LSTM
LIVERPOOL SCHOOL
OF TROPICAL MEDICINE

Understanding GRADE

Evidence-informed guidelines

Cochrane
Infectious Diseases

READ-It

Meta-analysis and Forest Plots

- A way of combining results from a number of individual trials to produce a summary result
- A forest plot displays the summary result of a meta-analysis and the results of the individual studies

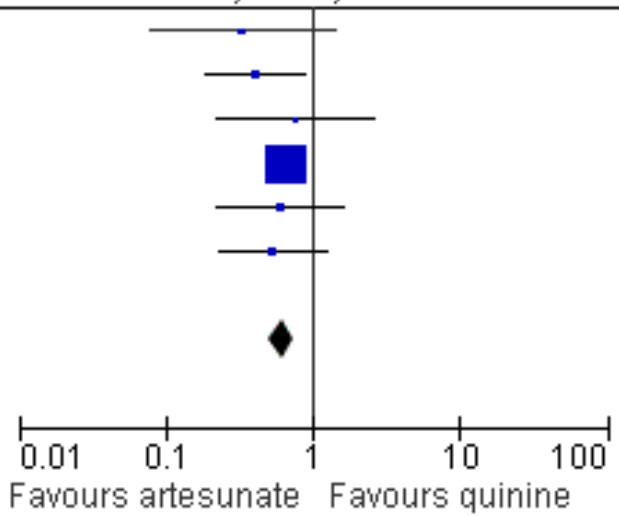
Artesunate versus quinine in severe malaria

Outcome: Death

Study or Subgroup	Artesunate Events	Artesunate Total	Quinine Events	Quinine Total	Weight	Risk Ratio M-H, Fixed, 95% CI
Anh 1989	2	19	7	22	3.0%	0.33 [0.08, 1.41]
Anh 1995	8	99	18	91	8.7%	0.41 [0.19, 0.89]
Cao 1997	4	37	5	35	2.4%	0.76 [0.22, 2.59]
Dondorp 2005	107	730	164	731	76.3%	0.65 [0.52, 0.81]
Hien 1992	5	31	8	30	3.8%	0.60 [0.22, 1.64]
Newton 2003	7	59	12	54	5.8%	0.53 [0.23, 1.26]
Total (95% CI)		**975**		**963**	**100.0%**	**0.62 [0.51, 0.75]**
Total events	133		214			

Heterogeneity: $Chi^2 = 2.26$, $df = 5$ ($P = 0.81$); $I^2 = 0\%$

Test for overall effect: $Z = 4.82$ ($P < 0.00001$)



Cochrane Infectious Diseases

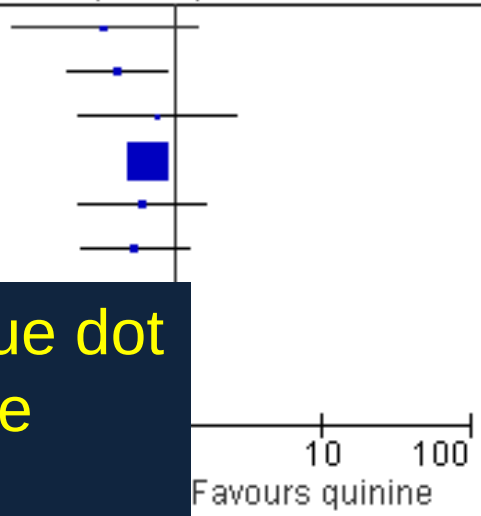
READ-It

Artesunate versus quinine in severe malaria
Outcome: Death

Study or Subgroup	Artesunate Events	Artesunate Total	Quinine Events	Quinine Total	Weight	Risk Ratio M-H, Fixed, 95% CI
Anh 1989	2	19	7	22	3.0%	0.33 [0.08, 1.41]
Anh 1995	8	99	18	91	8.7%	0.41 [0.19, 0.89]
Cao 1997	4	37	5	35	2.4%	0.76 [0.22, 2.59]
Dondorp 2005	107	730	164	731	76.3%	0.65 [0.52, 0.81]
Hien 1992	5	31	8	30	3.8%	0.60 [0.22, 1.64]
Newton 2003	7	59	12	54	5.8%	0.53 [0.23, 1.26]
Total (95% CI)		**975**		**963**	10	
Total events	133		214			

Heterogeneity: Chi² = 2.26, df = 5 (P = 0.81); I² = 0%

Test for overall effect: Z = 4.82 (P < 0.00001)



The size of the blue dot is the weight of the trial.

Weight takes into account the size of the trial, and the number of events (ie deaths)

Cochrane Infectious Diseases

READ-It

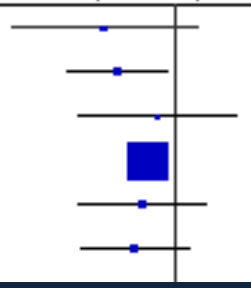
Artesunate versus quinine in severe malaria

Outcome: Death

Study or Subgroup	Artesunate Events	Artesunate Total	Quinine Events	Quinine Total	Weight	Risk Ratio M-H, Fixed, 95% CI
Anh 1989	2	19	7	22	3.0%	0.33 [0.08, 1.41]
Anh 1995	8	99	18	91	8.7%	0.41 [0.19, 0.89]
Cao 1997	4	37	5	35	2.4%	0.76 [0.22, 2.59]
Dondorp 2005	107	730	164	731	76.3%	0.65 [0.52, 0.81]
Hien 1992	5	31	8	30	3.8%	0.60 [0.22, 1.64]
Newton 2003	7	59	12	54	5.8%	0.53 [0.23, 1.26]
Total (95% CI)		**975**		**963**	**100.0%**	**0.62 [0.51,**
Total events	133		214			

Heterogeneity: $Chi^2 = 2.26$, $df = 5$ ($P = 0.81$); $I^2 = 0\%$

Test for overall effect: $Z = 4.82$ ($P < 0.00001$)



The 'whiskers' represent the 95% confidence interval

Cochrane Infectious Diseases

READ-It

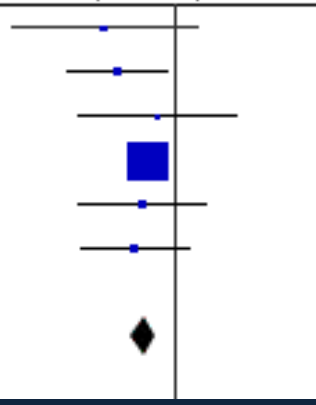
Artesunate versus quinine in severe malaria

Outcome: Death

Study or Subgroup	Artesunate Events	Artesunate Total	Quinine Events	Quinine Total	Weight	Risk Ratio M-H, Fixed, 95% CI
Anh 1989	2	19	7	22	3.0%	0.33 [0.08, 1.41]
Anh 1995	8	99	18	91	8.7%	0.41 [0.19, 0.89]
Cao 1997	4	37	5	35	2.4%	0.76 [0.22, 2.59]
Dondorp 2005	107	730	164	731	76.3%	0.65 [0.52, 0.81]
Hien 1992	5	31	8	30	3.8%	0.60 [0.22, 1.64]
Newton 2003	7	59	12	54	5.8%	0.53 [0.23, 1.26]
Total (95% CI)		**975**		**963**	**100.0%**	**0.62 [0.51, 0.75]**
Total events	133		214			

Heterogeneity: $Chi^2 = 2.26$, $df = 5$ ($P = 0.81$); $I^2 = 0\%$

Test for overall effect: $Z = 4.82$ ($P < 0.00001$)



The ‘diamond’ represents the point estimate and confidence intervals when you combine studies

Cochrane Infectious Diseases

READ-It

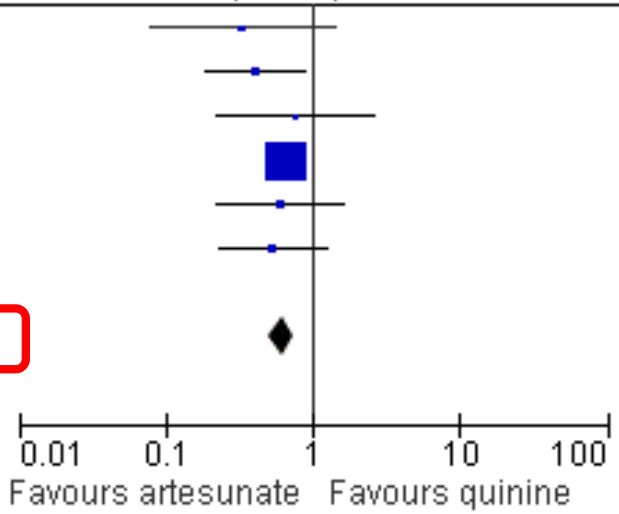
Artesunate versus quinine in severe malaria

Outcome: Death

Study or Subgroup	Artesunate Events	Artesunate Total	Quinine Events	Quinine Total	Weight	Risk Ratio M-H, Fixed, 95% CI
Anh 1989	2	19	7	22	3.0%	0.33 [0.08, 1.41]
Anh 1995	8	99	18	91	8.7%	0.41 [0.19, 0.89]
Cao 1997	4	37	5	35	2.4%	0.76 [0.22, 2.59]
Dondorp 2005	107	730	164	731	76.3%	0.65 [0.52, 0.81]
Hien 1992	5	31	8	30	3.8%	0.60 [0.22, 1.64]
Newton 2003	7	59	12	54	5.8%	0.53 [0.23, 1.26]
Total (95% CI)		**975**		**963**	**100.0%**	**0.62 [0.51, 0.75]**
Total events	133		214			

Heterogeneity: Chi² = 2.26, df = 5 (P = 0.81); I² = 0%

Test for overall effect: Z = 4.82 (P < 0.00001)



The RR in meta-analysis is 0.62. How do you interpret this?

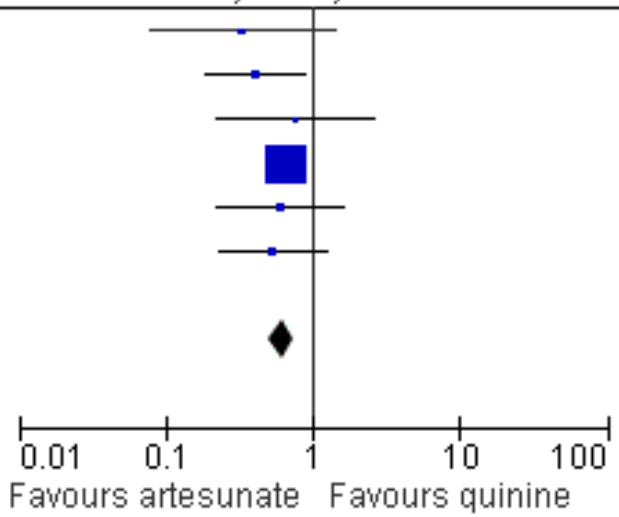
Cochrane Infectious Diseases

READ-It

Artesunate versus quinine in severe malaria

Outcome: Death

Study or Subgroup	Artesunate Events	Artesunate Total	Quinine Events	Quinine Total	Weight	Risk Ratio M-H, Fixed, 95% CI
Anh 1989	2	19	7	22	3.0%	0.33 [0.08, 1.41]
Anh 1995	8	99	18	91	8.7%	0.41 [0.19, 0.89]
Cao 1997	4	37	5	35	2.4%	0.76 [0.22, 2.59]
Dondorp 2005	107	730	164	731	76.3%	0.65 [0.52, 0.81]
Hien 1992	5	31	8	30	3.8%	0.60 [0.22, 1.64]
Newton 2003	7	59	12	54	5.8%	0.53 [0.23, 1.26]
Total (95% CI)		**975**		**963**	**100.0%**	**0.62 [0.51, 0.75]**
Total events						
Heterogeneity: Chi² = 2.26						
Test for overall effect: Z =						



Using Artesunate instead of quinine reduces the risk of death by 38%

- Do you believe this?
- What additional information would you want to know before you believe it?

Cochrane Infectious Diseases

READ-It

Use whiteboard

The usefulness of an estimate of the size of an effect depends on our certainty around that estimate.

READ-It

Gordon H Guyatt et al. BMJ 2008;336:924-926

Systematic review

Formulate question

Select outcomes

Rate importance

P I C O

Outcome Critical

Outcome Critical

Outcome Important

Outcome Not important

Outcomes across studies

Create evidence profile with GRADEpro

Summary of findings & estimate of effect for each outcome

Rate quality of evidence for each outcome

High
Moderate
Low
Very low

RCT start high, obs. data start low

Grade down

1. Risk of bias
2. Inconsistency
3. Indirectness
4. Imprecision
5. Publication bias

Grade up

1. Large effect
2. Dose response
3. Confounders

Guideline development

Rate
overall quality of evidence
across outcomes based on
lowest quality
of ***critical*** outcomes

Formulate recommendations:

- For or against (direction)
- Strong or weak (strength)

By considering:

- ❑ Quality of evidence
- ❑ Balance benefits/harms
- ❑ Values and preferences

Revise if necessary by considering:

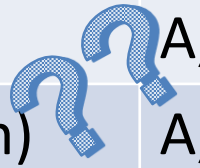
- ❑ Resource use (cost)

- "We recommend using…"
- "We suggest using…"
- "We recommend against using…"
- "We suggest against using…"

YFY 2008-11-11

The origin of GRADE: Which hierarchy?

Before GRADE	"Levels of evidence"	"Recommendation grades"
Oxford CEBM	1a, 1b, 2a, 2b, 3a, 3b, 4, 5	A, B, C, D
US PSTF	I, II-1, II-2, II-3, III	A, B, C, D, I
ACC/AHA	I, II-a, II-b, III, III (harm)	A, B-R, B-NR, C-LD, C-EO



Cochrane Infectious Diseases

READ-It

Oxford Centre for Evidence-Based Medicine

Level	Therapy / Prevention, Aetiology / Harm
1a	SR (with homogeneity) of RCTs
1b	Individual RCT (with narrow Confidence Interval”)
2a	SR (with homogeneity) of cohort studies
2b	Individual cohort study (including low quality RCT; e.g., <80% follow-up)
2c	“Outcomes” Research; Ecological studies
3a	SR (with homogeneity) of case-control studies
3b	Individual Case-Control Study
4	Case-series (and poor quality cohort and case-control studies)
5	Expert opinion without explicit critical appraisal, or based on physiology, bench research or “first principles”

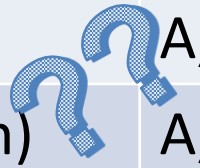
Oxford Centre for Evidence-Based Medicine

Grades of Recommendation

A	consistent level 1 studies
B	consistent level 2 or 3 studies or extrapolations from level 1 studies
C	level 4 studies or extrapolations from level 2 or 3 studies
D	level 5 evidence or troublingly inconsistent or inconclusive studies of any level

The origin of GRADE: Which hierarchy?

Before GRADE	"Levels of evidence"	"Recommendation grades"
Oxford CEBM	1a, 1b, 2a, 2b, 3a, 3b, 4, 5	A, B, C, D
US PSTF	I, II-1, II-2, II-3, III	A, B, C, D, I
ACC/AHA	I, II-a, II-b, III, III (harm)	A, B-R, B-NR, C-LD, C-EO



Cochrane Infectious Diseases

READ-It

GRADE

Grading of Recommendations Assessment, Development and Evaluation

An approach to formulating:

- Evidence-based recommendations
- Through a transparent and systematic process
- With an explicit link between evidence and recommendations

www.gradeworkinggroup.org

Cochrane Infectious Diseases

READ-It

Applying GRADE

Cochrane Infectious Diseases

READ-It

Levels of Certainty

Level	What it means
⊕⊕⊕⊕ HIGH	We have a lot of confidence that the true effect is similar to the estimated effect
⊕⊕⊕⊖ MODERATE	We believe that the true effect is probably close to the estimated effect
⊕⊕⊖⊖ LOW	The true effect might be markedly different from the estimated effect
⊕⊖⊖⊖ VERY LOW	The true effect is probably markedly different from the estimated effect

Cochrane Infectious Diseases

READ-It

Levels of Certainty: Plain language

Level	What it means ***Compared to quinine…***
⊕⊕⊕⊕ HIGH	Artesunate reduces mortality
⊕⊕⊕⊖ MODERATE	Artesunate probably reduces mortality
⊕⊕⊖⊖ LOW	Artesunate may reduce mortality
⊕⊖⊖⊖ VERY LOW	We *don't know* if artesunate reduces mortality

Cochrane Infectious Diseases

READ-It

How to GRADE

Level

⊕⊕⊕⊕
HIGH

Evidence from RCTs is considered **high certainty**, but may be downgraded

⊕⊕⊕⊖
MODERATE

⊕⊕⊖⊖
LOW

⊕⊖⊖⊖
VERY LOW

Evidence from observational studies is considered **low certainty,** but may be up- or down-graded

What would you take into account when considering how much confidence you have in the results of a RCT?

When to downgrade evidence

Level

⊕⊕⊕⊕
HIGH

Evidence from RCTs is considered **high certainty**, but may be downgraded

⊕⊕⊕⊖
MODERATE

⊕⊕⊖⊖
LOW

⊕⊖⊖⊖
VERY LOW

5 reasons to downgrade:

- Risk of bias
- Inconsistency
- Indirectness
- Imprecision
- Other (publication bias)

5 reasons to downgrade:

Risk of bias	Is the risk of bias in individual studies sufficiently large to reduce your confidence in the estimated effect?
Inconsistency	Do the studies have inconsistent effects? Are the studies and their outcomes too heterogenous to compare?
Indirectness	Do the trials reporting this outcome directly address the question we are asking?
Imprecision	Would your clinical action change if either the upper or lower boundary of the 95% confidence interval represented the truth?
Other	Should you suspect publication bias? For example, are the studies all small, and commercially funded?

When to upgrade evidence

Level

⊕⊕⊕⊕
HIGH

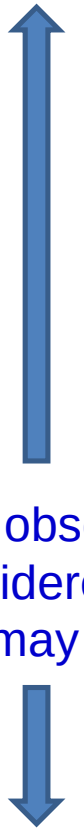
⊕⊕⊕⊖
MODERATE

⊕⊕⊖⊖
LOW

⊕⊖⊖⊖
VERY LOW

Reasons to upgrade:

- Strong association
- Confounders act to reduce observed effect
- Dose-response effect



Evidence from observational studies is considered **low certainty,** but may be up- or down-graded

Exercise

Use the five posters around the room to judge how well the evidence answers this question:

In people with severe malaria does treatment with Artesunate (i.v.) reduce death compared to treatment with Quinine (i.v.)?

Is the evidence:

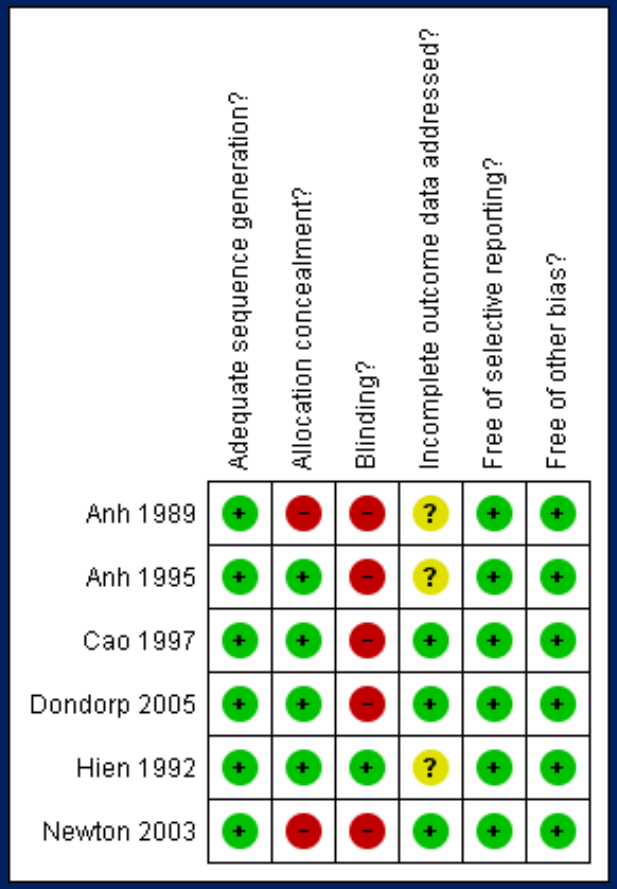
High Certainty: Artesunate reduces death compared to quinine

Moderate Certainty: Artesunate probably reduces death compared to quinine

Low Certainty: Artesunate may reduce death compared to quinine but…

Very Low Certainty: We don't know whether artesunate reduces death

Feedback (1): Risk of bias

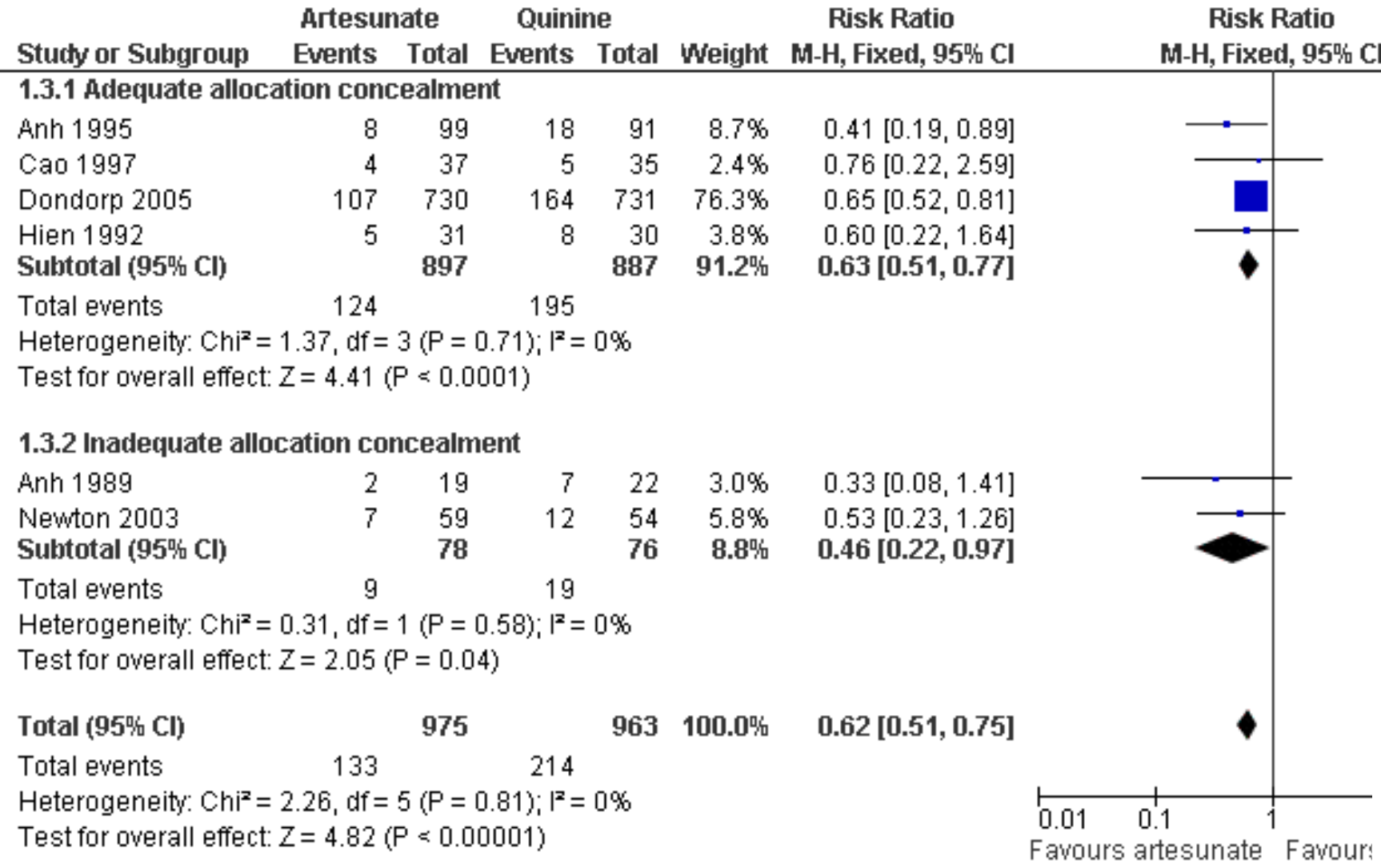


Is the risk of bias in individual studies sufficiently large to reduce your confidence in the estimated effect?

Sensitivity analyses, removing the trials at high risk of bias, can help inform the judgement

Risk of bias: sensitivity analysis

Artesunate versus quinine in severe malaria; Outcome: Death



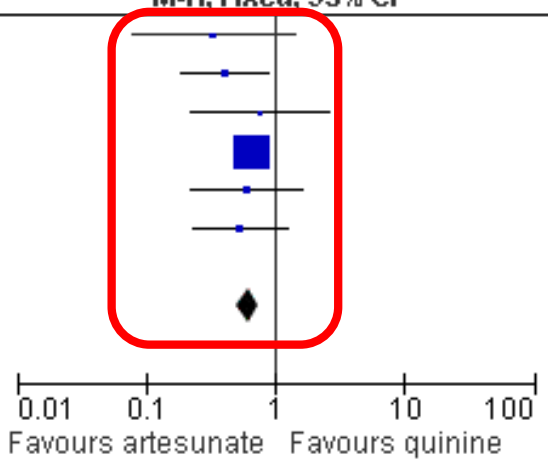
Feedback (2): Inconsistency

The **eyeball test:** Are estimates similar, and do CIs overlap?

Study or Subgroup	Artesunate Events	Total	Quinine Events	Total	Weight	Risk Ratio M-H, Fixed, 95% CI
Anh 1989	2	19	7	22	3.0%	0.33 [0.08, 1.41]
Anh 1995	8	99	18	91	8.7%	0.41 [0.19, 0.89]
Cao 1997	4	37	5	35	2.4%	0.76 [0.22, 2.59]
Dondorp 2005	107	730	164	731	76.3%	0.65 [0.52, 0.81]
Hien 1992	5	31	8	30	3.8%	0.60 [0.22, 1.64]
Newton 2003	7	59	12	54	5.8%	0.53 [0.23, 1.26]
Total (95% CI)		**975**		**963**	**100.0%**	**0.62 [0.51, 0.75]**
Total events	133		214			

Heterogeneity: Chi² = 2.26, df = 5 (P = 0.81); I² = 0%

Test for overall effect: Z = 4.82 (P < 0.00001)



The **statistical tests:** Is there significant unexplained heterogeneity?

Do the studies have inconsistent effects? Are the studies and their outcomes too heterogenous to compare?

Heterogeneity

Heterogeneity is observed differences in the results of different trials:

- A fixed effects model assumes that there is one true effect that the trials are attempting to measure – When heterogeneity is high this assumption no longer holds.
- When there is heterogeneity but it is still meaningful to combine trials a random effects model can be used
- If there is too much heterogeneity it may be inappropriate or meaningless to pool the trials

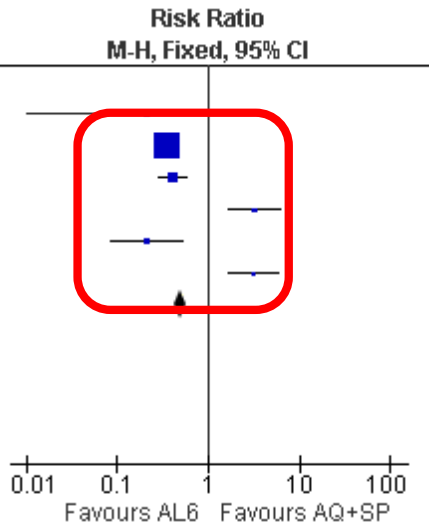
There are many causes of heterogeneity, including different populations, interventions, and outcomes.

Inconsistency: Example

Artemether-lumefantrine versus sulfadoxine-pyrimethamine/amodiaquine for uncomplicated malaria
Outcome: Treatment failure at day 28

The **eyeball test:** Are estimates similar, and do CIs overlap?

Study or Subgroup	AL6 Events	AL6 Total	AQ+SP Events	AQ+SP Total	Weight	Risk Ratio M-H, Fixed, 95% CI	Year
11.1.1 Africa							
Faye 2003 SEN	0	147	2	156	0.6%	0.21 [0.01, 4.38]	2003
Mutabingwa 2004 TZA	103	485	282	463	67.6%	0.35 [0.29, 0.42]	2004
Fanello 2004 RWA	36	246	89	247	20.8%	0.41 [0.29, 0.57]	2004
Zongo 2005 BFA	37	245	11	233	2.6%	3.20 [1.67, 6.12]	2005
Dorsey 2006 UGA	5	100	25	105	5.7%	0.21 [0.08, 0.53]	2006
Zongo 2007 BFA	36	178	11	171	2.6%	3.14 [1.65, 5.97]	2007
Subtotal (95% CI)		**1401**		**1375**	**100.0%**	**0.50 [0.43, 0.58]**	
Total events	217		420				
Heterogeneity: Chi² = 82.52, df = 5 (P < 0.00001); I² = 94%							
Test for overall effect: Z = 9.54 (P < 0.00001)							



The **statistical tests:** Is there significant unexplained heterogeneity?

Cochrane Infectious Diseases

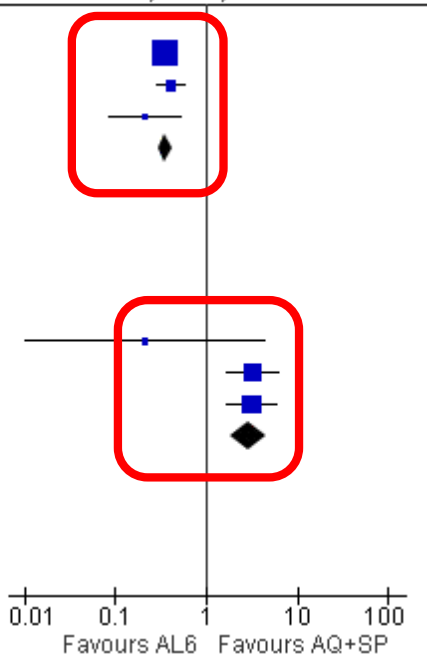
READ-It

Inconsistency: Example

Artemether-lumefantrine versus sulfadoxine-pyrimethamine/amodiaquine for uncomplicated malaria
Outcome: Treatment failure at day 28

The **eyeball test:** Are estimates similar, and do CIs overlap?

Study or Subgroup	AL6 Events	AL6 Total	AQ+SP Events	AQ+SP Total	Weight	M-H, Fixed, 95% CI	Year
11.1.1 East Africa							
Mutabingwa 2004 TZA	103	485	282	463	71.8%	0.35 [0.29, 0.42]	2004
Fanello 2004 RWA	36	246	89	247	22.1%	0.41 [0.29, 0.57]	2004
Dorsey 2006 UGA	5	100	25	105	6.1%	0.21 [0.08, 0.53]	2006
Subtotal (95% CI)		**831**		**815**	**100.0%**	**0.35 [0.30, 0.41]**	
Total events	144		396				
Heterogeneity: Chi² = 1.88, df = 2 (P = 0.39); I² = 0%							
Test for overall effect: Z = 12.61 (P < 0.00001)							
11.1.2 West Africa							
Faye 2003 SEN	0	147	2	156	9.7%	0.21 [0.01, 4.38]	2003
Zongo 2005 BFA	37	245	11	233	45.2%	3.20 [1.67, 6.12]	2005
Zongo 2007 BFA	36	178	11	171	45.0%	3.14 [1.65, 5.97]	2007
Subtotal (95% CI)		**570**		**560**	**100.0%**	**2.88 [1.86, 4.47]**	
Total events	73		24				
Heterogeneity: Chi² = 3.02, df = 1 (P = 0.22); I² = 34%							
Test for overall effect: Z = 4.72 (P < 0.00001)							

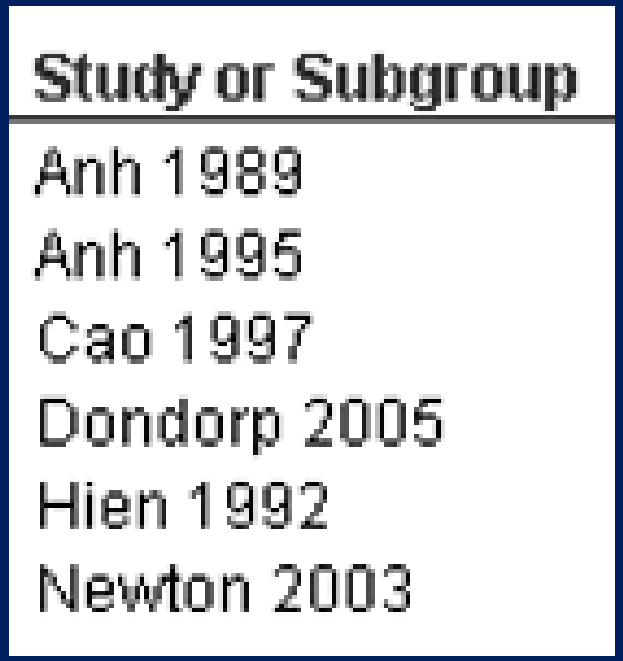


The **statistical tests:** Is there significant unexplained heterogeneity?

Cochrane Infectious Diseases

READ-It

Feedback (3): Indirectness



Do the trials reporting this outcome directly address the question we are asking?

- **Population:**
  - Right patients? Right country? Right illness severity / diagnosis?
- **Intervention:**
  - Right drug? Right dose?
- **Comparator:**
  - Did the control group receive current standard care?
- **Outcome:**
  - Direct measurement? Correct f/u?

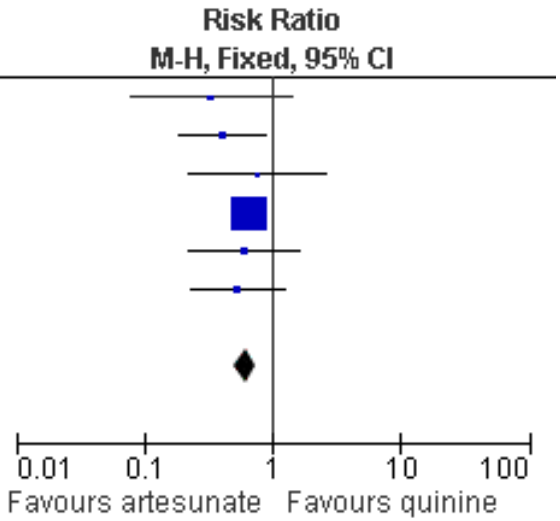
Indirectness

Artesunate versus quinine in severe malaria; Outcome: Death

Study or Subgroup	Artesunate Events	Artesunate Total	Quinine Events	Quinine Total	Weight	Risk Ratio M-H, Fixed, 95% CI
Anh 1989	2	19	7	22	3.0%	0.33 [0.08, 1.41]
Anh 1995	8	99	18	91	8.7%	0.41 [0.19, 0.89]
Cao 1997	4	37	5	35	2.4%	0.76 [0.22, 2.59]
Dondorp 2005	107	730	164	731	76.3%	0.65 [0.52, 0.81]
Hien 1992	5	31	8	30	3.8%	0.60 [0.22, 1.64]
Newton 2003	7	59	12	54	5.8%	0.53 [0.23, 1.26]
Total (95% CI)		**975**		**963**	**100.0%**	**0.62 [0.51, 0.75]**
Total events	133		214			

Heterogeneity: Chi² = 2.26, df = 5 (P = 0.81); I² = 0%

Test for overall effect: Z = 4.82 (P < 0.00001)



Population: Only 2 out of 6 trials included children.

All trials were conducted in Asia

Intervention: 5 out of 6 trials used IV artesunate, one used IM

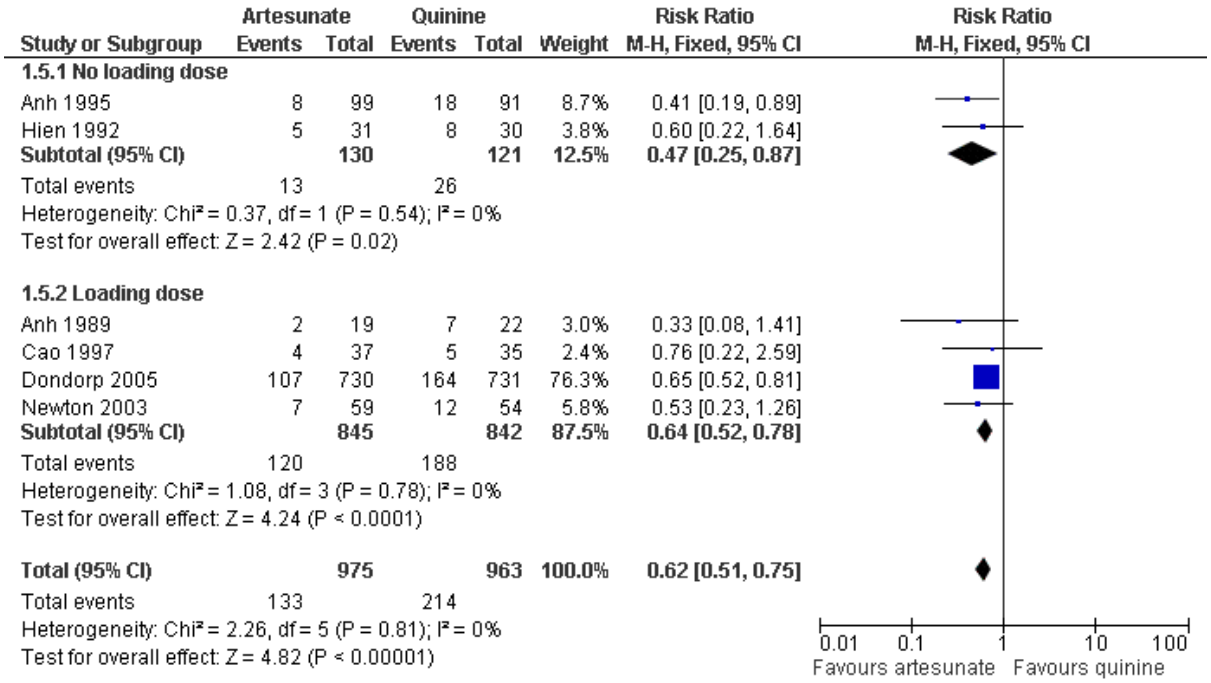
Control: Only 4 trials gave the loading dose of Quinine

Cochrane Infectious Diseases

READ-It

Indirectness

Artesunate versus quinine in severe malaria
Outcome: Death (sub-grouped by loading dose of quinine)



Study or Subgroup	Artesunate Events	Artesunate Total	Quinine Events	Quinine Total	Weight	Risk Ratio M-H, Fixed, 95% CI
1.5.1 No loading dose						
Anh 1995	8	99	18	91	8.7%	0.41 [0.19, 0.89]
Hien 1992	5	31	8	30	3.8%	0.60 [0.22, 1.64]
Subtotal (95% CI)		**130**		**121**	**12.5%**	**0.47 [0.25, 0.87]**
Total events	13		26			
Heterogeneity: Chi² = 0.37, df = 1 (P = 0.54); I² = 0%						
Test for overall effect: Z = 2.42 (P = 0.02)						
1.5.2 Loading dose						
Anh 1989	2	19	7	22	3.0%	0.33 [0.08, 1.41]
Cao 1997	4	37	5	35	2.4%	0.76 [0.22, 2.59]
Dondorp 2005	107	730	164	731	76.3%	0.65 [0.52, 0.81]
Newton 2003	7	59	12	54	5.8%	0.53 [0.23, 1.26]
Subtotal (95% CI)		**845**		**842**	**87.5%**	**0.64 [0.52, 0.78]**
Total events	120		188			
Heterogeneity: Chi² = 1.08, df = 3 (P = 0.78); I² = 0%						
Test for overall effect: Z = 4.24 (P < 0.0001)						
Total (95% CI)		**975**		**963**	**100.0%**	**0.62 [0.51, 0.75]**
Total events	133		214			
Heterogeneity: Chi² = 2.26, df = 5 (P = 0.81); I² = 0%						
Test for overall effect: Z = 4.82 (P < 0.00001)						

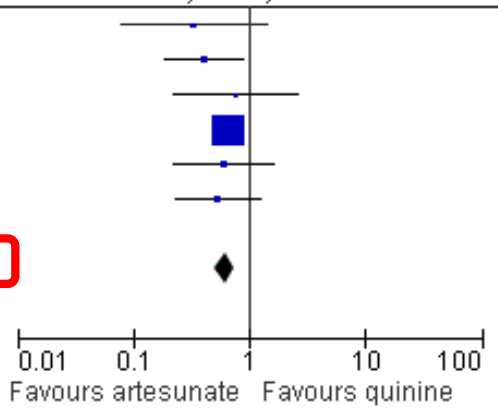
Did the inclusion of trials without a quinine loading dose effect the results?

Should you downgrade for indirectness?

Feedback (4): Imprecision

Study or Subgroup	Artesunate Events	Total	Quinine Events	Total	Weight	Risk Ratio M-H, Fixed, 95% CI
Anh 1989	2	19	7	22	3.0%	0.33 [0.08, 1.41]
Anh 1995	8	99	18	91	8.7%	0.41 [0.19, 0.89]
Cao 1997	4	37	5	35	2.4%	0.76 [0.22, 2.59]
Dondorp 2005	107	730	164	731	76.3%	0.65 [0.52, 0.81]
Hien 1992	5	31	8	30	3.8%	0.60 [0.22, 1.64]
Newton 2003	7	59	12	54	5.8%	0.53 [0.23, 1.26]
Total (95% CI)		**975**		**963**	**100.0%**	**0.62 [0.51, 0.75]**
Total events	133		214			

Heterogeneity: Chi² = 2.26, df = 5 (P = 0.81); I² = 0%
Test for overall effect: Z = 4.82 (P < 0.00001)



Would your clinical action change if either the upper or lower boundary of the 95% confidence interval represented the truth?

Does the CI include:

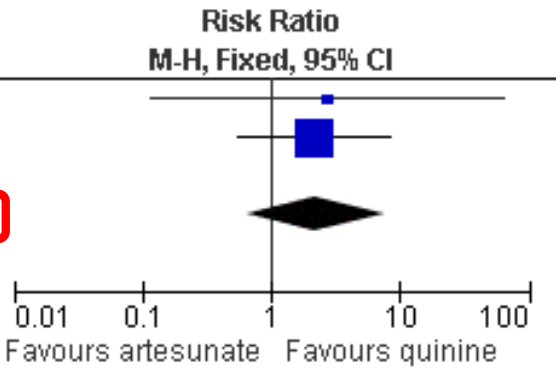
- Clinically important benefit?
- No clinically important difference?
- Clinically important harm?

Imprecision

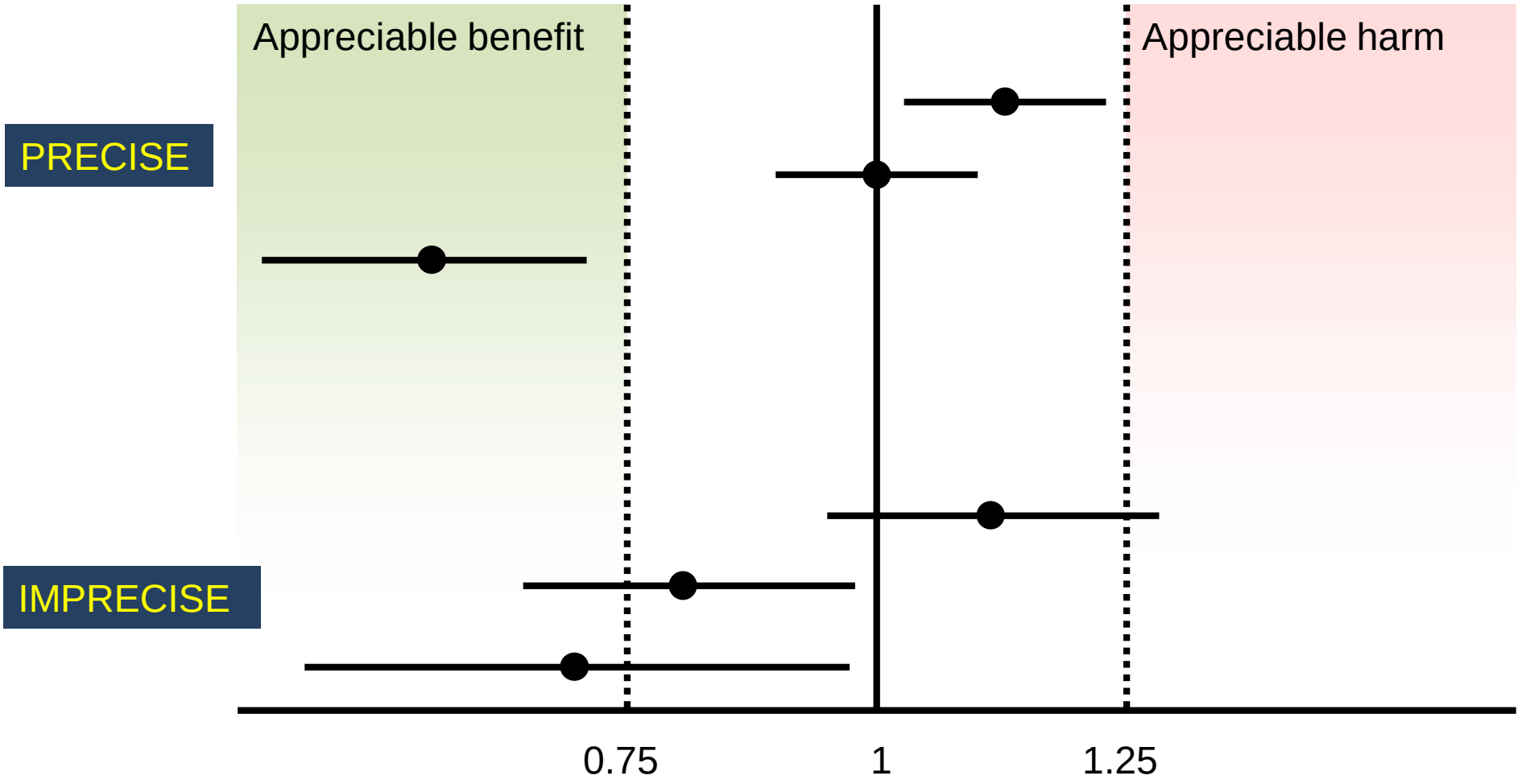
Artesunate versus quinine in severe malaria
Outcome: Neurological disability at discharge

Study or Subgroup	Artesunate Events	Artesunate Total	Quinine Events	Quinine Total	Weight	Risk Ratio M-H, Fixed, 95% CI
Cao 1997	1	33	0	30	14.3%	2.74 [0.12, 64.69]
Dondorp 2005	7	623	3	567	85.7%	2.12 [0.55, 8.17]
Total (95% CI)		**656**		**597**	**100.0%**	**2.21 [0.64, 7.63]**
Total events	8		3			

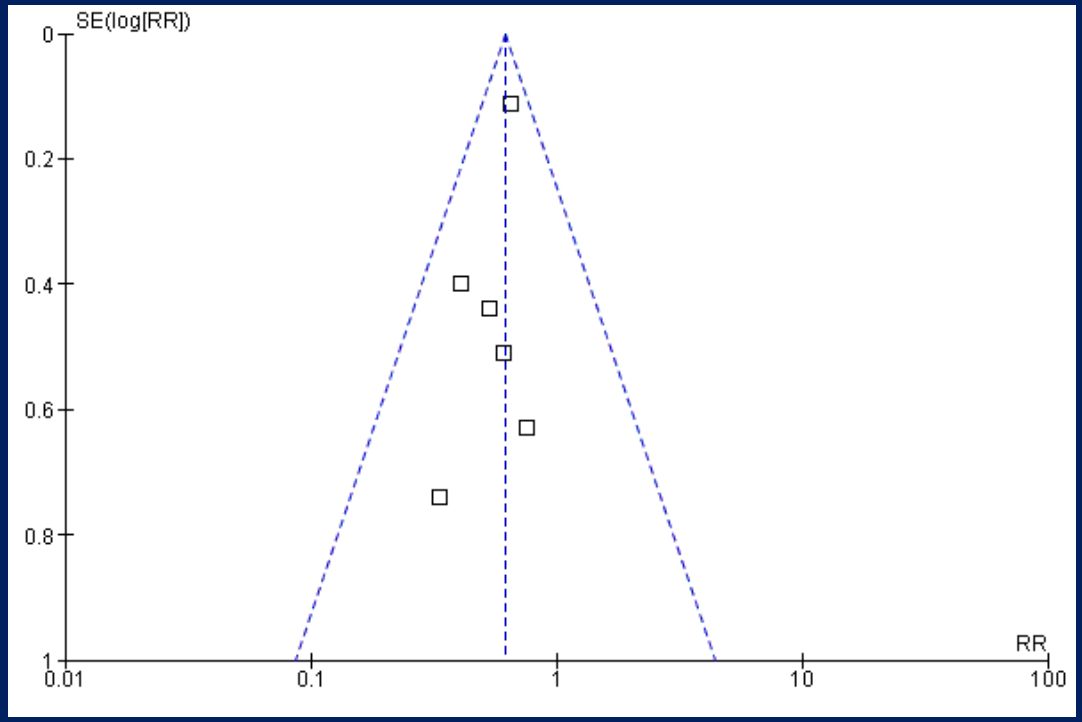
Heterogeneity: Chi² = 0.02, df = 1 (P = 0.89); I² = 0%
Test for overall effect: Z = 1.26 (P = 0.21)



What about for this outcome?


Appreciable benefit
Appreciable harm
PRECISE
IMPRECISE
0.75
1
1.25

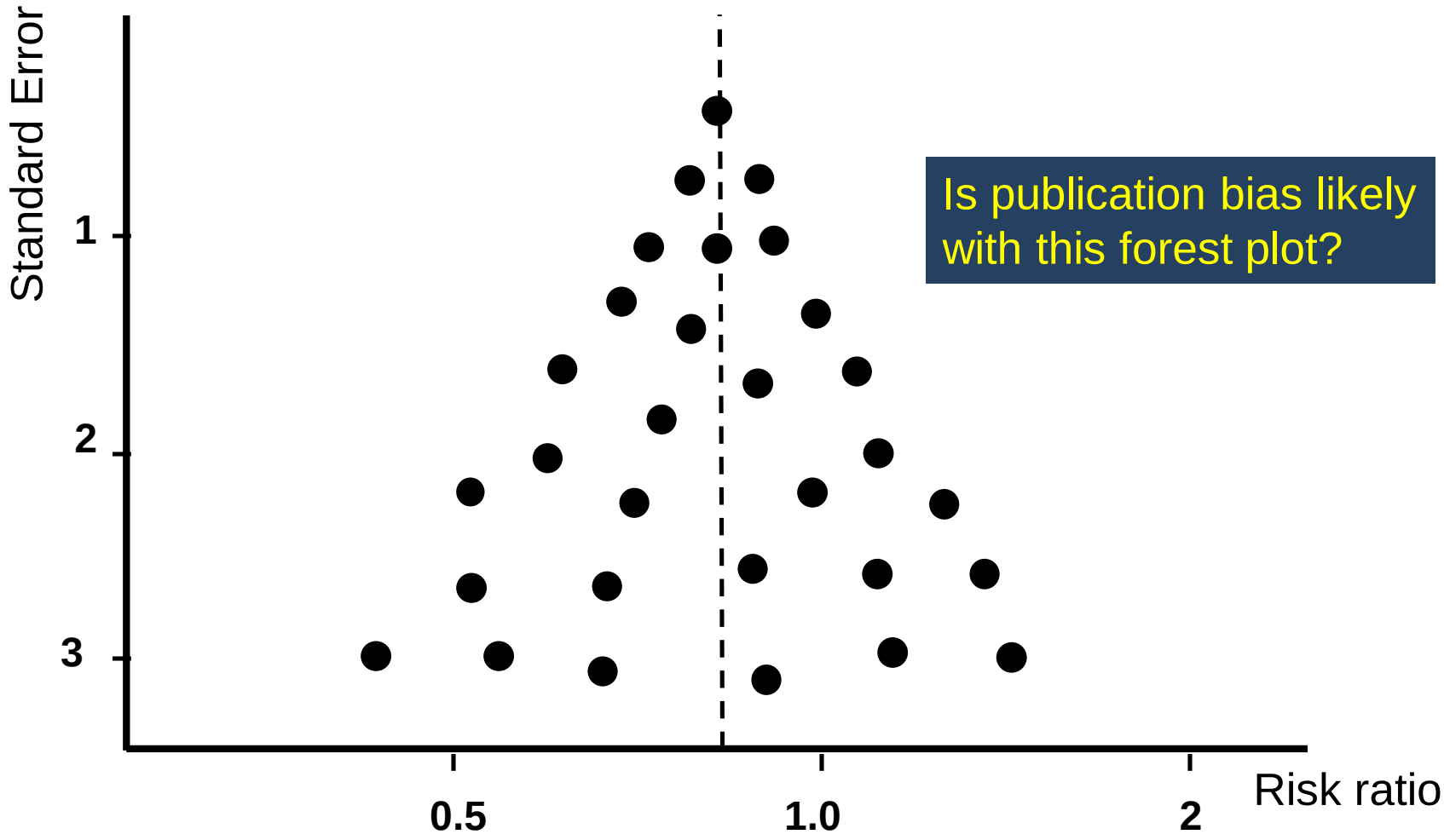
Feedback (5): Other bias



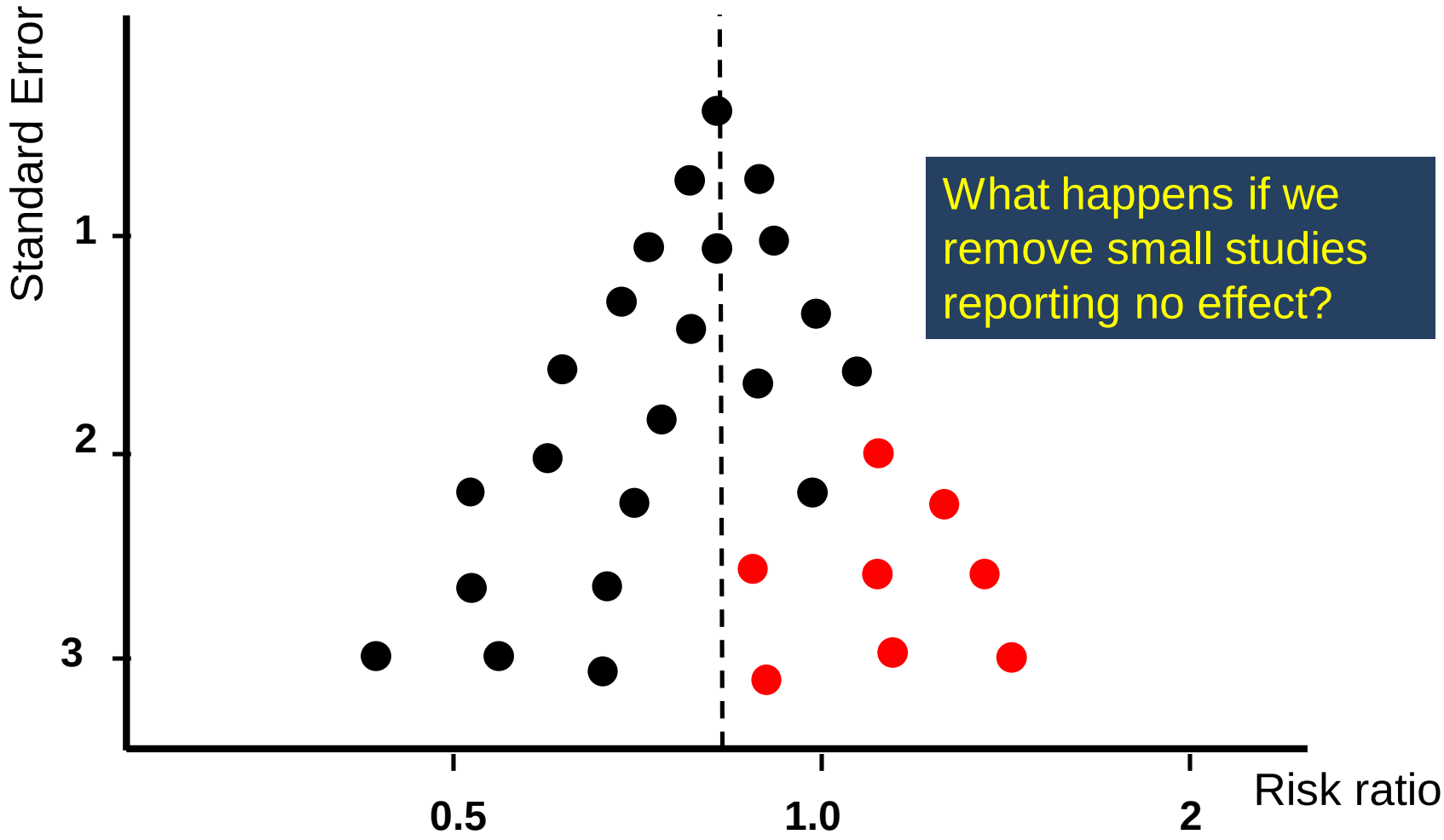
Should you suspect publication bias?

For example, are the studies all small, and commercially funded?

Other bias: publication bias



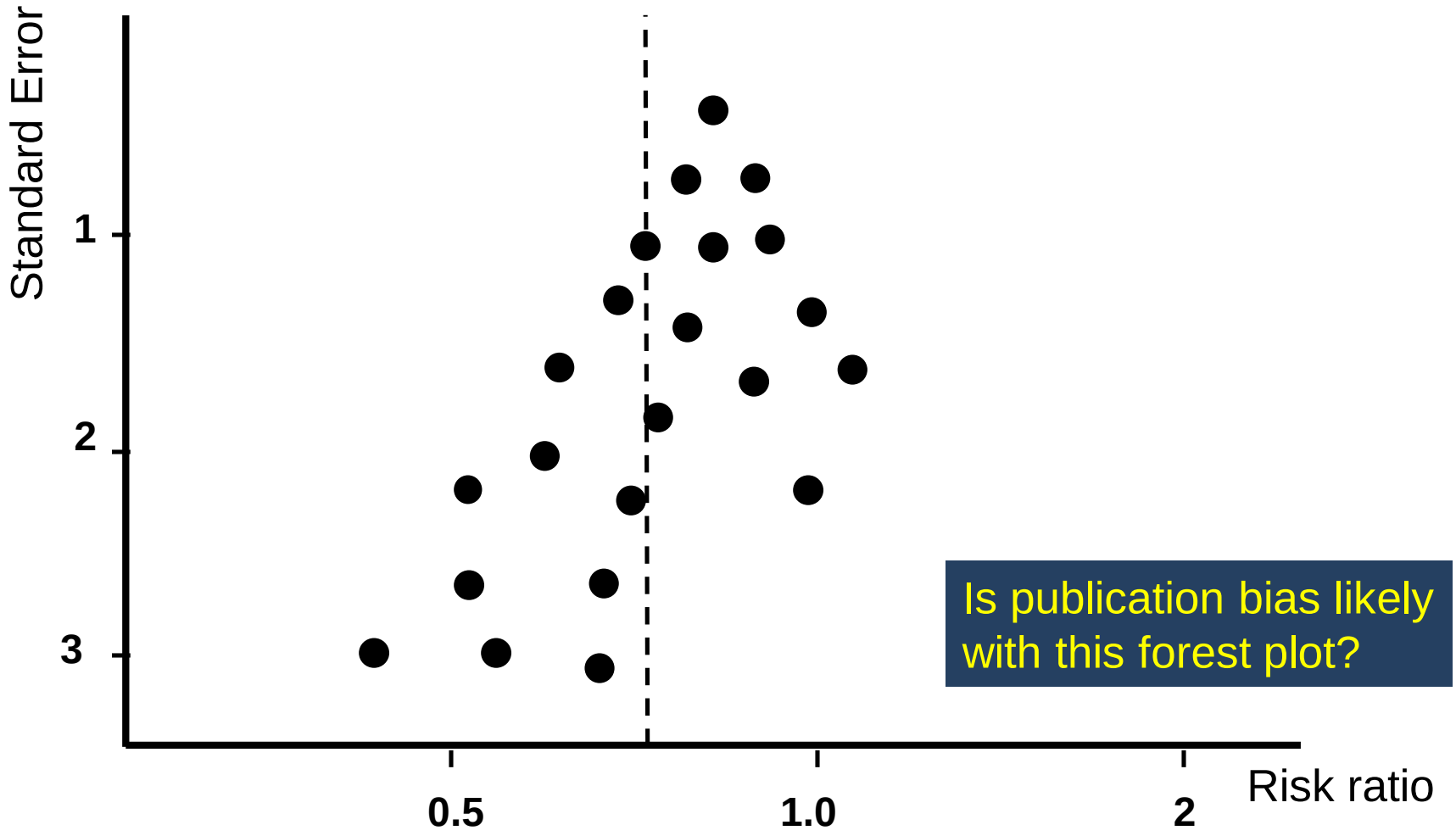
Other bias: publication bias



Cochrane Infectious Diseases

READ-It

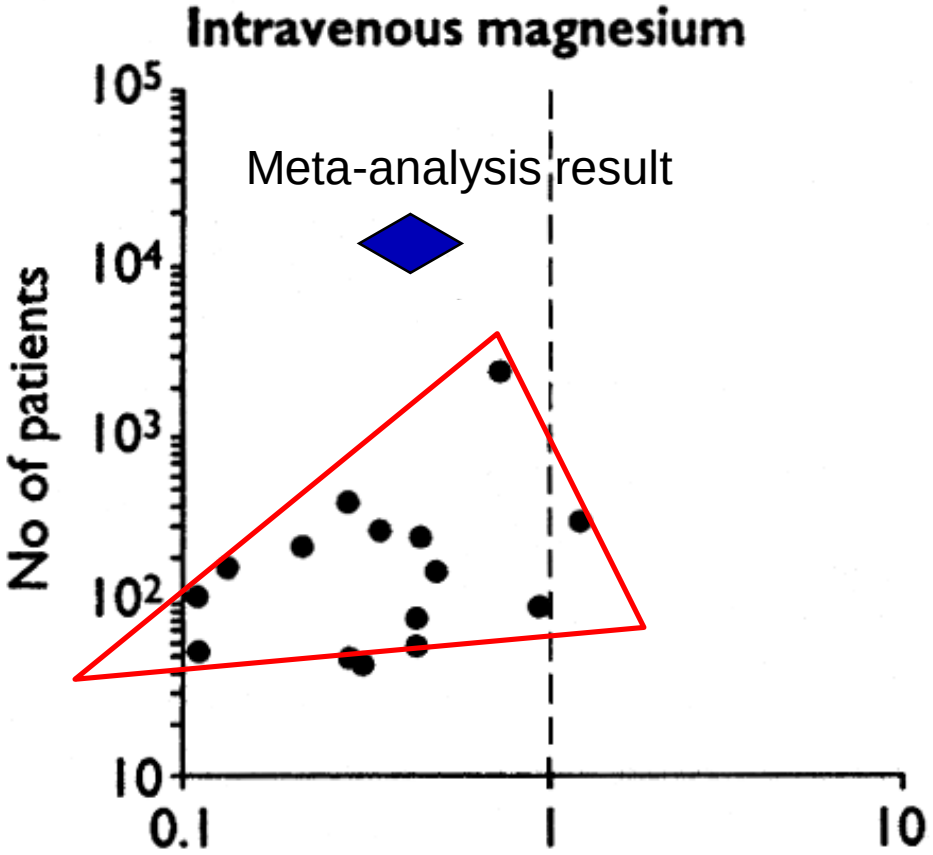
Other bias: publication bias



Cochrane Infectious Diseases

READ-It

Other bias: publication bias



Is publication bias likely with this forest plot?

Would you be certain in the results of the meta-analysis?

New search | Conclusions changed

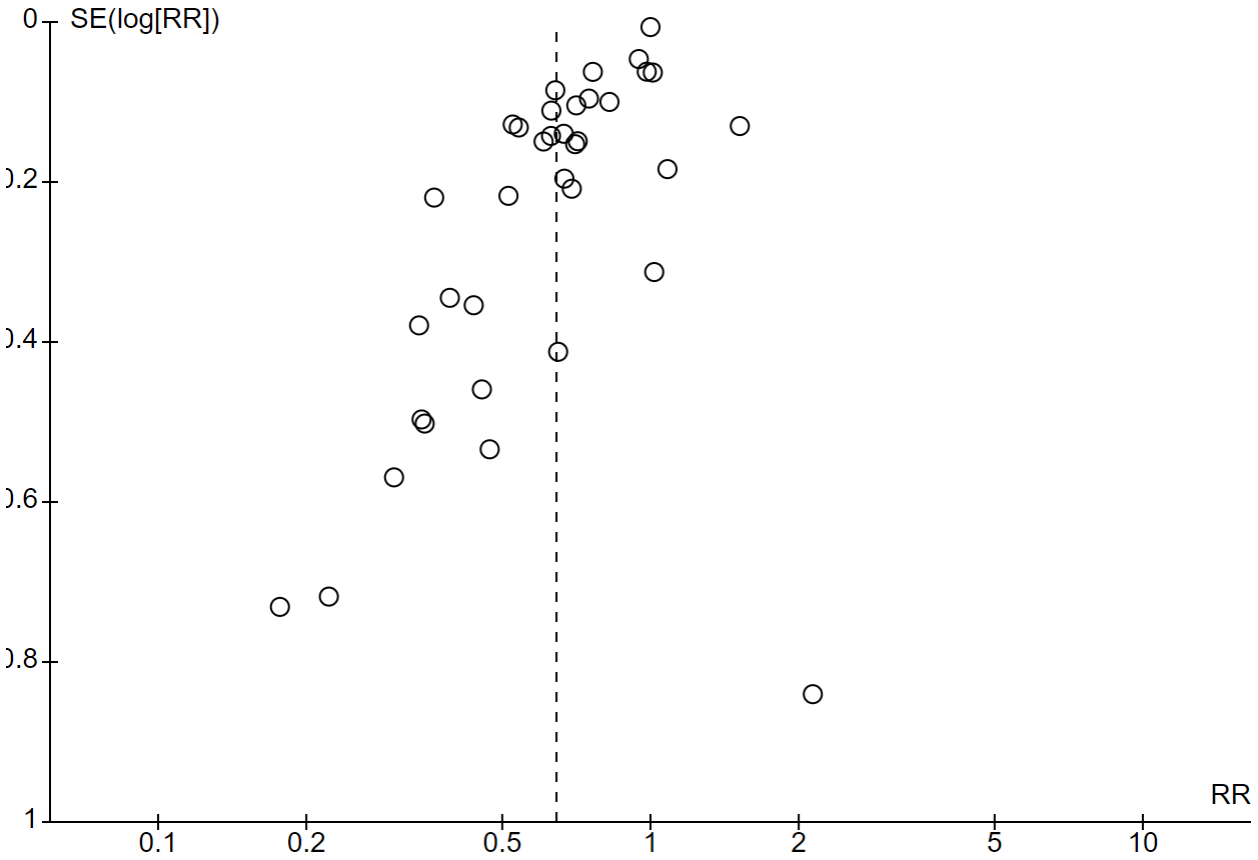
Probiotics for treating acute infectious diarrhoea

Shelui Collinson, Andrew Deans, April Padua-Zamora, Germana V Gregorio, Chao Li, Leonila F Dans, Stephen J Allen

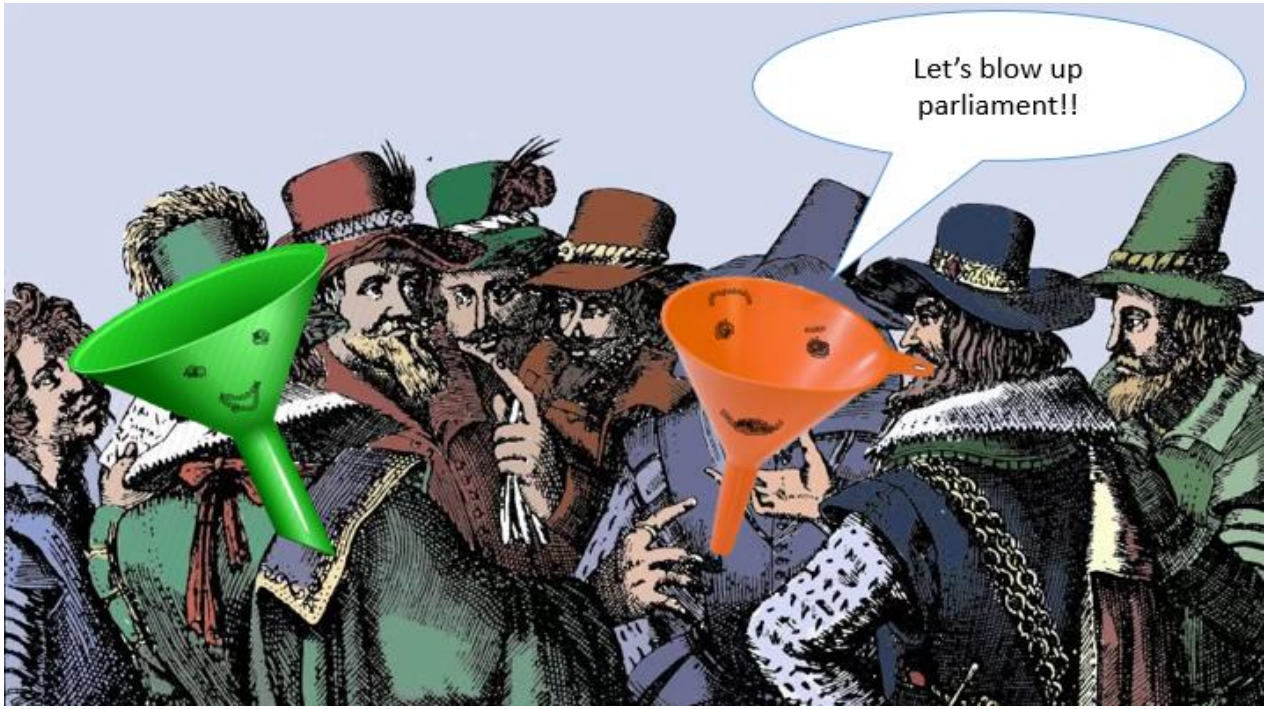
Authors' declarations of interest

Version published: 08 December 2020 Version history

https://doi.org/10.1002/14651858.CD003048.pub4



Gunpowder, funnels, and plot.



Feedback

Would you downgrade the certainty for the mortality outcome in this review?

GRADE

Artesunate versus quinine in severe malaria

Outcome: Death

Study limitations: Limiting trials to only those with adequate allocation concealment did not change result

Inconsistency: No statistical heterogeneity

Indirectness: Very little data from children, no African trials

Precision: Precise result of reduced deaths in adults in Asia

Publication bias: Possibly some evidence of publication bias, but result from largest trial still indicates benefit

NNT = 12 (95%CI: 9 to 18)

In adults: High Certainty evidence

In children: Low Certainty evidence

Would you recommend Artesunate in adults?

Would you recommend Artesunate in children?

What other factors might you want to consider?

What if…

Artesunate was 10 x more expensive?

(resource use/cost)

Artesunate required specialised monitoring?

(feasibility)

Artesunate caused more neurological sequelae?

(balance between benefits and harms)

Formulate question

Select outcomes

Rate importance

Outcomes across studies

Create evidence profile with GRADEpro

Rate quality of evidence for each outcome

P
I
C
O

Outcome Critical

Outcome Critical

Outcome Important

Outcome Not important

Summary of findings & estimate of effect for each outcome

High
Moderate
Low
Very low

RCT start high, obs. data start low

Grade down

1. Risk of bias
2. Inconsistency
3. Indirectness
4. Imprecision
5. Publication bias

Grade up

1. Large effect
2. Dose response
3. Confounders

Systematic review

Guideline development

Rate
overall quality of evidence
across outcomes based on
lowest quality
of ***critical*** outcomes

Formulate recommendations:

- For or against (direction)
- Strong or weak (strength)

By considering:

- ❑ Quality of evidence
- ❑ Balance benefits/harms
- ❑ Values and preferences

Revise if necessary by considering:

- ❑ Resource use (cost)

- "We recommend using…"
- "We suggest using…"
- "We recommend against using…"
- "We suggest against using…"

YFY 2008-11-11

Moving from evidence to recommendations

Requires further consideration of:

- The balance of benefits and harms
- Feasibility
- Resource implications/costs

It is therefore possible to make:

- STRONG recommendations based on LOW certainty evidence
- Recommendations NOT to do something even with HIGH certainty evidence that it works

Questions to ask…

- Is the problem a priority?
- How substantial are the desirable anticipated effects?
- How substantial are the undesirable anticipated effects?
- What is the overall certainty of the evidence of effects?
- Does the balance between desirable and undesirable effects favour the intervention or the comparison?