Getting Started with Systematic Reviews

W. Scott Richardson, M.D.

AU/UGA Medical Partnership

Three Owl Learning Institute

Conflicts of Interest

- I have no financial ties with industry that pose a conflict of interest regarding the content of this presentation
- I will not be discussing “off label” uses of any medications or devices
- Image copyrights are retained by their original creators, publishers, etc.

Session Aims

- Why bother … ?
- Structure & function
- Critical appraisal
- Forest plots
- Heterogeneity
- Further learning about systematic reviews

EBCP: Here is Why

- **We can't make well-informed decisions without information**
- **Not all information is created equal**
- **Misinformation can be worse than no information**
- **Better information → better informed decisions → better outcomes**

Individual randomized trials of treatment …

- Each trial is one experiment, one new chance to get closer to the ‘truth’
- One trial ~ one race
- Often, more than one trial is done
- Will all trial results agree (even by chance)?

As trials accumulate …

- Seldom is one trial definitive (“One ring to rule them all …”)
- In science, as experiments accrue, knowledge is built cumulatively
- Is there a scientific way to combine results of individual trials?
- Yes! Systematic reviews (we’ll abbreviate “SRs”)

‘Narrative’ vs. ‘Systematic’

- Address disorder as a whole – overview
- Or, tell a ‘story’
- Variety of questions
- No methods section
- No formal pooling
- Thus, may be cumulative but not comprehensive

- Address focused question (e.g. effect of therapy, accuracy of diagnostic test)
- Methods section
- Formal pooling, when appropriate
- Thus, cumulative *and* comprehensive

SR Methods

- Formulate questions
- Define eligibility criteria for study inclusion
- Develop *a priori* hypotheses to explain heterogeneity
- Conduct search
- Screen titles, abstracts for inclusion, exclusion
- Review full text
- Assess the risk of bias
- Abstract data
- When meta-analysis is performed:
  - Summary estimates, confidence intervals
  - Explain heterogeneity
  - Rate confidence in estimates of effect
- Report results
- Update review as needed

‘PRISMA’

- ‘Preferred Reporting Items for Systematic reviews and Meta-Analyses’
- Incorporates evolutionary advances
- Specifies 27 item checklist for reporting, e.g. standardizes figures, etc.
- Since 2009, has replaced ‘QUOROM’, has been adopted by many journals
- *Ann Intern Med 2009; 151: 264 – 269*

Finding SRs

- Cochrane Library
  - CDSR – Cochrane Database of Systematic Reviews
  - DARE – Database of Abstracts of Reviews of Effects
- PubMed
  - Publication types
  - Clinical queries
- Work with your librarians

Critical Appraisal of SRs

Credibility:

- Sensible question?
- Exhaustive search?
- Selection, assessments reproducible?
- Present results ready for application?
- Address confidence in estimates of effect?

Confidence in Estimates:

- Risk of bias?
- Consistent across studies?
- Effect: RR, OR, WMD
- Precision: 95% CI
- Apply to my patient?
- Reporting bias?
- Reasons to increase confidence rating?

‘Risk of bias’

The Cochrane Collaboration’s tool for assessing risk of bias in randomised trials

OPEN ACCESS

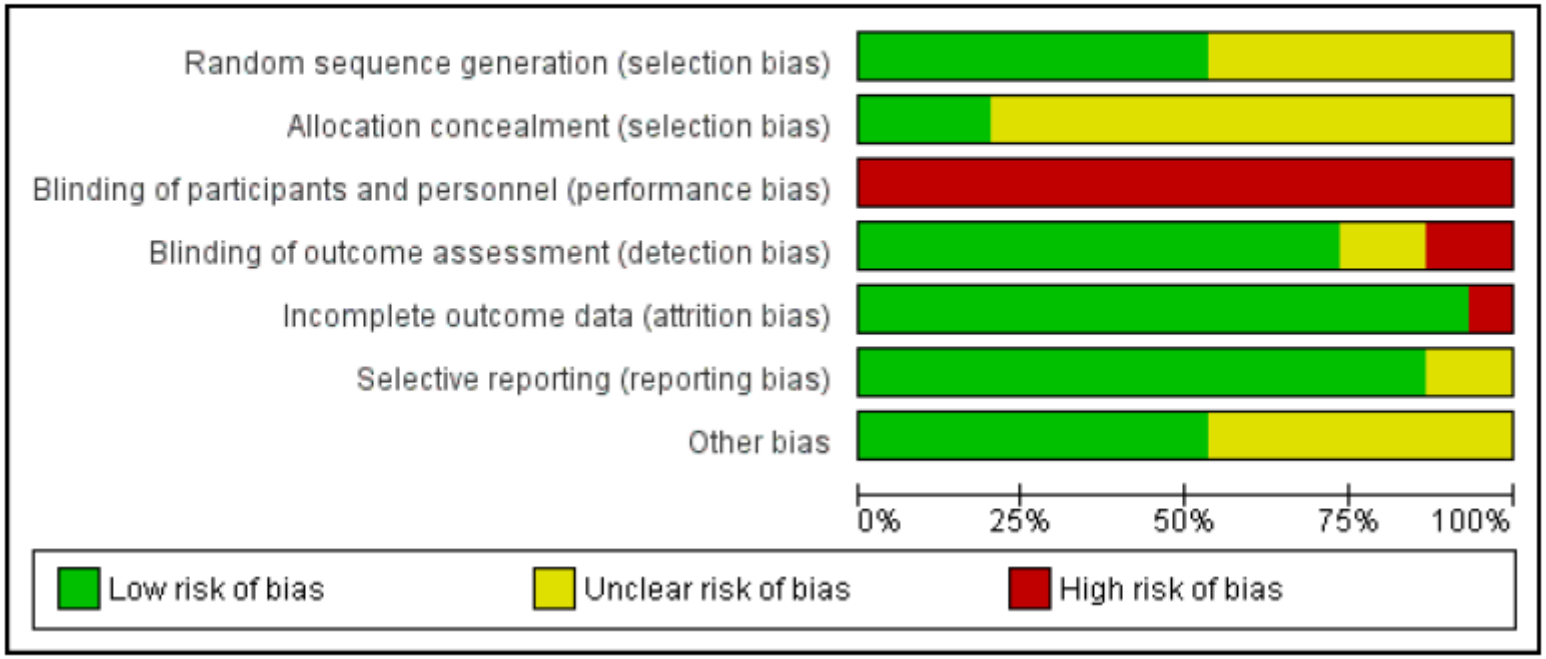
Flaws in the design, conduct, analysis, and reporting of randomised trials can cause the effect of an intervention to be underestimated or overestimated. The Cochrane Collaboration’s tool for assessing risk of bias aims to make the process clearer and more accurate

Julian P T Higgins *senior statistician*[1], Douglas G Altman *director*[2], Peter C Gøtzsche *director*[3], Peter Jüni *head of division*[4], David Moher *senior scientist*[5 6], Andrew D Oxman *senior researcher*[7], Jelena Savović *postdoctoral fellow*[8], Kenneth F Schulz *vice president*[9], Laura Weeks *research associate*[5], Jonathan A C Sterne *professor of medical statistics and epidemiology*[8], Cochrane Bias Methods Group, Cochrane Statistical Methods Group

- Moves away from dichotomous “yes/no” to explicit rating of risk of bias
- At both study-level and outcome-level
- *BMJ 2011; 343: d5928 doi*

Risk of bias graphs

Figure 1. Risk of bias graph: review authors' judgements about each risk of bias item presented as percentages across all included trials.

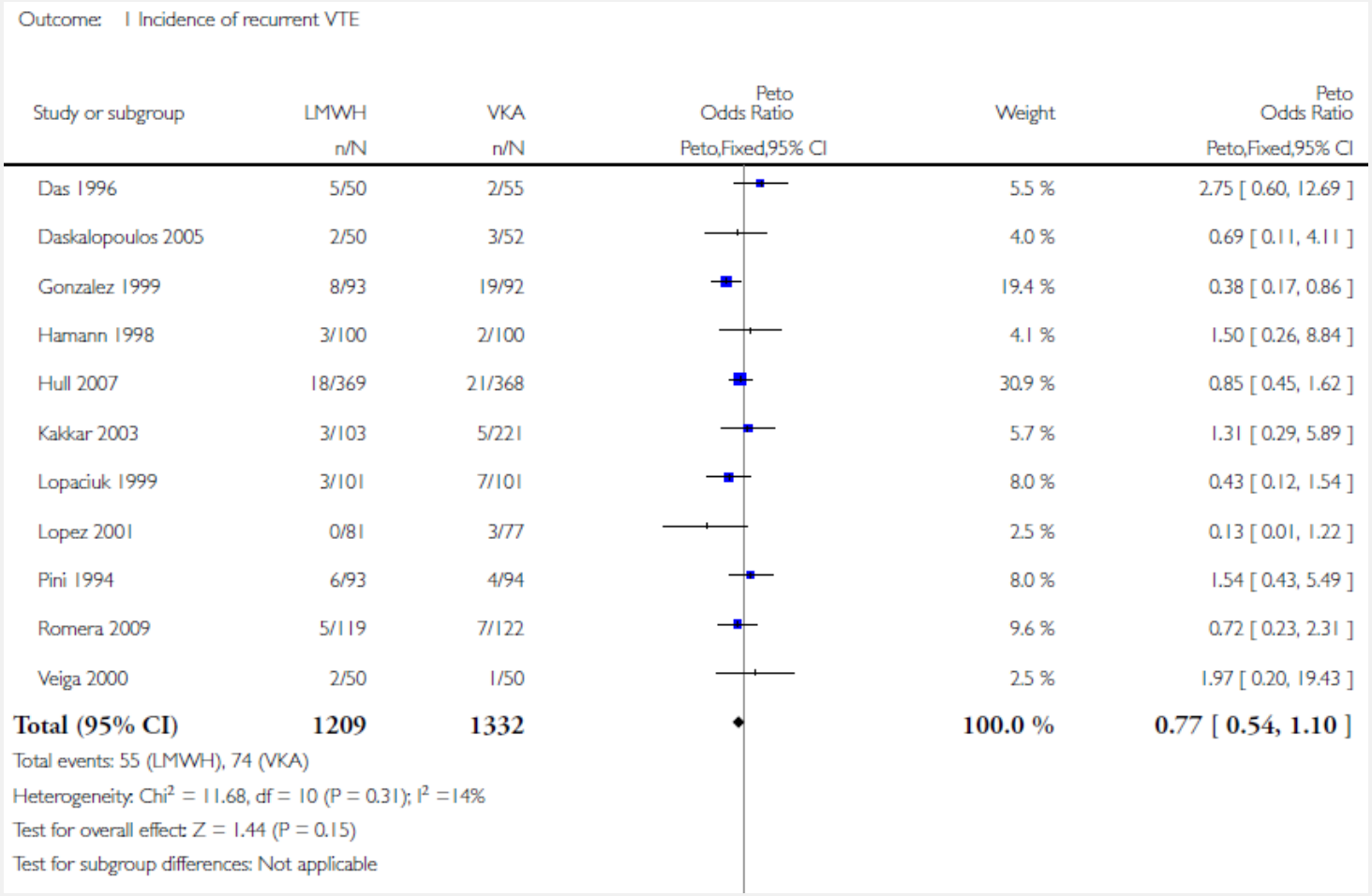


Reporting Biases

- Selective reporting of studies
  - Delayed (or never)
  - Location, language
- Selective reporting of outcomes, times
- Selective reporting of analyses
- *UG 3/e Box 23-2*

- Empirical evidence
- Distort the ‘body of evidence’ in the literature
- Can lead to wrong conclusions about the benefits and harms

Forest Plot – a



Outcome: 1 Incidence of recurrent VTE

Study or subgroup	LMWH n/N	VKA n/N	Peto Odds Ratio Peto,Fixed,95% CI	Weight	Peto Odds Ratio Peto,Fixed,95% CI
Das 1996	5/50	2/55		5.5 %	2.75 [0.60, 12.69]
Daskalopoulos 2005	2/50	3/52		4.0 %	0.69 [0.11, 4.11]
Gonzalez 1999	8/93	19/92		19.4 %	0.38 [0.17, 0.86]
Hamann 1998	3/100	2/100		4.1 %	1.50 [0.26, 8.84]
Hull 2007	18/369	21/368		30.9 %	0.85 [0.45, 1.62]
Kakkar 2003	3/103	5/221		5.7 %	1.31 [0.29, 5.89]
Lopaciuk 1999	3/101	7/101		8.0 %	0.43 [0.12, 1.54]
Lopez 2001	0/81	3/77		2.5 %	0.13 [0.01, 1.22]
Pini 1994	6/93	4/94		8.0 %	1.54 [0.43, 5.49]
Romera 2009	5/119	7/122		9.6 %	0.72 [0.23, 2.31]
Veiga 2000	2/50	1/50		2.5 %	1.97 [0.20, 19.43]
Total (95% CI)	**1209**	**1332**		**100.0 %**	**0.77 [0.54, 1.10]**

Total events: 55 (LMWH), 74 (VKA)

Heterogeneity: $Chi^2 = 11.68$, df = 10 (P = 0.31); $I^2 = 14\%$

Test for overall effect: Z = 1.44 (P = 0.15)

Test for subgroup differences: Not applicable

Forest plot – b

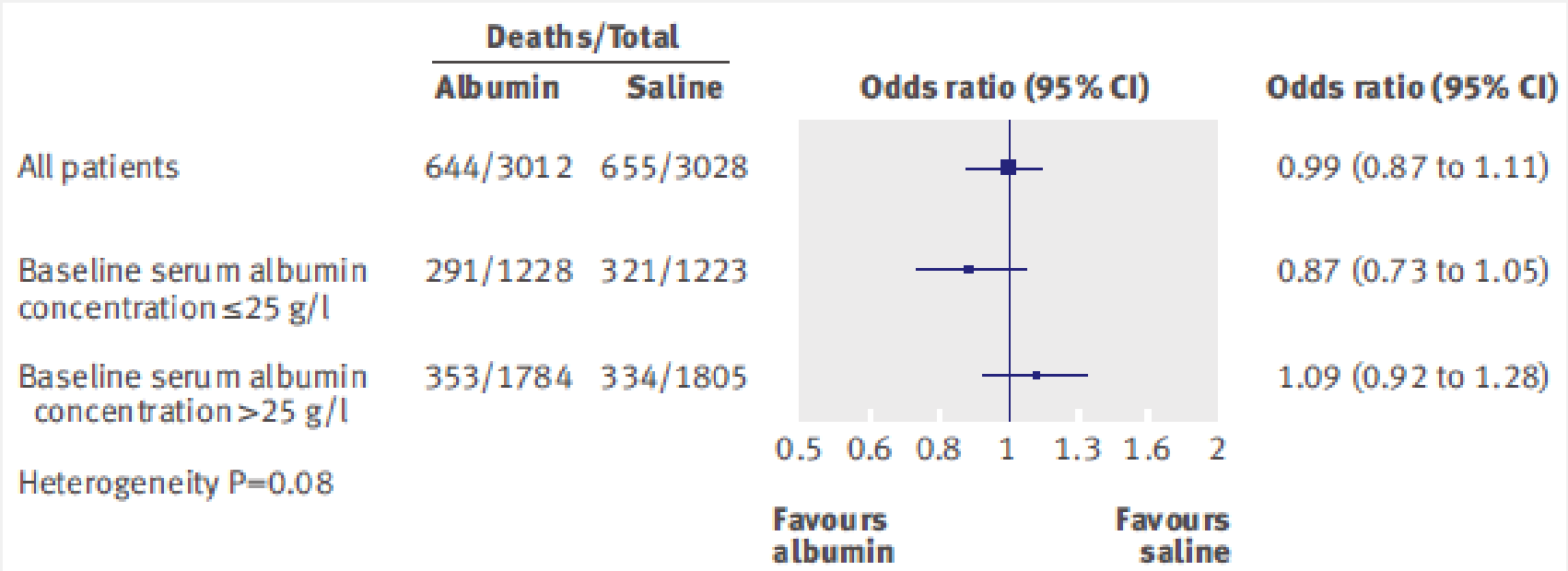
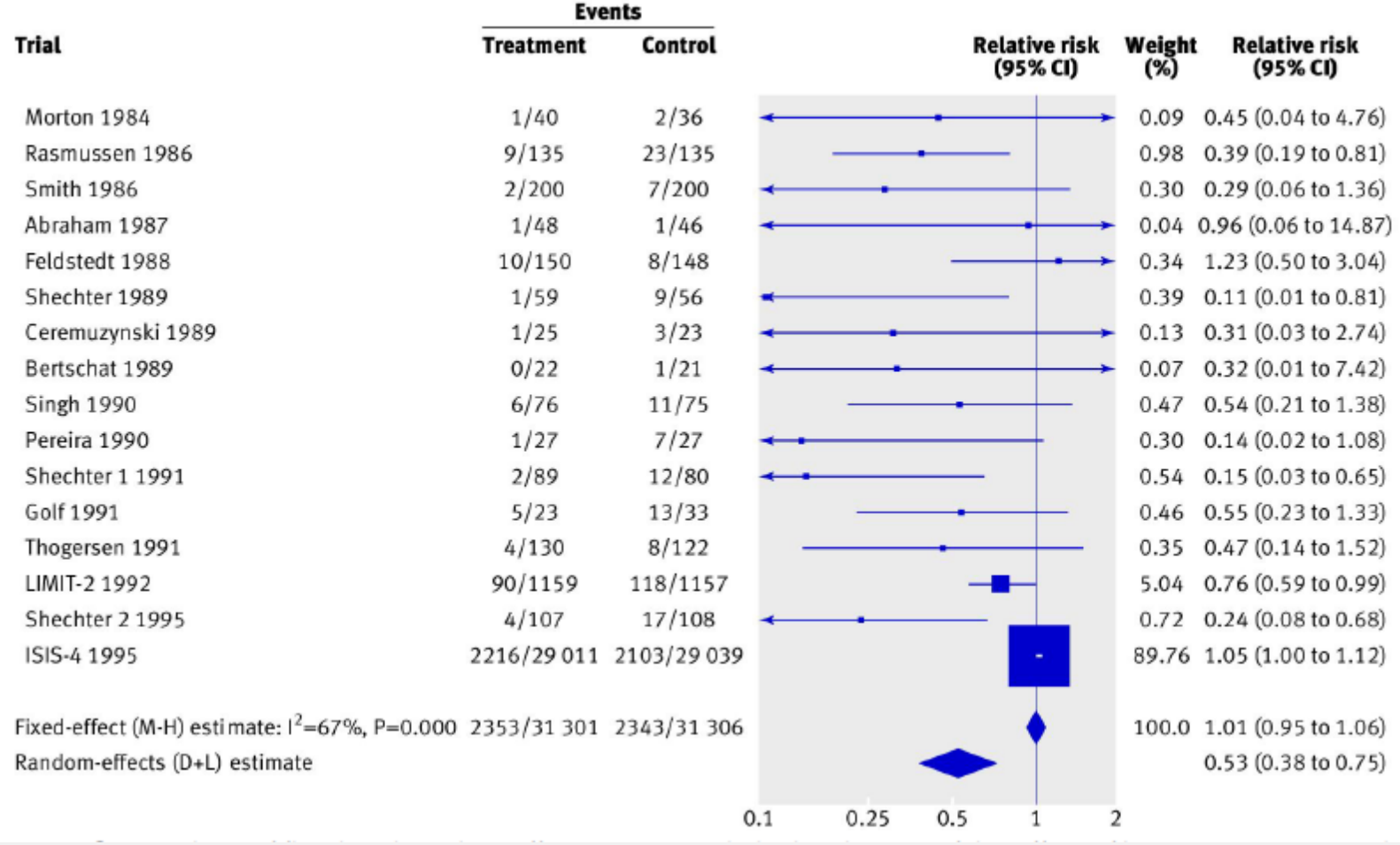
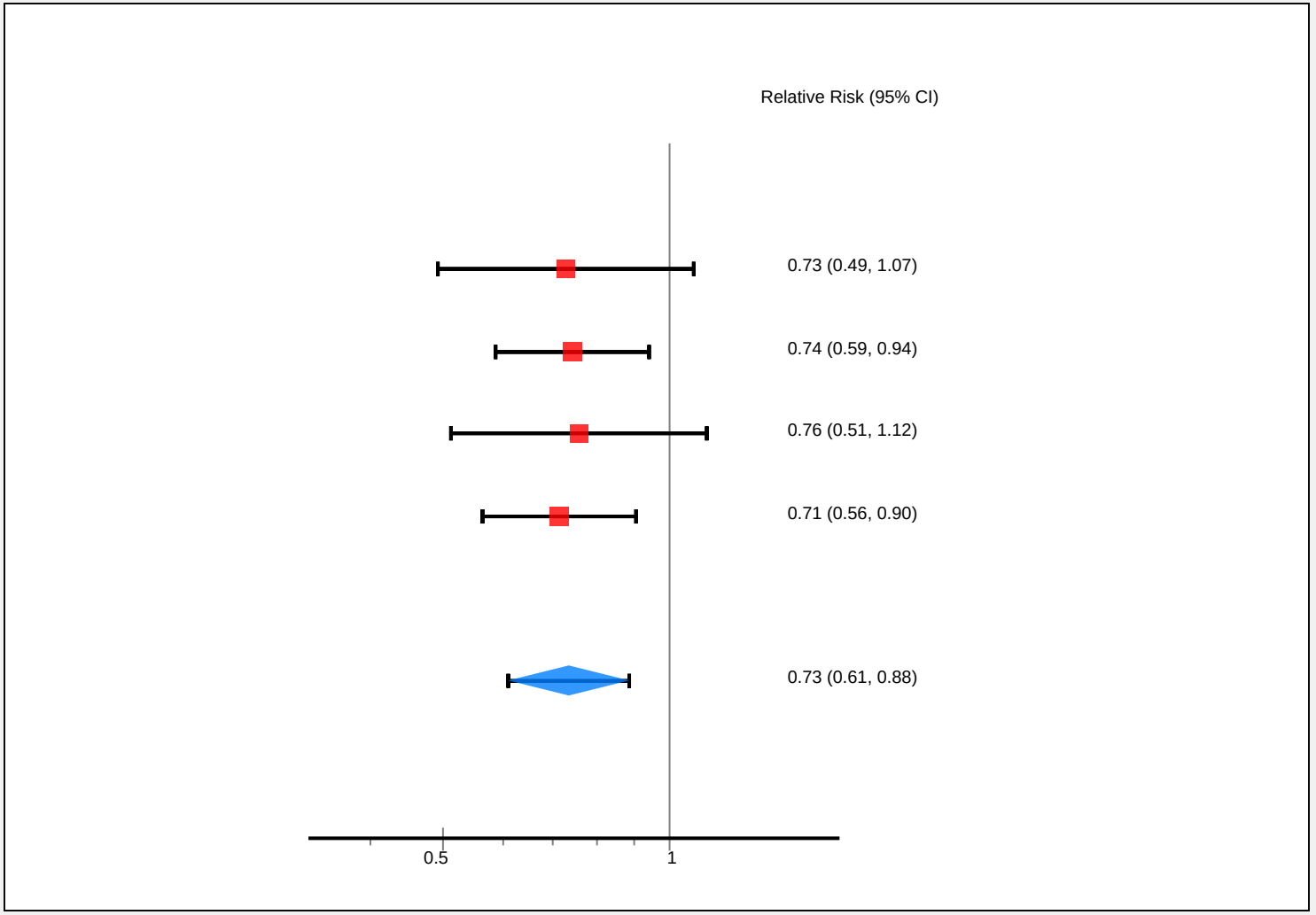


Fig 5 | Forest plot from study comparing resuscitation with albumin or saline in intensive care showing unadjusted odds ratio of death stratified by baseline albumin concentration[18]

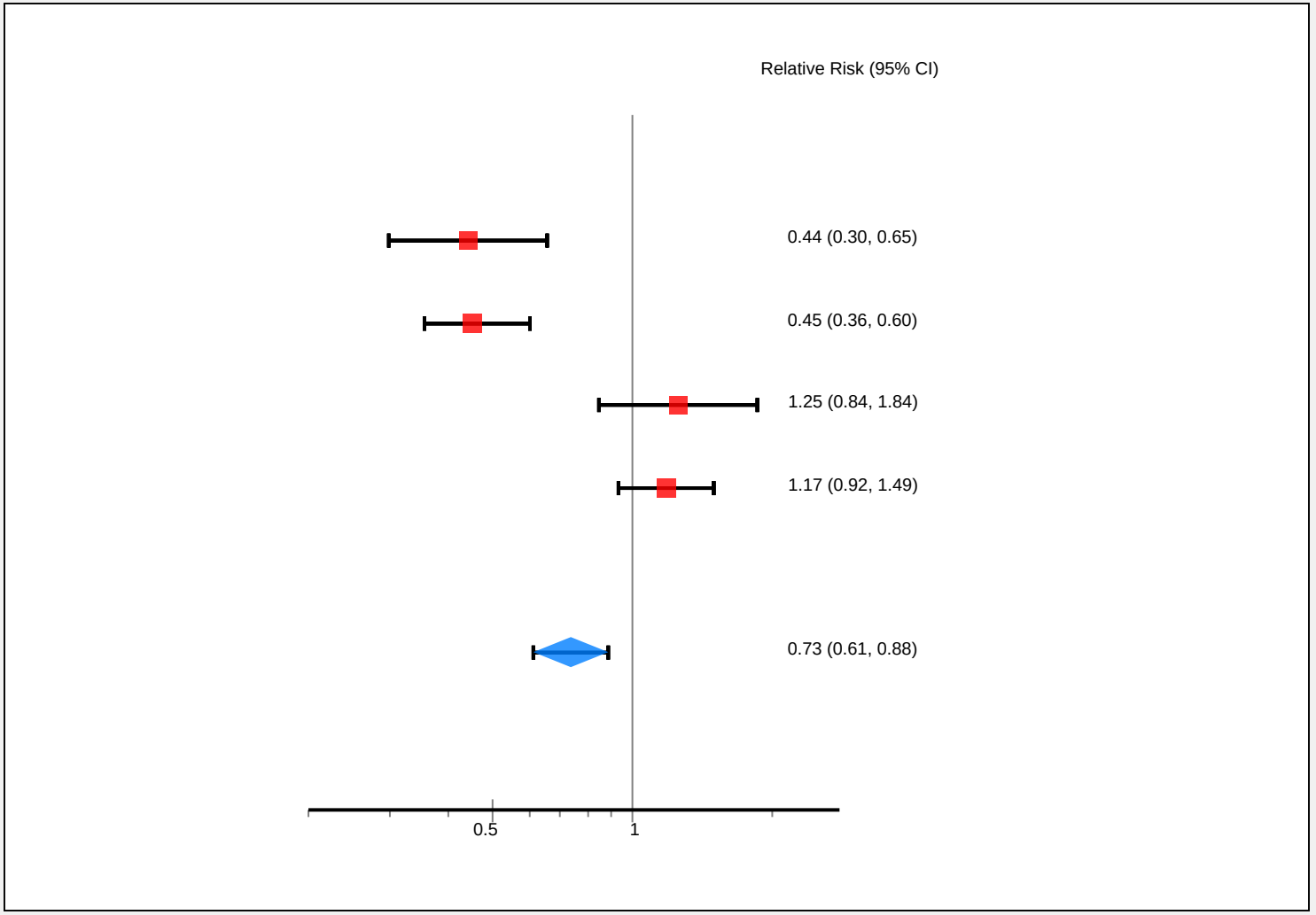
Forest plot – c



Are you happy pooling?



Are you happy pooling?



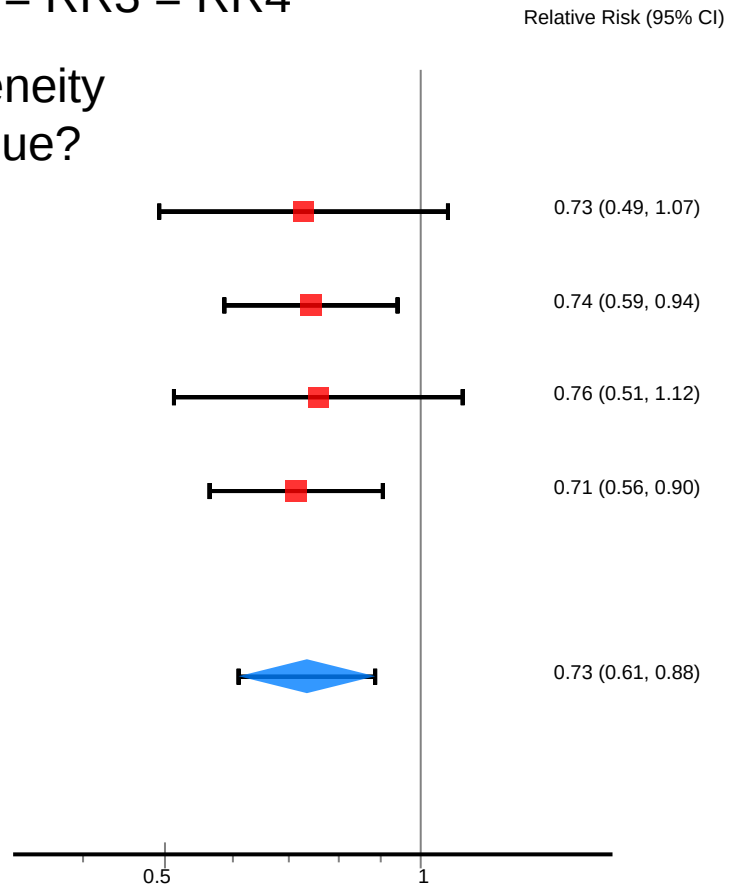
What criteria were you using?

- similarity of point estimates
  - less similar, less happy
- overlap of confidence intervals
  - less overlap, less happy

Homogenous

Ho: RR1 = RR2 = RR3 = RR4

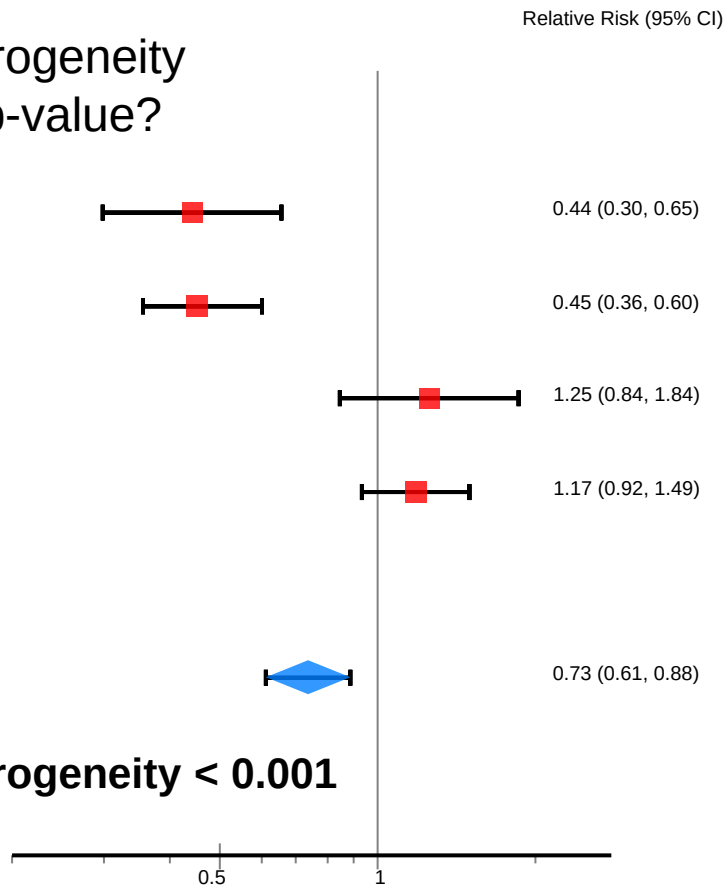
test for heterogeneity
what is the p-value?



p=0.99 for heterogeneity

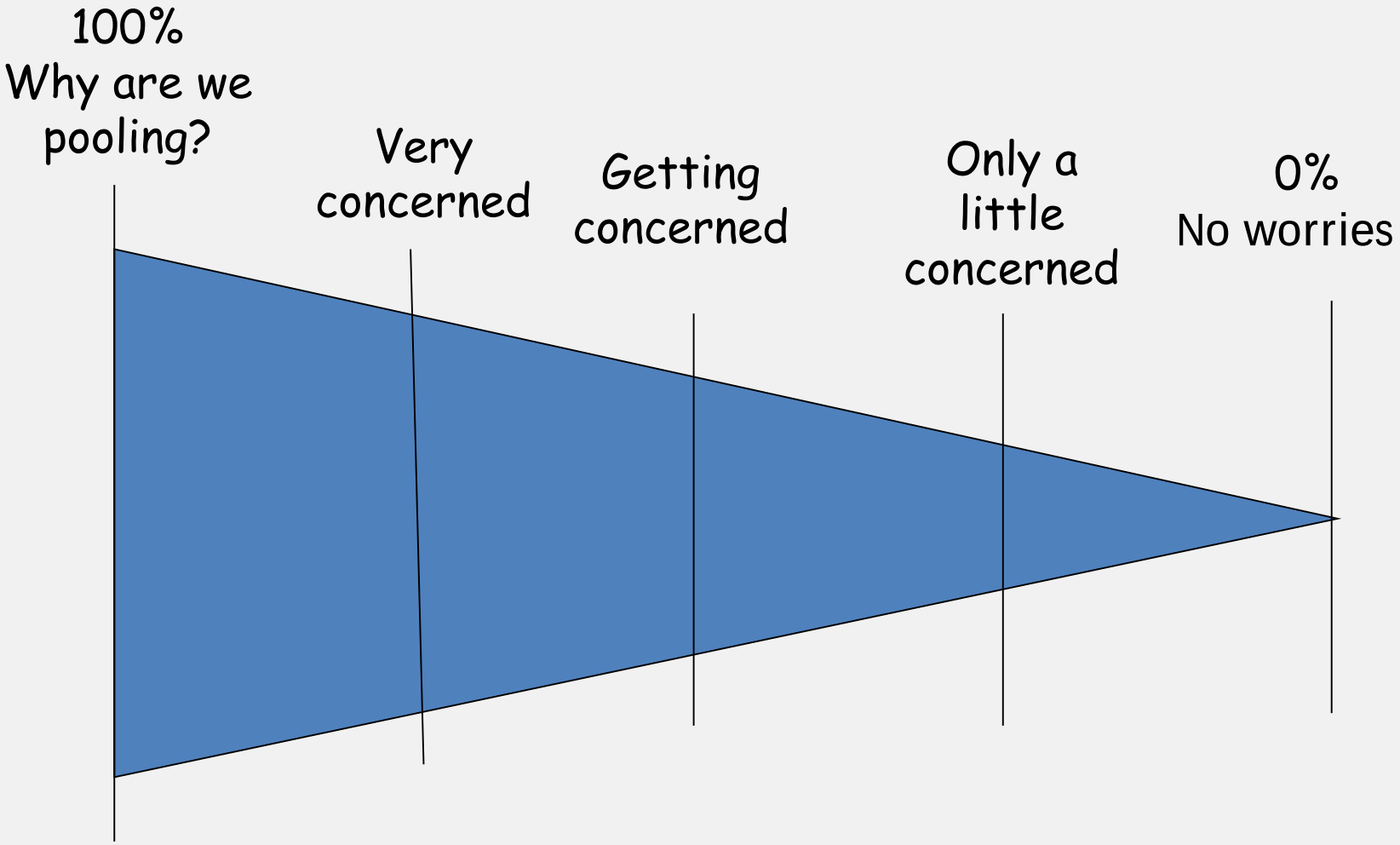
Heterogeneous

test for heterogeneity
what is the p-value?



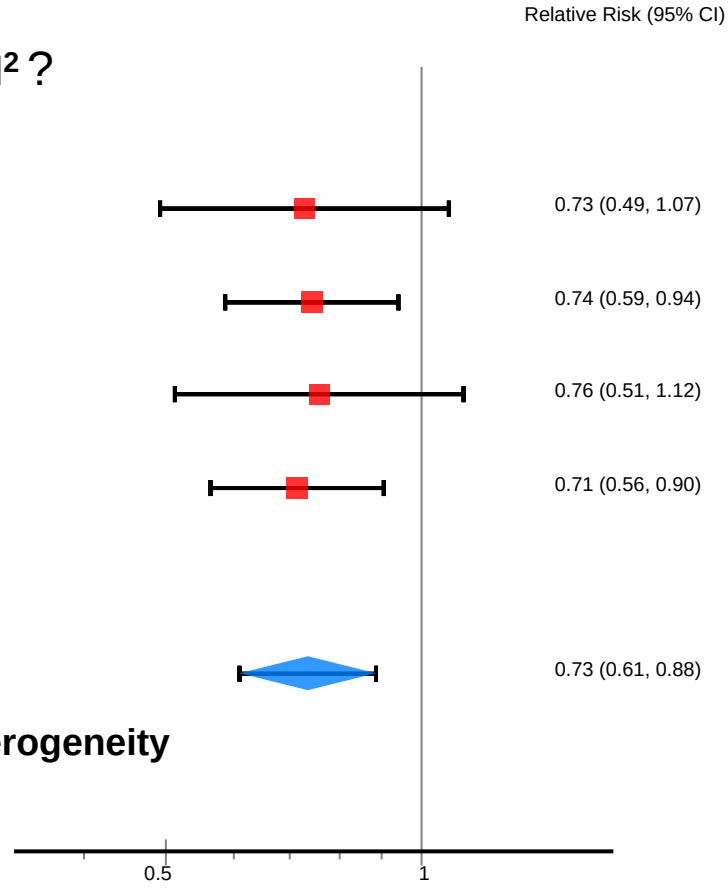
p-value for heterogeneity < 0.001

I^2 Interpretation



Homogenous

What is the I^2?

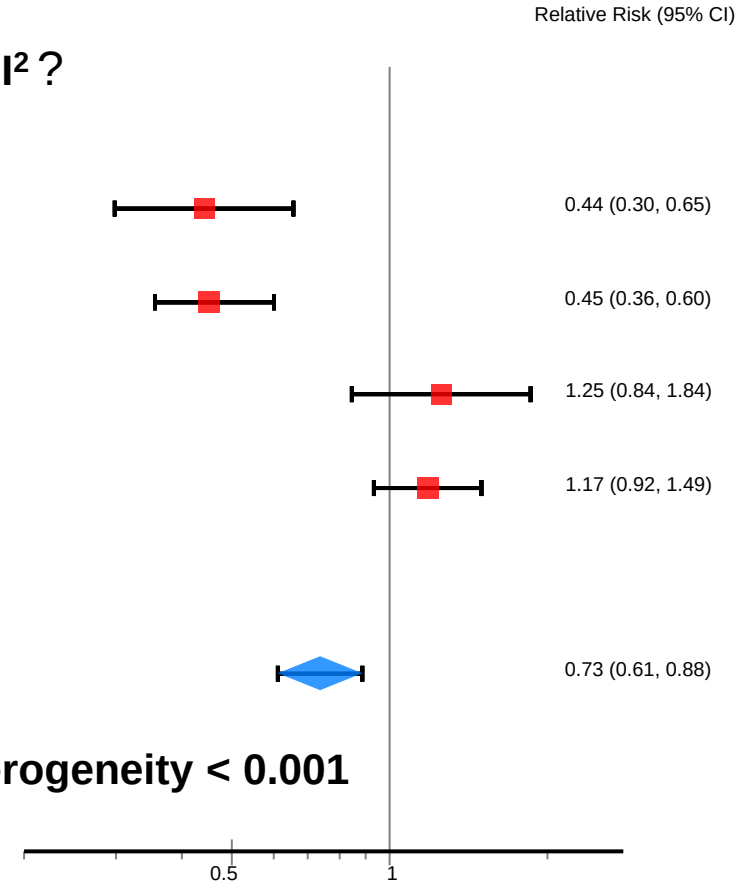


p=0.99 for heterogeneity

I^2=0%

Heterogeneous

What is the I^2?



p-value for heterogeneity < 0.001

I^2=89%

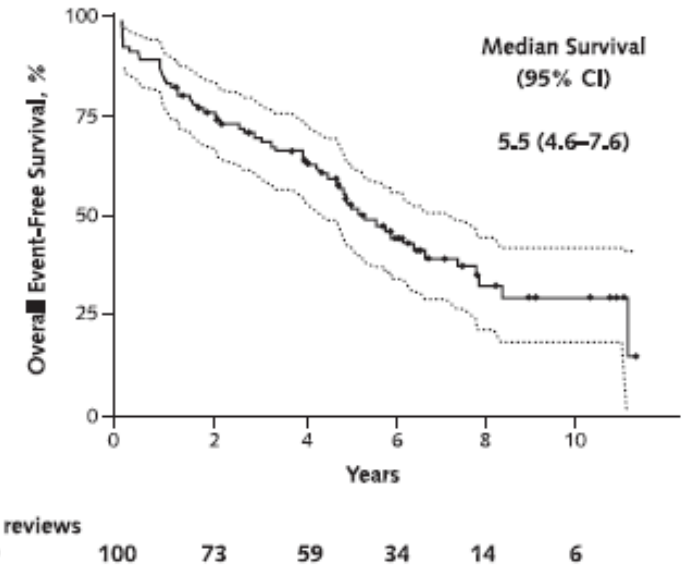
Heterogeneity

- Humans vary, e.g. in risk of poor outcomes from disease, in response to therapy, and in vulnerability to adverse effects
- Heterogeneity represents this variation in results
- Affects certainty about estimates of effect

- Identified by:
  - Visual inspection
  - Chi^2: “yes” or “no”
  - I^2: 0 to 100%
- Explored by:
  - Patients
  - Interventions
  - Comparisons
  - Outcomes
  - Methods, Systems, +

How quickly do systematic reviews go out of date?

Figure 2. **Overall survival time (95% CI) free of signals for updating.**



The immediate decrease in survival at time zero reflects the 7 systematic reviews for which signals for updating had already occurred at the time of publication. The low number of reviews at risk after 10 years reflects the fact that the sample spanned 1995 to 2005 and censoring occurred on 1 September 2006. Thus, only reviews published before September 1996 and having no signals for updating could have more than 10 years of observation.

- Survival analysis
- 100 systematic reviews, 1995 – 2005
- Searched for ‘update signals’ (i.e. new trial evidence)

- *Ann Intern Med 2007*

Learning more about SRs

Next steps

- In your groups, work through SR package
- Build skills in areas:
  - Finding SRs
  - Appraising SRs critically
  - Interpreting results
  - Applying results
  - Explaining to others

Further steps

- Subgroup analyses
- Cumulative meta-analysis
- Individual patient-data meta-analysis
- Network meta-analysis
- Translating SR evidence into action

Taking SRs home …

- When well-made and current, SRs synthesize the body of research evidence that can guide important decisions
- SRs have limits, yet we should start with them: ‘how well does this work?’
- We can (and must!) appraise SRs for risk of bias, estimates of effect, and confidence in these estimates

Thank You