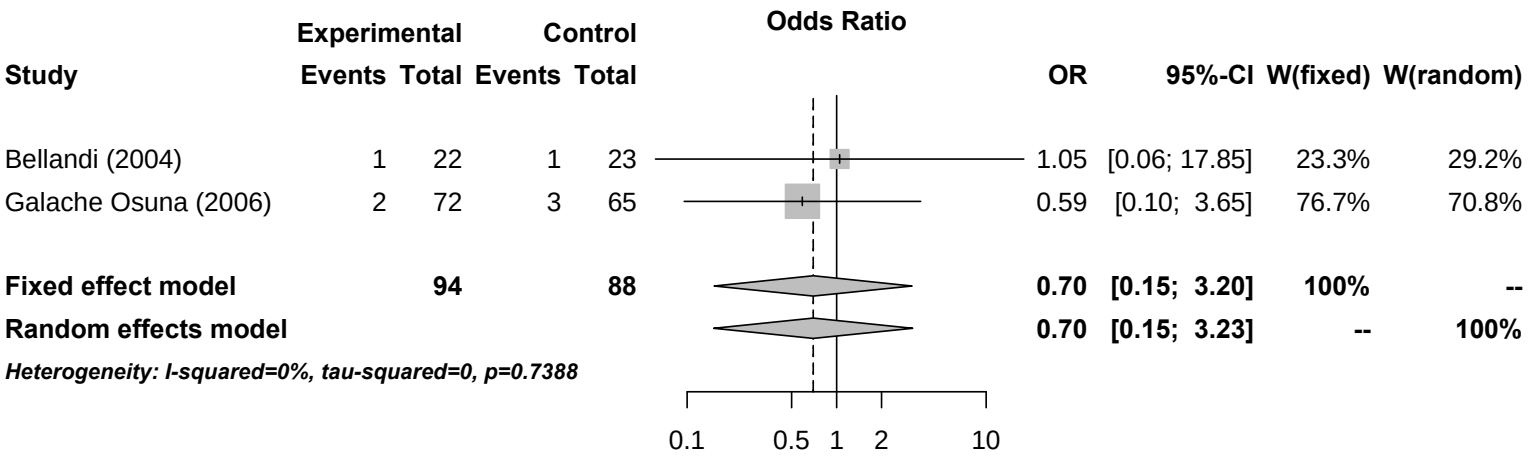


Mortality

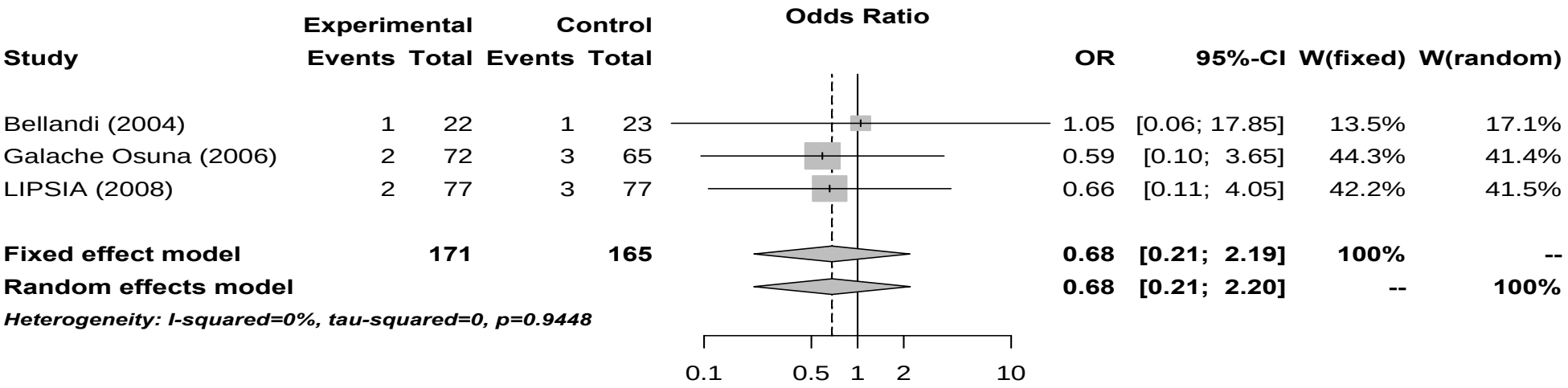
Trials to 2004: no published meta-analyses (cumulative results not statistically significant, no evidence of heterogeneity)



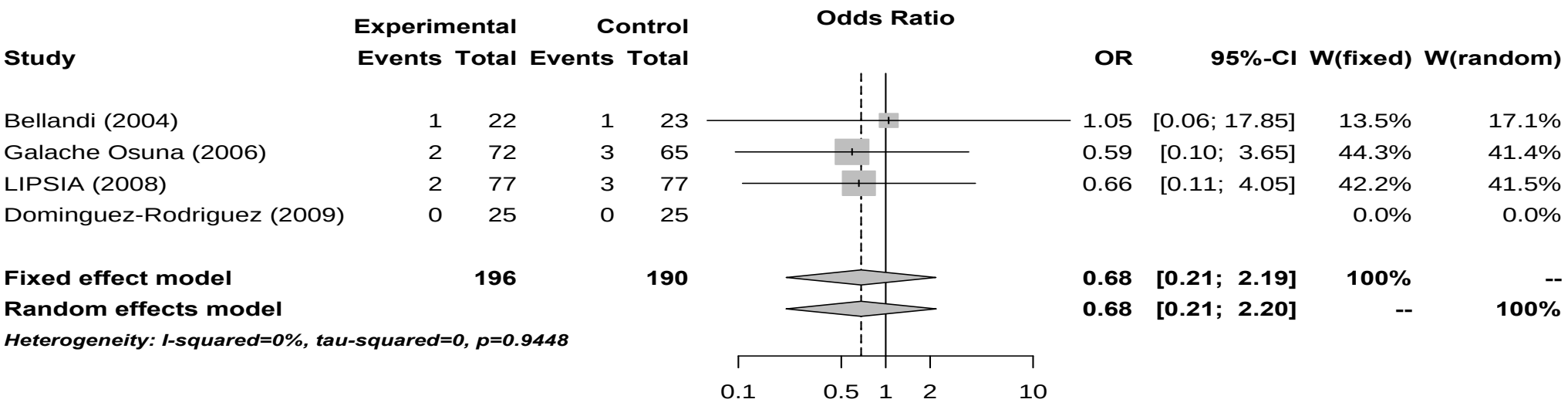
Trials to 2006: no published meta-analyses (cumulative results not statistically significant, no evidence of heterogeneity)



Trials to 2008: no published meta-analyses (cumulative results not statistically significant, no evidence of heterogeneity)



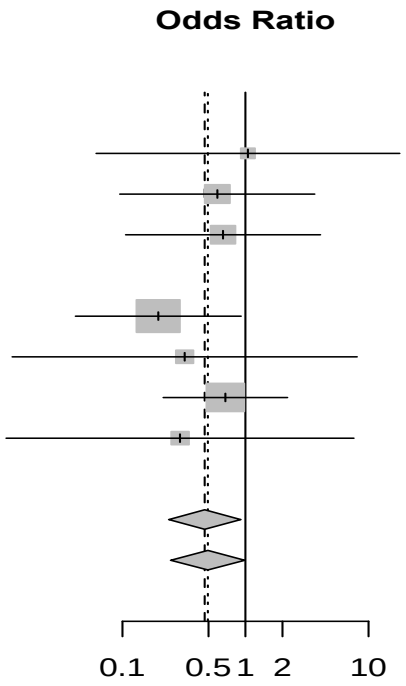
Trials to 2009: no published meta-analyses (cumulative results not statistically significant, no evidence of heterogeneity)



Trials to 2010: First published meta-analysis (cumulative results statistically significant, no evidence of heterogeneity)

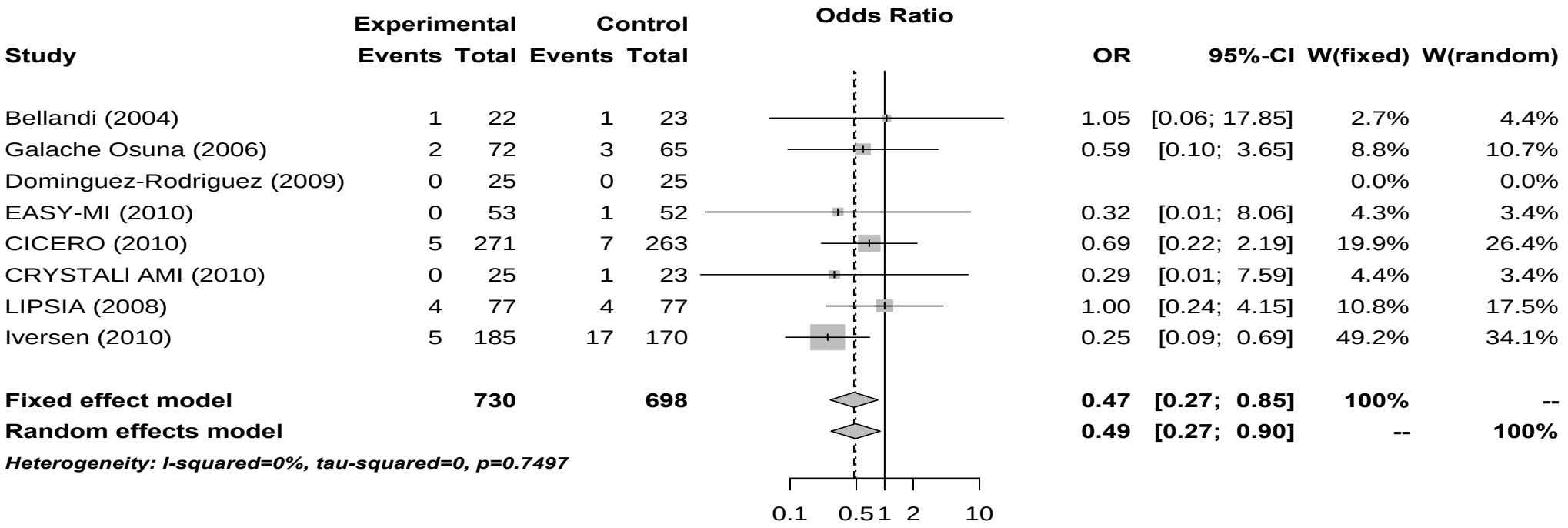
Study	Experimental Events	Experimental Total	Control Events	Control Total	OR	95%-CI	W(fixed)	W(random)
Bellandi (2004)	1	22	1	23	1.05	[0.06; 17.85]	3.6%	6.0%
Galache Osuna (2006)	2	72	3	65	0.59	[0.10; 3.65]	11.7%	14.5%
LIPSIA (2008)	2	77	3	77	0.66	[0.11; 4.05]	11.2%	14.6%
Dominguez-Rodriguez (2009)	0	25	0	25			0.0%	0.0%
Iversen (2010)	2	185	9	170	0.20	[0.04; 0.92]	35.4%	20.1%
EASY-MI (2010)	0	53	1	52	0.32	[0.01; 8.06]	5.7%	4.6%
CICERO (2010)	5	271	7	263	0.69	[0.22; 2.19]	26.6%	35.7%
CRYSTALI AMI (2010)	0	25	1	23	0.29	[0.01; 7.59]	5.8%	4.6%
Fixed effect model		**730**		**698**	**0.47**	**[0.24; 0.92]**	**100%**	**--**
Random effects model					**0.49**	**[0.25; 0.99]**	**--**	**100%**

Heterogeneity: I-squared=0%, tau-squared=0, p=0.8917



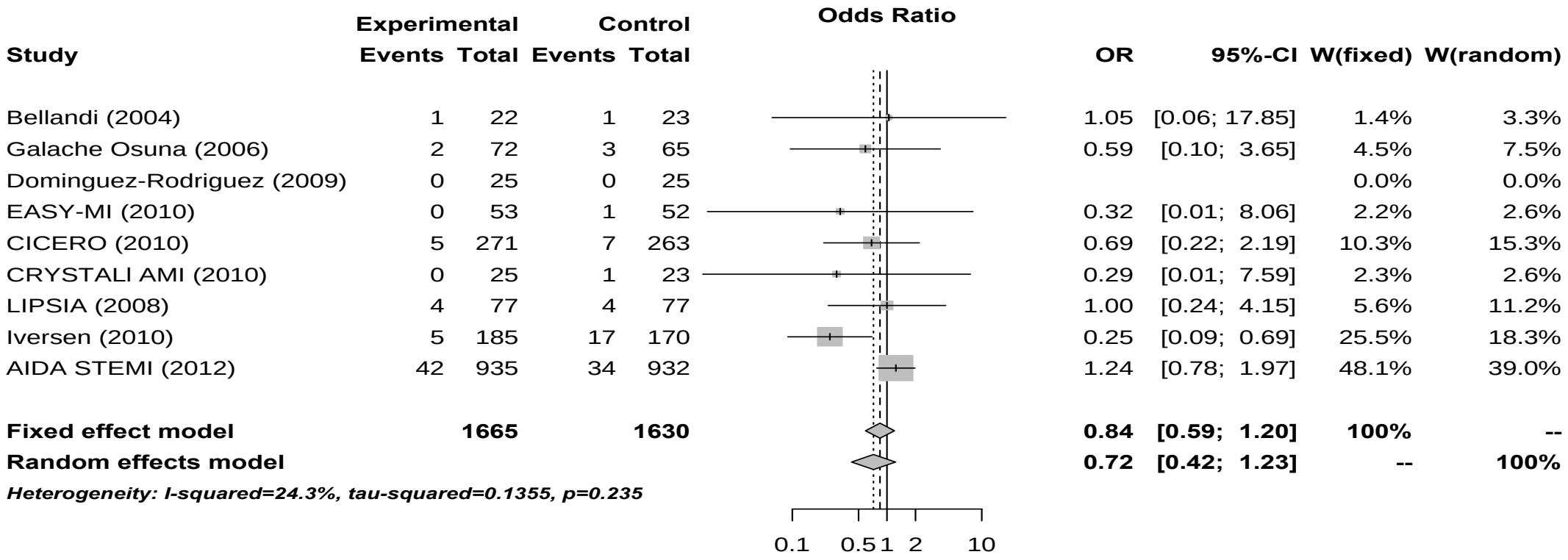
Should we stop here?

Trial to 2011: Three published meta-analyses (cumulative results statistically significant, no evidence of heterogeneity)



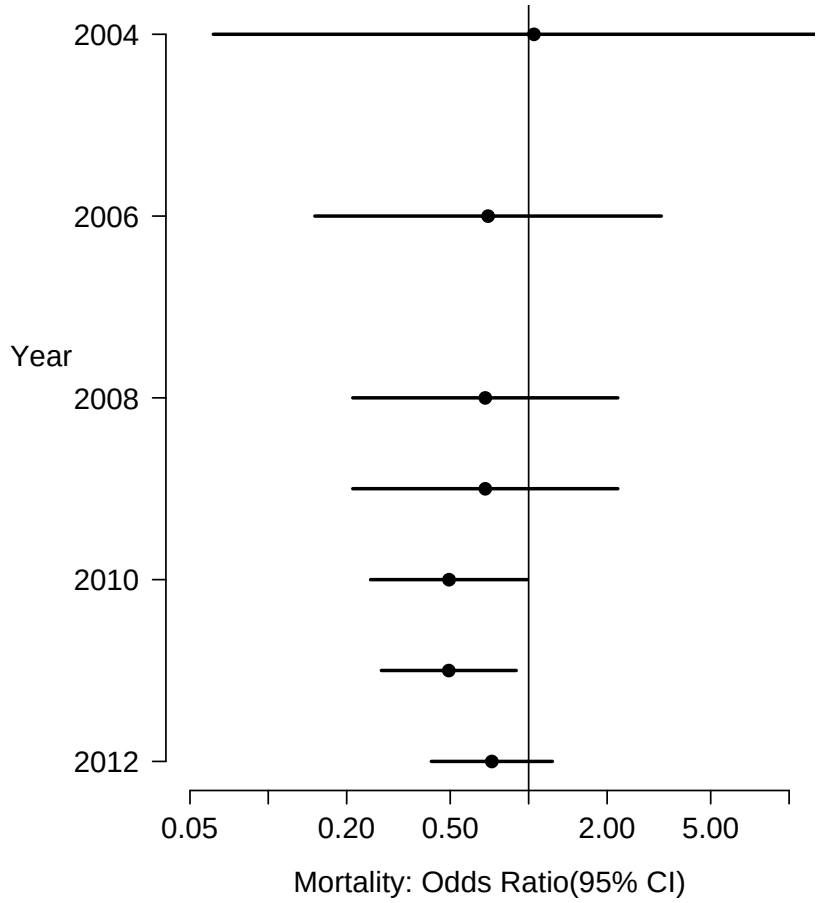
Should we stop here?

Trials to 2012 Eight published meta-analyses (cumulative results not statistically significant, evidence of heterogeneity)



Should we did stop here?

Conventional meta-analysis does show statistical significance during the sequence of trials



Research Article

Statistics in Medicine

Received 28 August 2008, Accepted 23 August 2010 Published online 28 December 2010 in Wiley Online Library

(wileyonlinelibrary.com) DOI: 10.1002/sim.4088

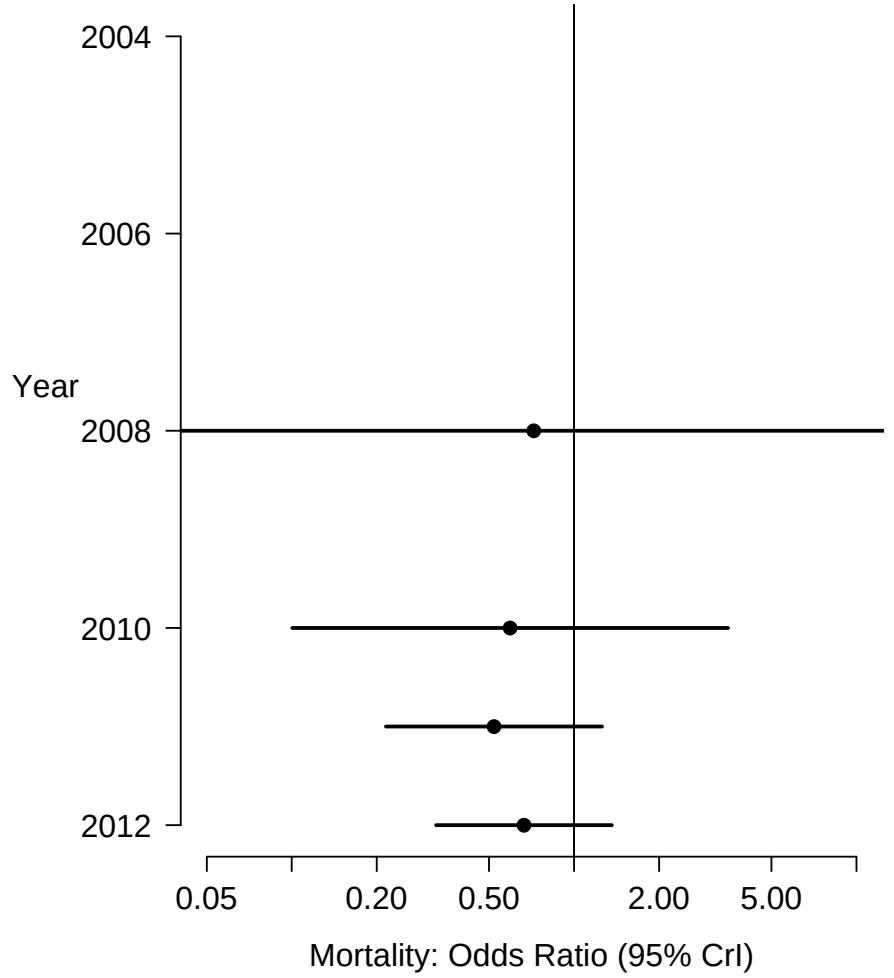
Sequential methods for random-effects meta-analysis

Julian P. T. Higgins,[a*†] Anne Whitehead[b] and Mark Simmonds[c]

Although meta-analyses are typically viewed as retrospective activities, they are increasingly being applied prospectively to provide up-to-date evidence on specific research questions. When meta-analyses are updated account should be taken of the possibility of false-positive findings due to repeated significance tests. We discuss the use of sequential methods for meta-analyses that incorporate random effects to allow for heterogeneity across studies. We propose a method that uses an approximate semi-Bayes procedure to update evidence on the among-study variance, starting with an informative prior distribution that might be based on findings from previous meta-analyses. We compare our methods with other approaches, including the traditional method of cumulative meta-analysis, in a simulation study and observe that it has Type I and Type II error rates close to the nominal level. We illustrate the method using an example in the treatment of bleeding peptic ulcers.
Copyright © 2010 John Wiley & Sons, Ltd.

Keywords: meta-analysis; sequential methods; cumulative meta-analysis; prospective meta-analysis; prior distributions

However, sequential methods meta-analysis didnot show statistical significance at any time point

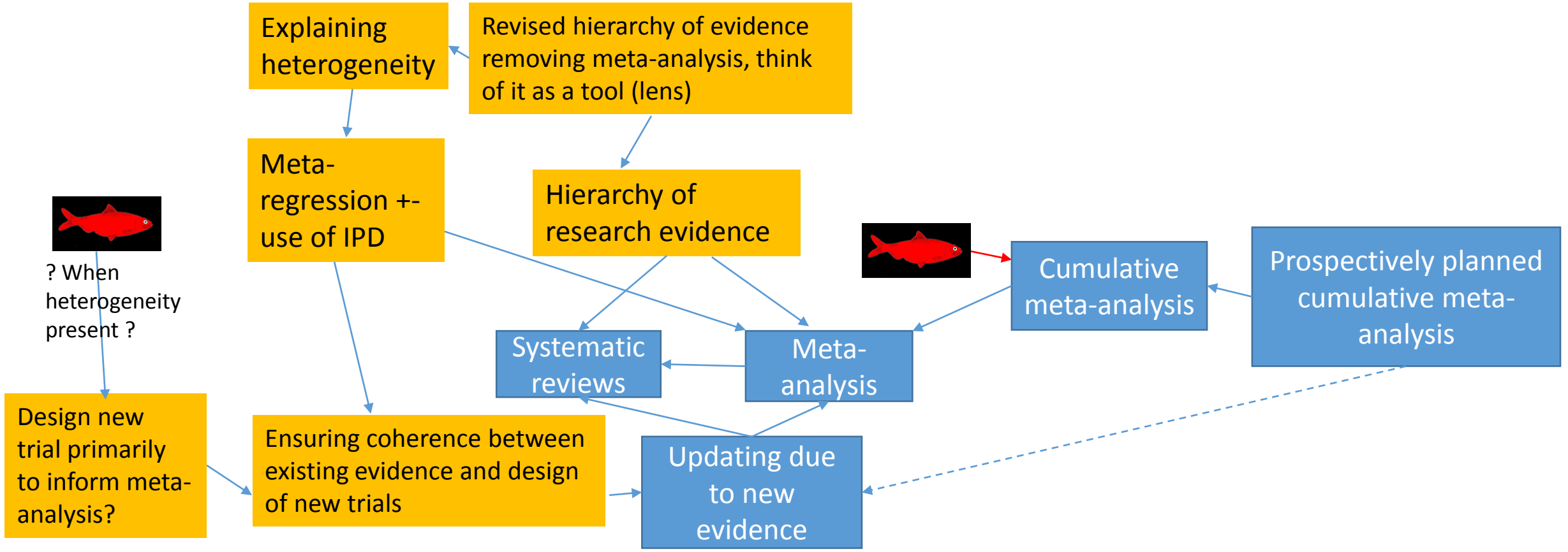


Discussion

- Meta-analyses are potentially/essentially sequential
- “Interim” meta-analyses may not be published
- If we do not account for sequential analysis we may under-estimate uncertainty and obtain biased treatment effects estimates

- Should we require “sequential” methods for meta-analysis?
- Would pre-regsistration help?

Schematic – Key concepts and approximate relationships



Schematic – Key concepts and approximate relationships

? When heterogeneity present ?

Explaining heterogeneity

Revised hierarchy of evidence removing meta-analysis, think of it as a tool (lens)

Meta-regression +- use of IPD

Hierarchy of research evidence

Cumulative meta-analysis

Prospectively planned cumulative meta-analysis

Systematic reviews

Meta-analysis

Design new trial primarily to inform meta-analysis?

Ensuring coherence between existing evidence and design of new trials

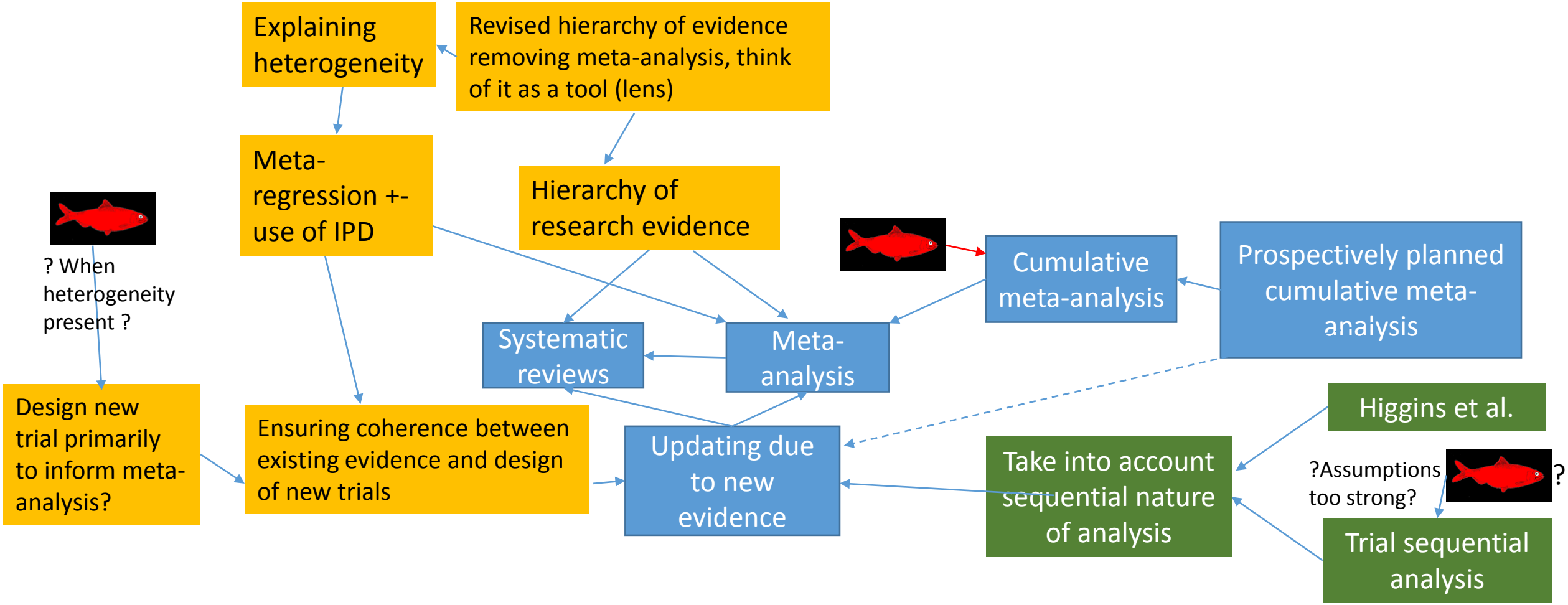
Updating due to new evidence

Take into account sequential nature of analysis

Higgins et al.

Trial sequential analysis

Schematic – Key concepts and approximate relationships



Overall conclusions of the workshop

- Evidence Based Research
  - Formalise the approach to the accumulation of evidence in order to improve quality and reduce wastage

- Issues we need to think (more) about include:
  1) Develop review updating strategies
  2) Extend / formalise how a current systematic review can help inform future trials

     - Including (among many other things) sample size calculations
  3) Address the issue of multiple overlapping analyses - when updating meta-analysis - potentially producing misleading results
- Methods starting to emerge to address all the above
  - But, no one has examined how the issues interact with each other
  - So more research is required on a framework which integrates all 3.
    - Although, if all trials are powered adequately to inform meta-analysis, then updating strategies are irrelevant (always update!) and this will reduce the number of updates and reduce the impact of multiplicity.
  - Two “old” issues that need addressing
    - A strategy for acknowledging variable primary study quality for meta-analysis
    - How to proceed when existing studies heterogeneous and variability can’t be explained

References

- Clarke L, Clarke M, Clarke T: How useful are Cochrane reviews in identifying research needs? *Journal of Health Services Research & Policy*. 2007, 12: 101-103.
- Clark M, Brice A, Chalmers I. Accumulating Research: A Systematic Account of How Cumulative Meta-Analyses Would Have Provided Knowledge, Improved Health, Reduced Harm and Saved Resources. *Plos One* 2014
- IntHout J, Ioannidis JPA, Rovers MM, et al. Plea for routinely presenting prediction intervals in meta-analysis. BMJ Open 2016;6:e010247. doi:10.1136/bmjopen-2015-010247
    - After reading this, next read the first paper this paper cites also
- Jones AP, Conroy E, Williamson PR, Clarke M, Gamble C (2013) The use of systematic reviews in the planning, design and conduct of randomised trials: a retrospective cohort of NIHR HTA funded trials. *BMC Medical Research Methodology* 13: 50.
- Langan D, Higgins JPT, Gregory W, Sutton AJ. Graphical augmentations to the funnel plot assess the impact of additional evidence on a meta-analysis. *Journal of Clinical Epidemiology* 2012; 65 (5): 511-9.
- Takwoingi Y, Hopewell S, Tovey D, Sutton AJ. A multicomponent decision tool for prioritising the updating of systematic reviews. *BMJ* 2013;**347**:f7191. doi:10.1136/bmj.f7191 pmid:24336453
- Welton N, Ades AE. (2011) Research decision in the face of heterogeneity: What can a new study tell us? *Health Economics*. DOI: 10.1002/hec.1797

The coherent design of future trials

Sutton, A.J., Cooper, N.J. Jones, D.R. Role of existing systematic reviews and meta-analyses in the design of future research *BMC Medical Research Methodology* 2009; 9:29.

A framework for evaluation of health interventions, aimed at increasing coherence and efficiency through:

i) making better use of information contained within the existing evidence-base when designing future studies

ii) maximising the information available and thus potentially reducing the need for future studies.

iii) Basing design of future trials with respect to impact on the (updated) meta-analysis