This session is going to be interactive. You will be breaking into groups to solve problems. So please sit with people with whom you will enjoy a conversation

Plan

- GRADE background
- two steps
  - quality of evidence
  - strength of recommendation
- evidence profiles
- an exercise in applying GRADE

Why this talk?

- you and your students are likely to be seeing a lot of it
- exemplifies three key principles of EBHC
  - need for systematic reviews of best evidence
  - hierarchy of evidence
  - need for values and preferences
- if you understand GRADE you understand how to use evidence to inform practice

Why Grade Recommendations?

- strong recommendations
  - strong methods
  - large precise effect
  - few down sides of therapy
- weak recommendations
  - weak methods
  - imprecise estimate
  - small effect
  - substantial down sides

Grading good idea, but which grading system to use?

- many available
  - Australian National and MRC
  - Oxford Center for Evidence-based Medicine
  - Scottish Intercollegiate Guidelines (SIGN)
  - US Preventative Services Task Force
  - American professional organizations
    - AHA/ACC, ACCP, AAP, Endocrine society, etc....
- cause of confusion, dismay

Common international grading system?

- GRADE (*G*rades of *r*ecommendation, *a*ssessment, *d*evelopment and *e*valuation)
- international group
  - Australian NMRC, SIGN, USPSTF, WHO, NICE, Oxford CEBM, CDC, CC
- ~ 30 meetings over last 13 years
  - (~10 – 50 attendants)

GRADE GUIDANCE

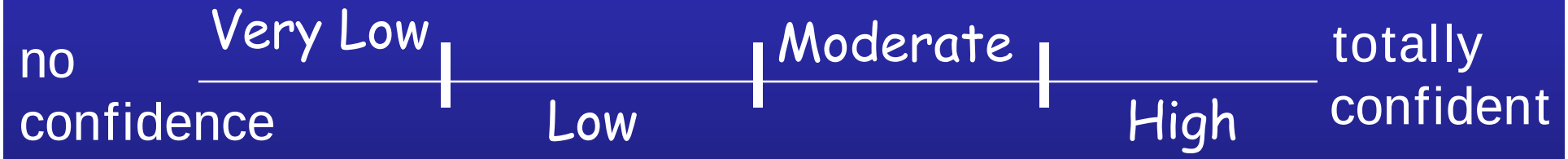
- 2004 BMJ, first description
- 2008 BMJ six part series
  - for guideline users
- 2010-13, 21 part series, 15 published
  - for systematic review authors, HTA practitioners, guideline developers

70+ Organizations

GRADE uptake

What are we grading?

two components



strength of recommendation:
strong and weak

Determinants of quality

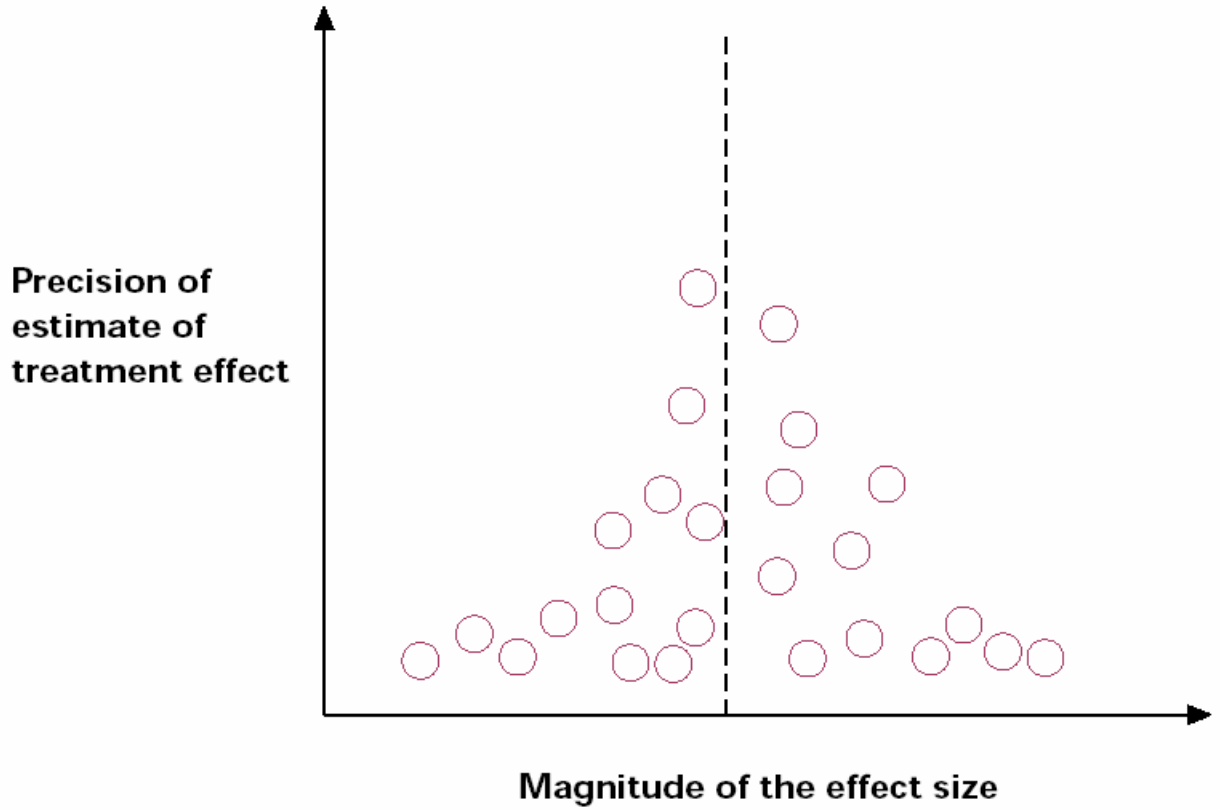
- RCTs start high
- observational studies start low
- what can lower quality?

- risk of bias
- inconsistency
- indirectness
- imprecision
- publication bias

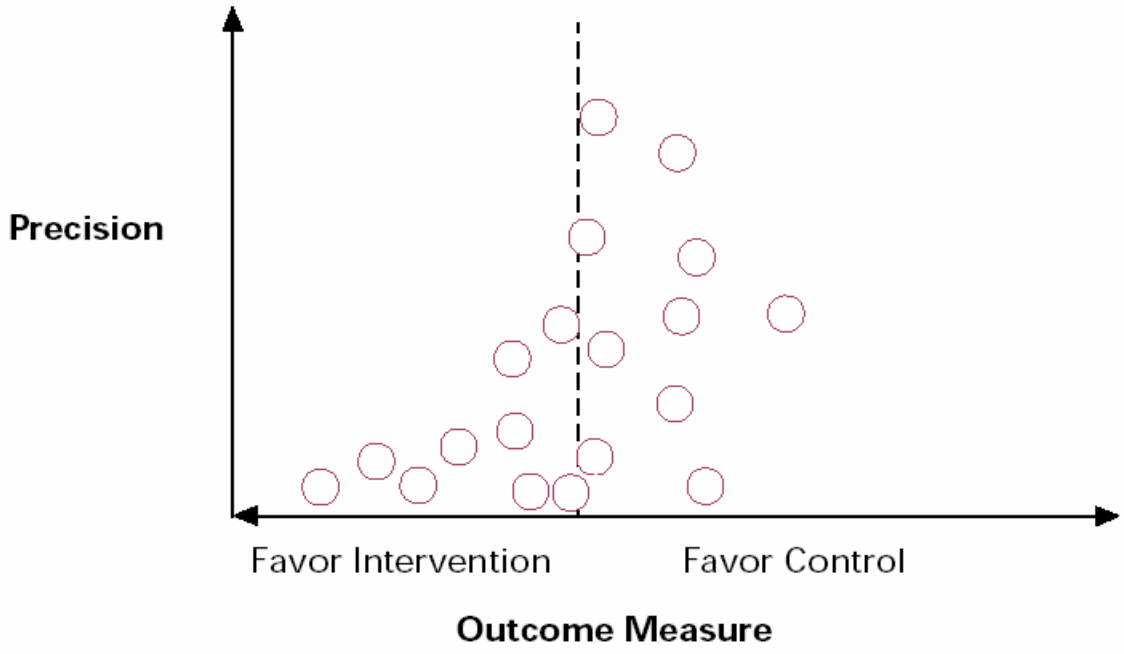
Determinants of quality

- imprecision
  - wide confidence intervals
- risk of bias
  - concealment, blinding, loss to follow-up
- publication bias

Funnel Plot


Precision of estimate of treatment effect
Magnitude of the effect size

Publication Bias



Consistency of results

- if inconsistency, look for explanation
  - patients, intervention, outcome, methods
- judgment of consistency
  - variation in size of effect
  - overlap in confidence intervals
  - statistical significance of heterogeneity
  - I^2

Relative Risk with 95% CI for Vitamin D Non-vertebral Fractures

Chapuy et al, (1994) 0.79 (0.69, 0.92)

Lips et al, (1996) 1.10 (0.87, 1.39)

Dawson-Hughes et al, (1997) 0.46 (0.24, 0.88)

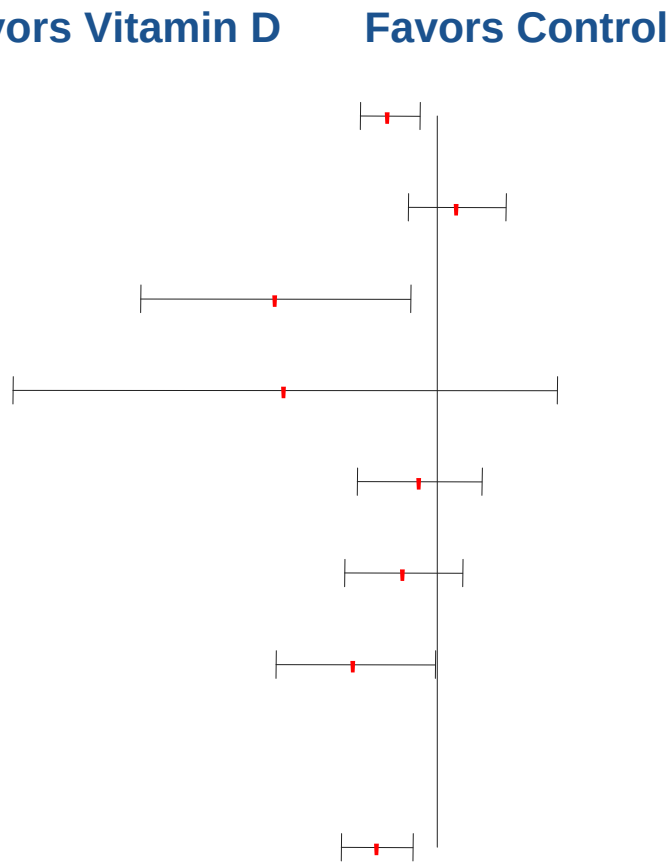
Pfeifer et al, (2000) 0.48 (0.13, 1.78)

Meyer et al, (2002) 0.92 (0.68, 1.24)

Chapuy et al, (2002) 0.85 (0.64, 1.13)

Trivedi et al, (2003) 0.67 (0.46, 0.99)

Pooled Random Effect Model
0.82 (0.69 to 0.98)
p= 0.05 for heterogeneity, I^2=53%

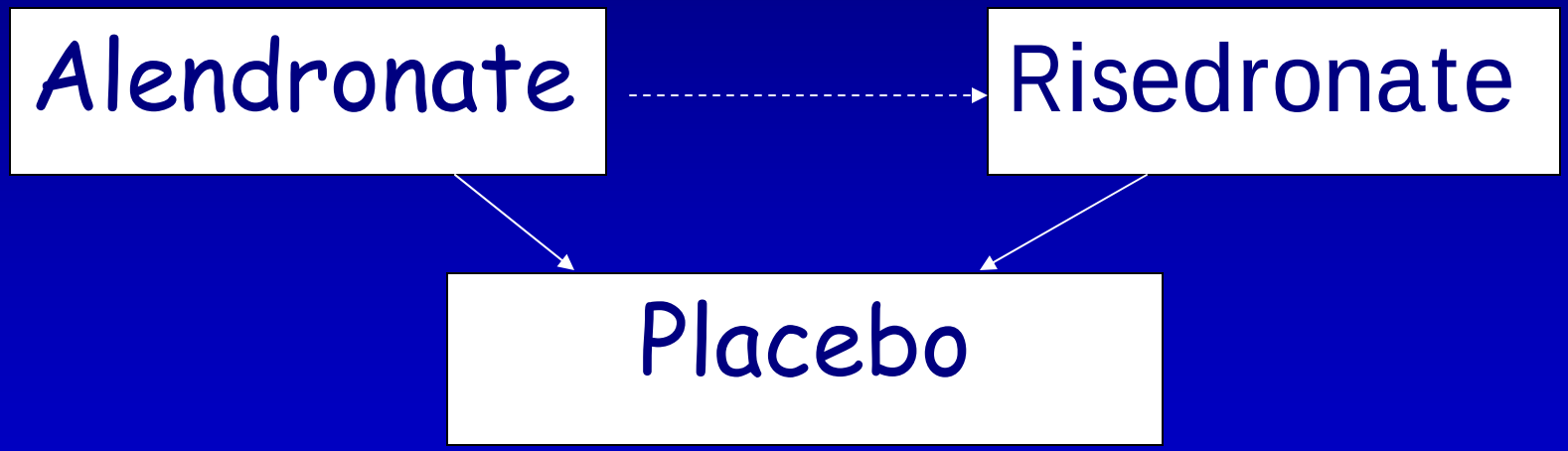


Quality judgments: Directness

- populations
  - older, sicker or more co-morbidity
- interventions
  - new statins versus old
- outcomes
  - important versus surrogate outcomes
  - glucose control versus CV events

Directness

interested in A versus B
available data A vs C, B vs C



What can raise quality?

- large magnitude can rate up one level
  - very large two levels
- common criteria
  - everyone used to do badly
  - almost everyone does well
  - quick action
- hip replacement for hip osteoarthritis
- mechanical ventilation in respiratory failure

Confidence assessment criteria

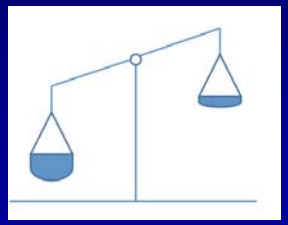
Study Design	Confidence in estimates	Lower if	Higher if
Randomised trial →	High	Risk of bias - 1 Serious -2 Very serious Inconsistency -1 Serious -2 Very serious Indirectness -1 Serious -2 Very serious Imprecision -1 Serious -2 Very serious Publication bias -1 Likely -2 Very likely	Large effect +1 Large +2 Very large Dose response +1 Evidence of a gradient All plausible confounding +1 Would reduce a demonstrated effect or +1 Would suggest a spurious effect when results show no effect
	Moderate		
Observational study →	Low		
	Very low		

Beta blockers in non-cardiac surgery

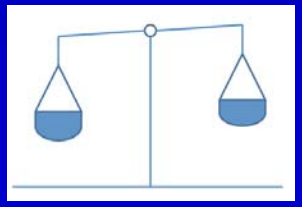
Quality Assessment							Summary of Findings		
Outcome	Number of participants (studies)	Risk of Bias	Consistency	Directness	Precision	Publication Bias	Quality	Relative Effect (95% CI)	Absolute risk difference
Myocardial infarction	10,125 (9)	No serious limitations	No serious imitations	No serious limitations	No serious limitations	Not detected	High	0.71 (0.57 to 0.86)	1.5% fewer (0.7% fewer to 2.1% fewer)
Mortality	10,205 (7)	No serious limitations	No serious limiations	No serious limitations	Imprecise	Not detected	Moderate	1.23 (0.98 – 1.55)	0.5% more (0.1% fewer to 1.3% more)
Stroke	10,889 (5)	No serious limitaions	No serious limitations	No serious limitations	No serious limitations	Not detected	High	2.21 (1.37 – 3.55)	0.5% more (0.2% more to 1.3% more0

Strength of Recommendation

- strong recommendation
  - benefits clearly outweigh risks/hassle/cost
  - risk/hassle/cost clearly outweighs benefit



- what can downgrade strength?
- low confidence in estimates
- close balance between up and downsides



Flavanoids for Hemorrhoids

- venotonic agents
  - mechanism unclear, increase venous return
- popularity
  - 90 venotonics commercialized in France
  - none in Sweden and Norway
  - France 70% of world market
- possibilities
  - French misguided
  - rest of world missing out

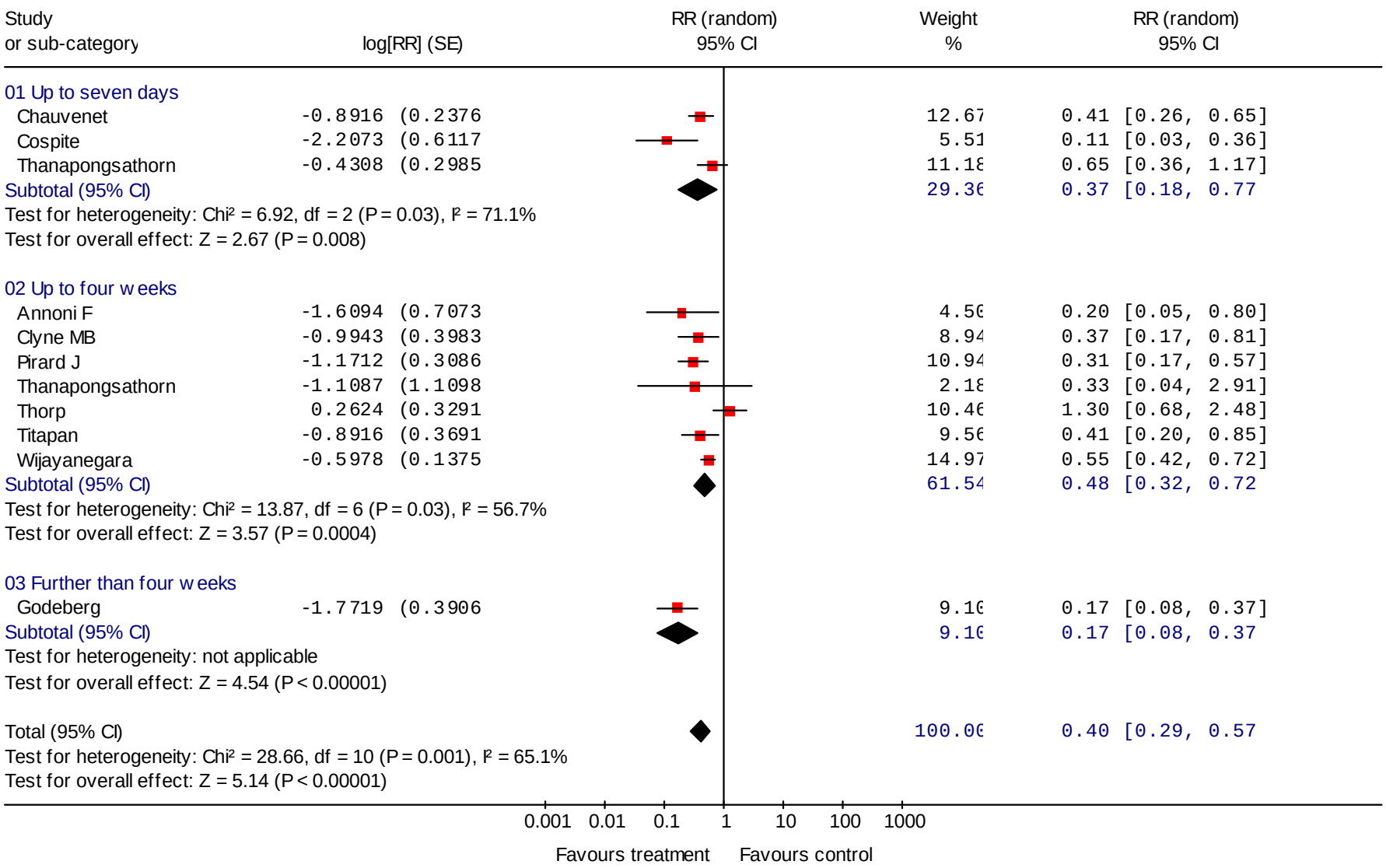
Systematic Review

- 14 trials, 1432 patients
- key outcome
  - risk not improving/persistent symptoms
  - 11 studies, 1002 patients, 375 events
  - RR 0.4, 95% CI 0.29 to 0.57
- minimal side effects
- is France right?
- what is the quality of evidence?

What can lower quality?

- risk of bias
  - lack of detail re concealment
  - questionnaires not validated
- indirectness – no problem
- inconsistency, need to look at the results

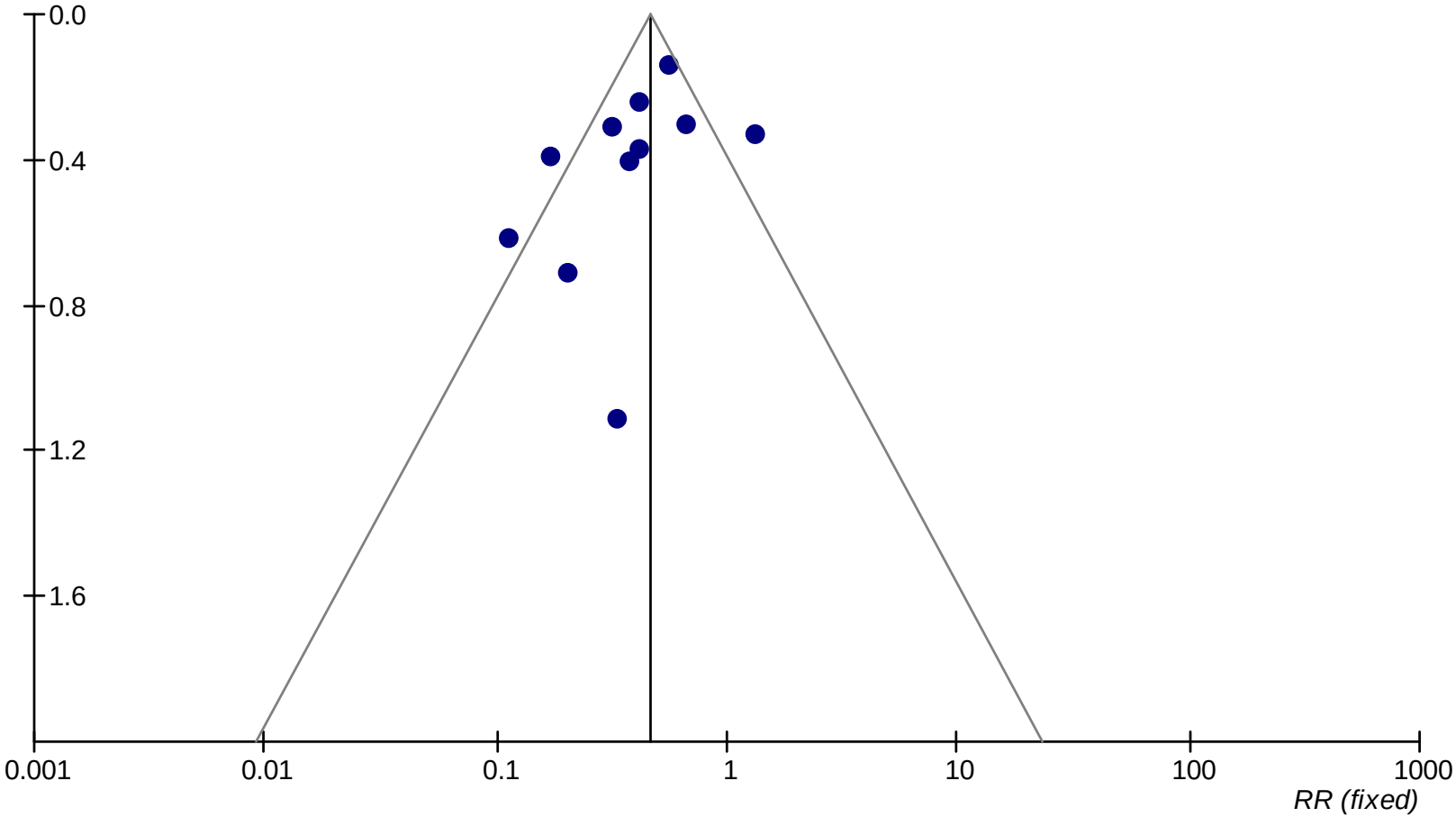
Review: Phlebotonics for hemorrhoids
Comparison: 01 Venotonics vs placebp
Outcome: 08 Overall improvement: no improvement/some improvement



Publication bias?

- size of studies
  - 40 to 234 patients, most around 100
- all industry sponsored

Review: Phlebotonics for hemorrhoids
Comparison: 01 Venotonics vs placebp
Outcome: 08 Overall improvement: no improvement/some improvement



What can lower quality?

- detailed design and execution
  - lack of detail re concealment
  - questionnaires not validated
- inconsistency
  - almost all show positive effect, trend
  - heterogeneity $p < 0.001$; I^2 65.1%
- indirectness
- imprecision
  - RR 0.4, 95% CI 0.29 to 0.57
- publication bias
  - 40 to 234 patients, most around 100

Is France right?

- recommendation
  - yes
  - no against use
- strength
  - strong
  - weak

Conclusion

- clinicians, policy makers need summaries
  - quality of evidence
  - strength of recommendations
- explicit rules
  - transparent, informative
- GRADE
  - simple, transparent, systematic
  - increasing wide adoption
  - great opportunity for teaching EBHC