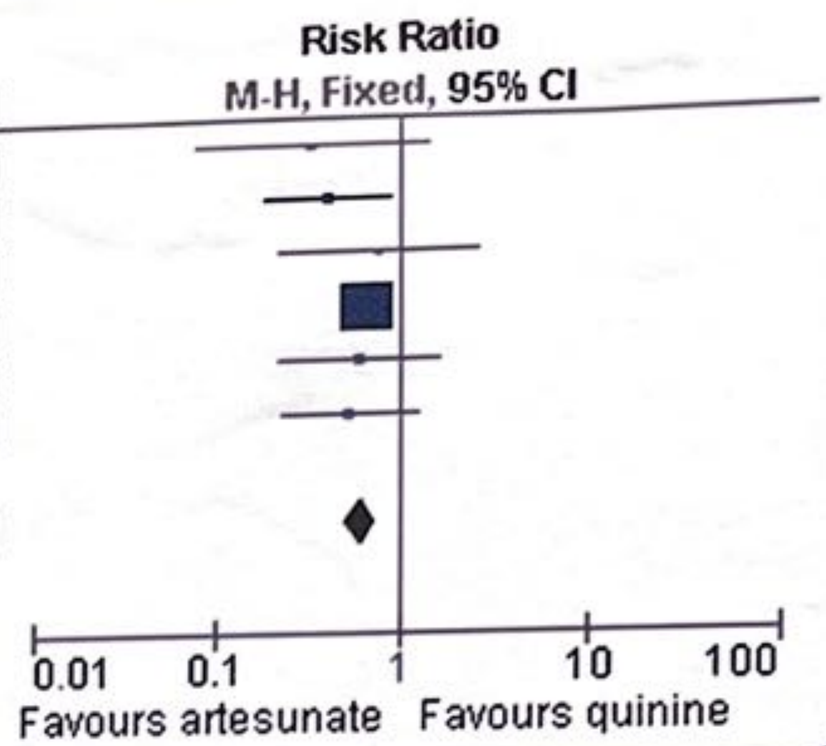
Artesunate compared to Quinine for treating adults and children with severe malaria

Patient or population: Adults and children with severe malaria
Settings: Malaria endemic settings in sub-Saharan Africa
Intervention: Artesunate 2.4 mg/kg repeated at 12, 24 and then every 24 hours
Comparison: Quinine 20 mg/kg loading dose then 10 mg/kg every 8 hours

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: $Chi^2 = 2.26$, $df = 5$ ($P = 0.81$); $I^2 = 0\%$
Test for overall effect: $Z = 4.82$ ($P < 0.00001$)



GRADE Assessment: Risk of bias

A high quality systematic review will assess the risk of bias of each individual included study, and consider this risk of bias when making conclusions about the results.

If the included studies are at high risk of bias, the result of the meta-analysis will also be at high risk of bias, and this may decrease your confidence in the results.

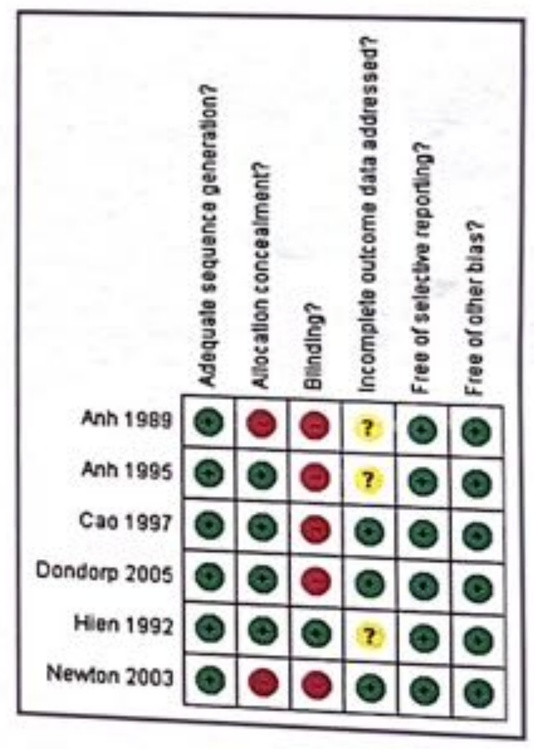
In Cochrane Reviews six criteria are commonly used to asses the risk of bias of randomized controlled trials (see figure):

- **Random sequence generation** and **allocation concealment** are methods to prevent selection bias
- **Blinding** is a method to reduce performance or detection bias

Questions to consider:

1) Does a lack of blinding in 5 out of 6 studies decrease your confidence that artesunate reduces mortality compared to quinine?
2) Does a lack of allocation concealment in two out of six studies decrease your confidence that artesunate reduces mortality compared to quinine?
3) Does it matter which studies lack allocation concealment?
4) Is there any alternative analysis you would like to see?

Overall: Is there a serious risk of bias that decreases your confidence that artesunate reduces mortality compared to quinine?



Red = High risk of bias
Yellow = Unclear risk of bias
Green = Low risk of bias

Your judgement:

- Don't downgrade
- Downgrade by 1 for serious risk of bias
- Downgrade by 2 for very serious risk of bias

GRADE levels of 'Quality'		
4	High	RCTs start as high quality evidence and are downgraded when there are **serious** problems which decrease your confidence in the results of the systematic review
3	Moderate	
2	Low	
1	Very Low	

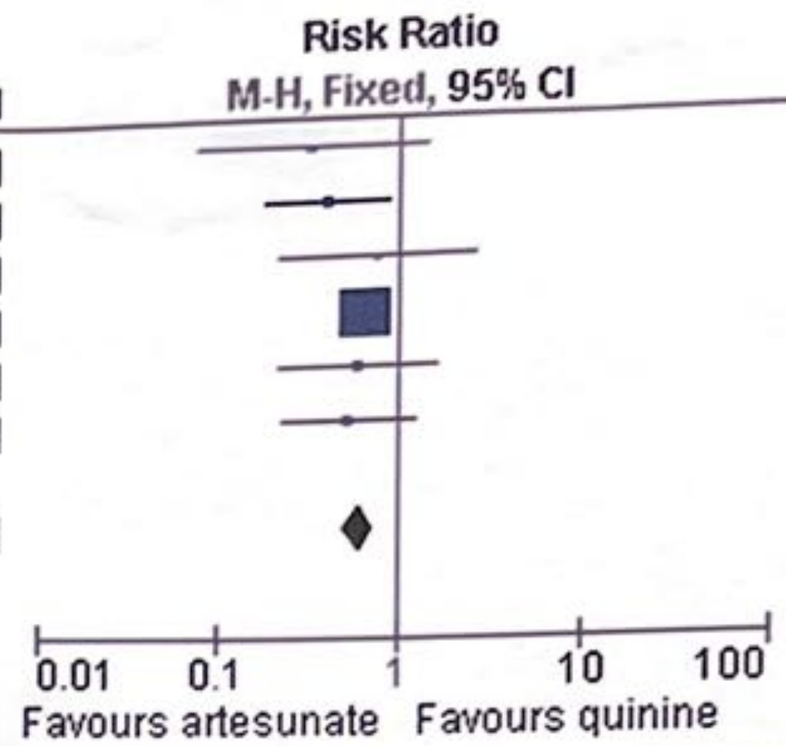
The South African Cochrane Centre
effective health care
Centre for Evidence Based Health Care
LSTM

Artesunate compared to Quinine for treating adults and children with severe malaria

Patient or population: Adults and children with severe malaria
Settings: Malaria endemic settings in sub-Saharan Africa
Intervention: Artesunate 2.4 mg/kg repeated at 12, 24 and then every 24 hours
Comparison: Quinine 20 mg/kg loading dose then 10 mg/kg every 8 hours

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)



GRADE Assessment: Consistency

In a meta-analysis the trials will always have slightly different results due to random error related to the sample size and the play of chance. Sometimes larger differences in trial results will be seen and this is called heterogeneity.

The presence of heterogeneity can be roughly assessed by looking at the forest plot (the Eye ball test). If the confidence intervals of the individual trials all overlap, then the differences can be explained by random error and heterogeneity is low.

There are also two statistical tests which help evaluate whether the observed differences in trial results can be explained by random error. These two tests are interpreted slightly differently:

Chi² p-value = The probability that the observed differences between trials occurred by chance.
I² test = The percentage of the observed differences between trials that is not due to chance

A high quality systematic review will investigate for causes of heterogeneity by conducting a series of sub-group analyses for potential effect modifiers such as participant age, country, year of study, drug dose, co-morbidities. When the heterogeneity cannot be explained by these sub-group analyses, this is called **Inconsistency**.

Questions to consider

1) Is there heterogeneity on the eyeball test? – Do the 95% confidence intervals overlap?
2) Is there heterogeneity as assessed by the Chi² p-value? – What is the probability that the observed differences in trial results are due to chance?
3) Is there heterogeneity as assessed by the I² - test? – What percentage of the observed differences in trial results is not due to chance?
4) Are there any sub-group analyses you would like to see to help you make a decision?

Overall: Is there serious inconsistency which would decrease your confidence that artesunate reduces mortality compared to quinine?

Your judgement:

- Don't downgrade
- Downgrade by 1 for serious inconsistency
- Downgrade by 2 for very serious inconsistency

GRADE levels of 'Quality'		
4	High	RCTs start as high quality evidence and are downgraded when there are **serious** problems which decrease your confidence in the results of the systematic review
3	Moderate	
2	Low	
1	Very Low	

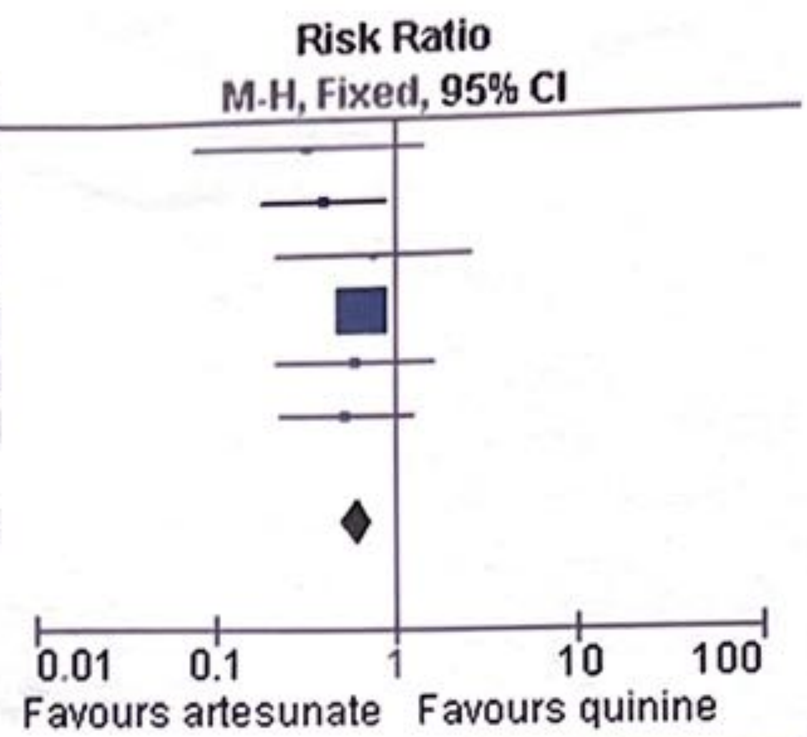
The South African Cochrane Centre
effective health care RESEARCH CONSORTIUM
Centre for Evidence Based Health Care
LSTM LIVERPOOL SCHOOL OF TROPICAL MEDICINE

Artesunate compared to Quinine for treating adults and children with severe malaria

Patient or population: Adults and children with severe malaria
Settings: Malaria endemic settings in sub-Saharan Africa
Intervention: Artesunate 2.4 mg/kg repeated at 12, 24 and then every 24 hours
Comparison: Quinine 20 mg/kg loading dose then 10 mg/kg every 8 hours

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: $Z = 4.82$ ($P < 0.00001$)



GRADE Assessment: Directness

Directness considers how well the evidence we have matches the participants, interventions, comparator and outcome of our original question. Directness is a similar concept to 'generalizability', or 'applicability' of the results.

A high quality systematic review will present the important characteristics of each individual trial in a table, allowing the reader to assess to whom, and where, the results may be applied.

Study ID	Year	Country	Participants	Artemisinin (Dose and Route)	Quinine (Dose and Route)
Anh 1989	1989	Vietnam	Adults (> 16 years)	60 mg i.v. Then at 4, 24, 48 hours	20 mg loading dose i.v. Then 10 mg/kg i.v. every 8 hours
Anh 1995	1995	Vietnam	Adults (> 15 years)	60 mg i.v. Then at 4, 24, 48 hours	20 mg loading dose i.v. Then 10 mg/kg i.v. every 8 hours
Cao 1997	1995	Vietnam	Children (< 15 years)	3 mg/kg i.m. Then 2 mg/kg IM at 12, 24, 48, 72 hours	20 mg/kg loading dose i.v. Then 10 mg/kg i.v. every 8 hours
Dondorp 2005	2005	Bangaldesh, Myanmar, India, Indonesia	Adults and children (> 2 years)	2.4 mg/kg i.v Then at 12, 24 hours	20 mg/kg loading dose i.v. Then 10 mg/kg i.v. every 8 hours
Hien 1992	1990	Vietnam	Adults (not specified)	60 mg i.v. Then at 4, 24, 48 hours	500 mg i.v. Then 500 mg i.v. every 8 hours
Newton 2003	2001	Thailand	Adults (> 15 years)	2.4 mg/kg i.v. Then 1.2 mg/kg at 12, 24 hours	20 mg/kg loading dose i.v. Then 10 mg/kg i.v every 8 hours

Questions to consider

1. Are these the participants we are interested in? (age, country)
2. Is it the right intervention?
3. Is it the right comparator?
4. Are there any other analyses you would like to see? (Sub-groups)

Overall: Is there any serious indirectness that would decrease your confidence that artesunate reduces mortality compared to quinine in adults or children in sub-Saharan Africa?

Your judgement:

- Don't downgrade
- Downgrade by 1 for serious indirectness
- Downgrade by 2 for very serious indirectness

GRADE levels of 'Quality'		
4	High	RCTs start as high quality evidence and are downgraded when there are **serious** problems which decrease your confidence in the results of the systematic review
3	Moderate	
2	Low	
1	Very Low	

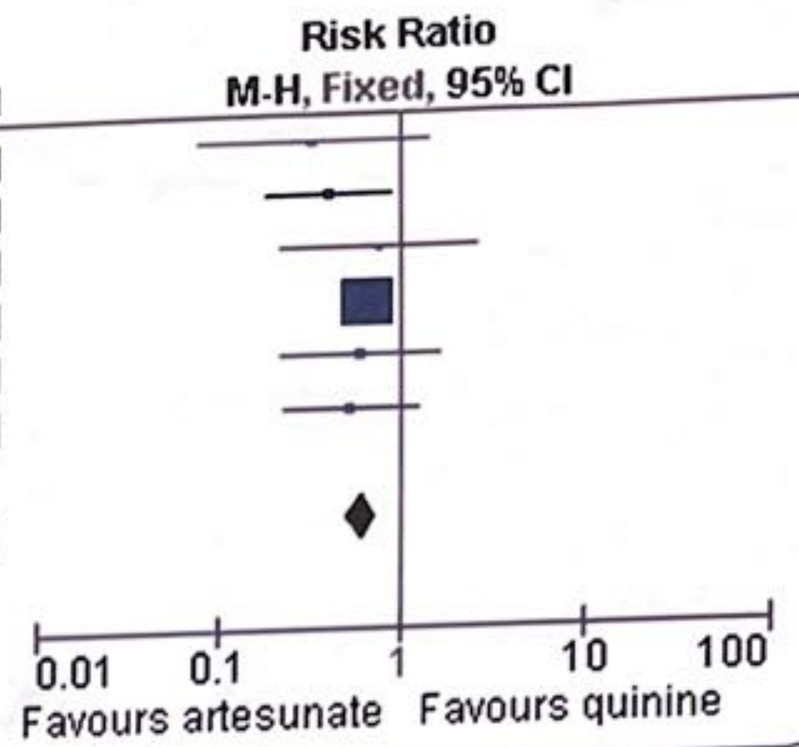
The South African Cochrane Centre
effective health care RESEARCH CONSORTIUM
Centre for Evidence Based Health Care
LSTM LIVERPOOL SCHOOL OF TROPICAL MEDICINE

Artesunate compared to Quinine for treating adults and children with severe malaria

Patient or population: Adults and children with severe malaria
Settings: Malaria endemic settings in sub-Saharan Africa
Intervention: Artesunate 2.4 mg/kg repeated at 12, 24 and then every 24 hours
Comparison: Quinine 20 mg/kg loading dose then 10 mg/kg every 8 hours

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: $Z = 4.82$ ($P < 0.00001$)



GRADE Assessment: Publication bias

Publication bias occurs when trials showing positive results are published (because they are newsworthy), and trials showing no effect, or negative effects, are not published (either because they are uninteresting, or because the findings don't support the interests of the researchers or drug companies).

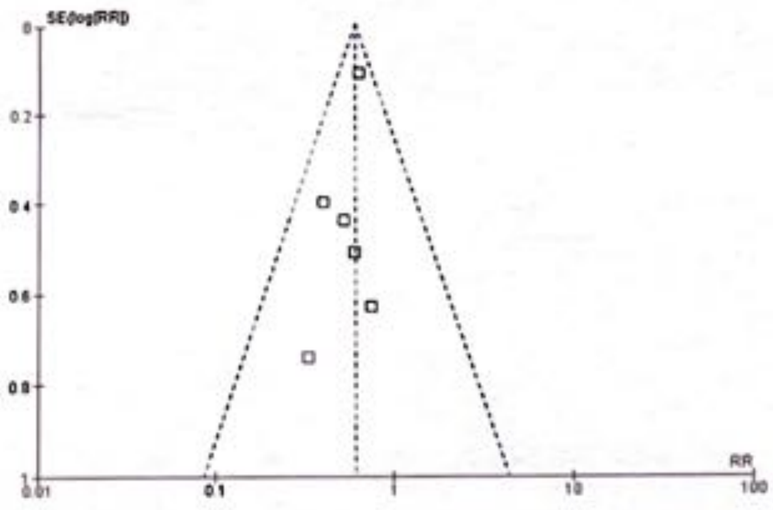
High quality systematic reviews will attempt to reduce the risk of publication bias by searching trial registers, contacting leading researchers in the field, and contacting drug companies, in order to find and include unpublished studies.

High quality systematic reviews will also look for evidence of publication bias by creating **funnel plots**. Each dot in the funnel plot represents a study. Small studies appear at the bottom of the plot, and due to their small sample size may be very inaccurate and should vary widely around the true effect. Large studies appear towards the tip of the plot and should be close to the true effect.

If small studies finding no effect, or harmful effects, are not published, there will be asymmetry in the plot, and the small published studies showing positive effects will lead to an overestimation of the effect.

Questions to consider

1. Are there enough trials to assess funnel plot asymmetry?
2. Is there asymmetry in the funnel plot?
3. Are small trials with large positive effects leading to an overestimate of the effect?



Overall: Is there enough evidence of publication bias to decrease your confidence in the effect size?

Your judgement:

- Don't downgrade
- Downgrade by 1 for serious risk of publication bias

GRADE levels of 'Quality'		
4	High	RCTs start as high quality evidence and are downgraded when there are **serious** problems which decrease your confidence in the results of the systematic review
3	Moderate	
2	Low	
1	Very Low	

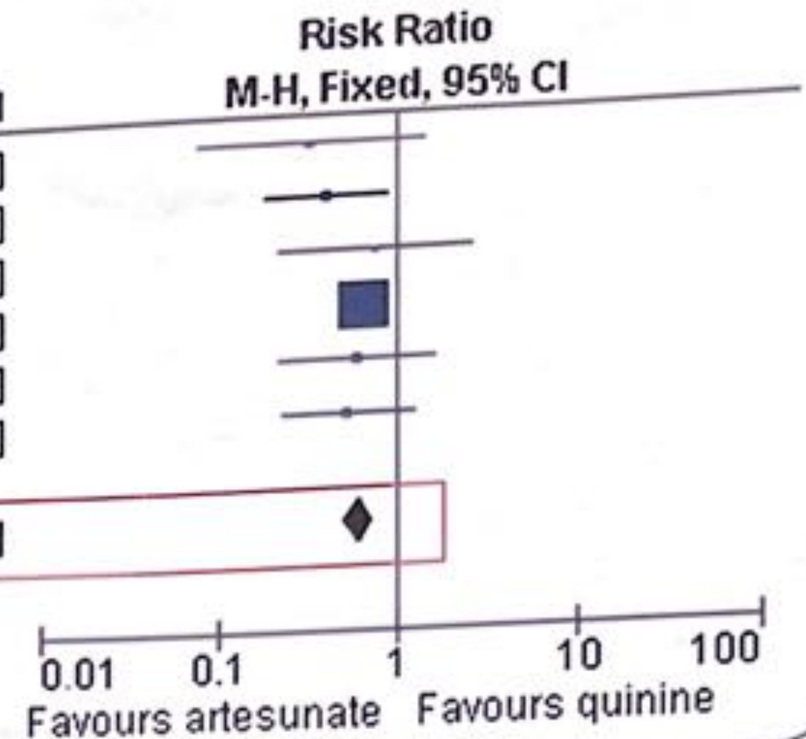
The South African Cochrane Centre
effective health care RESEARCH CONSORTIUM
Centre for Evidence Based Health Care
LSTM LIVERPOOL SCHOOL OF TROPICAL MEDICINE

Artesunate compared to Quinine for treating adults and children with severe malaria

Patient or population: Adults and children with severe malaria
Settings: Malaria endemic settings in sub-Saharan Africa
Intervention: Artesunate 2.4 mg/kg repeated at 12, 24 and then every 24 hours
Comparison: Quinine 20 mg/kg loading dose then 10 mg/kg every 8 hours

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			

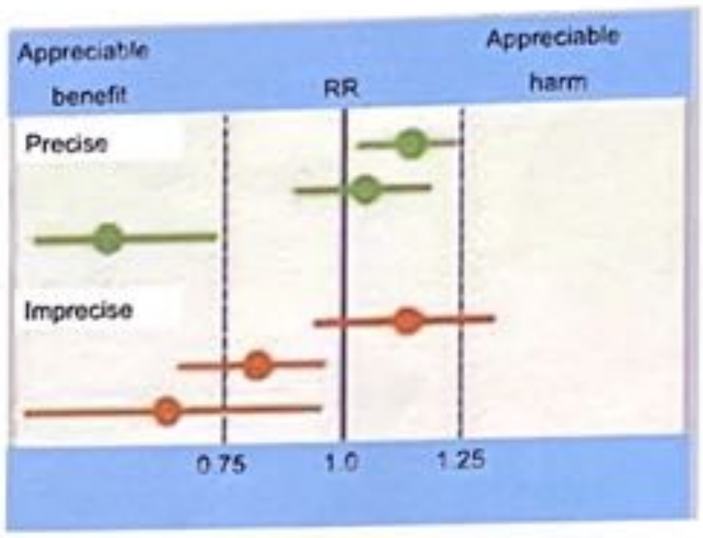
Heterogeneity: $Chi^2 = 2.26$, $df = 5$ ($P = 0.81$); $I^2 = 0\%$
Test for overall effect: $Z = 4.82$ ($P < 0.00001$)



GRADE Assessment: Precision

The results of individual trials and meta-analyses are usually presented with a 95% confidence interval (CI), which gives a measure of the statistical precision of the result. If the 95% CI of the result doesn't cross the line of no effect (RR = 1), then the result is said to be 'statistically significant'. However, guideline groups also have to decide if the result is 'clinically important' i.e. of an 'appreciable size'.

The two limits of the 95% CI, the 'best' and 'worst' case scenarios, can be used to guide decisions on 'precision'. If both limits of the 95% CI lie within what would be considered to be an 'appreciable benefit', then the result is a 'precise result' of a clinically important effect.



When assessing precision, it is also important to consider whether the trials/meta-analysis are adequately powered to detect the observed effect (sometimes small trials will find large effects just by chance). To do this an 'optimal information size' can be calculated, which is the same as a sample size calculation for an individual trial. To confidently detect a 25% reduction in mortality, in a setting where the risk of mortality after treatment with quinine is 214/963 (22%), would require 796 participants in each treatment arm.

Questions to consider

1. Is the result statistically significant?
2. Look at the two limits of the 95% CI. What is the best case scenario? What is the worst case scenario? Would these both be considered clinically important effects?
3. Which studies are adequately powered to detect a 25% reduction in mortality? Is the overall meta-analysis adequately powered?

Overall: Is there serious imprecision which would decrease your confidence that artesunate reduces mortality compared to quinine?

Your judgement:

- Don't downgrade
- Downgrade by 1 for serious imprecision
- Downgrade by 2 for very serious imprecision

GRADE levels of 'Quality'		
4	High	RCTs start as high quality evidence and are downgraded when there are **serious** problems which decrease your confidence in the results of the systematic review
3	Moderate	
2	Low	
1	Very Low	

The South African Cochrane Centre
effective health care RESEARCH CONSORTIUM
Centre for Evidence Based Health Care
LSTM