CINeMA

Confidence in Network Meta-Analysis

cinema.ispm.unibe.ch

University of Bern

Presenters: Georgia Salanti, Virginia Chiocchia, Adriani Nikolakopoulou
Programmer: Theodore Papakonstantinou
Contributors: Julian Higgins, Anna Chaimani, Matthias Egger, Cinzia Del Giovane

CINeMA framework

Comparison	Number of Studies	Within-study bias	Reporting bias	Indirectness	Imprecision	Heterogeneity	Incoherence	Confidence rating
Mixed evidence								
ACE vs BBlocker	3	Some concerns	Undetected	No concerns	No concerns	No concerns	Some concerns	High
ACE vs CCB	3	No concerns	Undetected	No concerns	No concerns	Some concerns	No concerns	High
ACE vs Diuretic	2	No concerns	Undetected	No concerns	No concerns	No concerns	No concerns	High
ACE vs Placebo	3	No					concerns	High
ARB vs BBlocker	1	No					concerns	High
ARB vs CCB	1	Som					concerns	High
ARB vs Diuretic	1	No					concerns	High
ARB vs Placebo	2	No					concerns	High
BBlocker vs CCB	5	No					concerns	High
BBlocker vs Diuretic	2	No					concerns	High
BBlocker vs Placebo	1	No concerns	Suspected	No concerns	No concerns	Major concerns	Some concerns	High
CCB vs Diuretic	2	No concerns	Undetected	No concerns	No concerns	Some concerns	No concerns	High
CCB vs Placebo	1	No concerns	Suspected	No concerns	Major concerns	Major concerns	No concerns	High
Diuretic vs Placebo	3	No concerns	Suspected	No concerns	No concerns	Some concerns	No concerns	High
Indirect evidence								
ACE vs ARB	--	No concerns	Undetected	No concerns	Major concerns	Some concerns	No concerns	High

Process

Explicit rules that classify each network meta-analysis effect for each domain to
No concerns, Some concerns, Major concerns
as described in the documentation

The rules can be overwritten!

Incident diabetes in clinical trials of antihypertensive drugs: a network meta-analysis

William J Elliott, Peter M Meyer

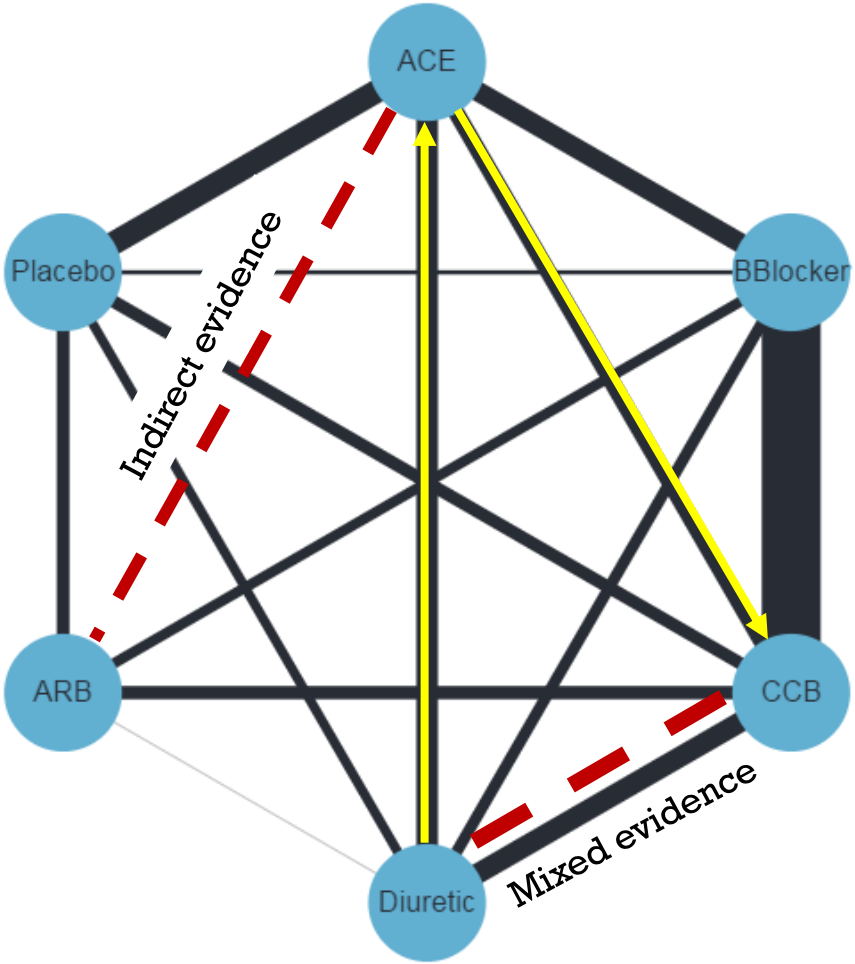
Summary

Background The effect of different classes of antihypertensive drugs on incident diabetes mellitus is controversial because traditional meta-analyses are hindered by heterogeneity across trials and the absence of trials comparing angiotensin-converting-enzyme (ACE) inhibitors with angiotensin-receptor blockers (ARB). We therefore undertook a network meta-analysis, which accounts for both direct and indirect comparisons to assess the effects of antihypertensive agents on incident diabetes.

Lancet 2007; 369: 201–07

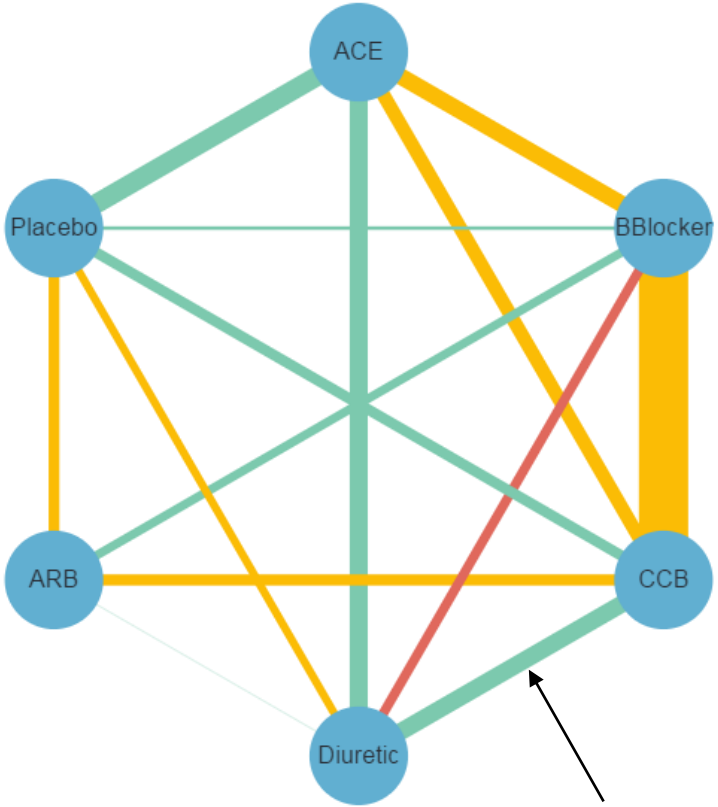
Department of Preventive Medicine, Rush Medical College of Rush University at Rush University Medical Center, Chicago, IL 60612, USA

Number of studies	22
Number of treatment nodes	6
Primary outcome	Effect of antihypertensives on incidence diabetes mellitus - proportion of patients who developed diabetes
Measurement	Binary
Intervention comparison type	pharmacological vs placebo


ACE
Placebo
BBlocker
ARB
CCB
Diuretic
Indirect evidence
Mixed evidence

WITHIN-STUDY BIAS

- ❑ Major concerns
- ❑ Some concerns
- ❑ No concerns

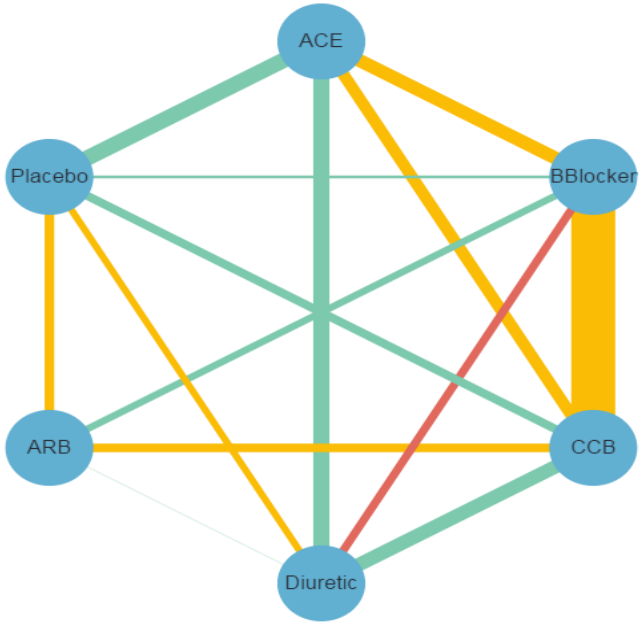


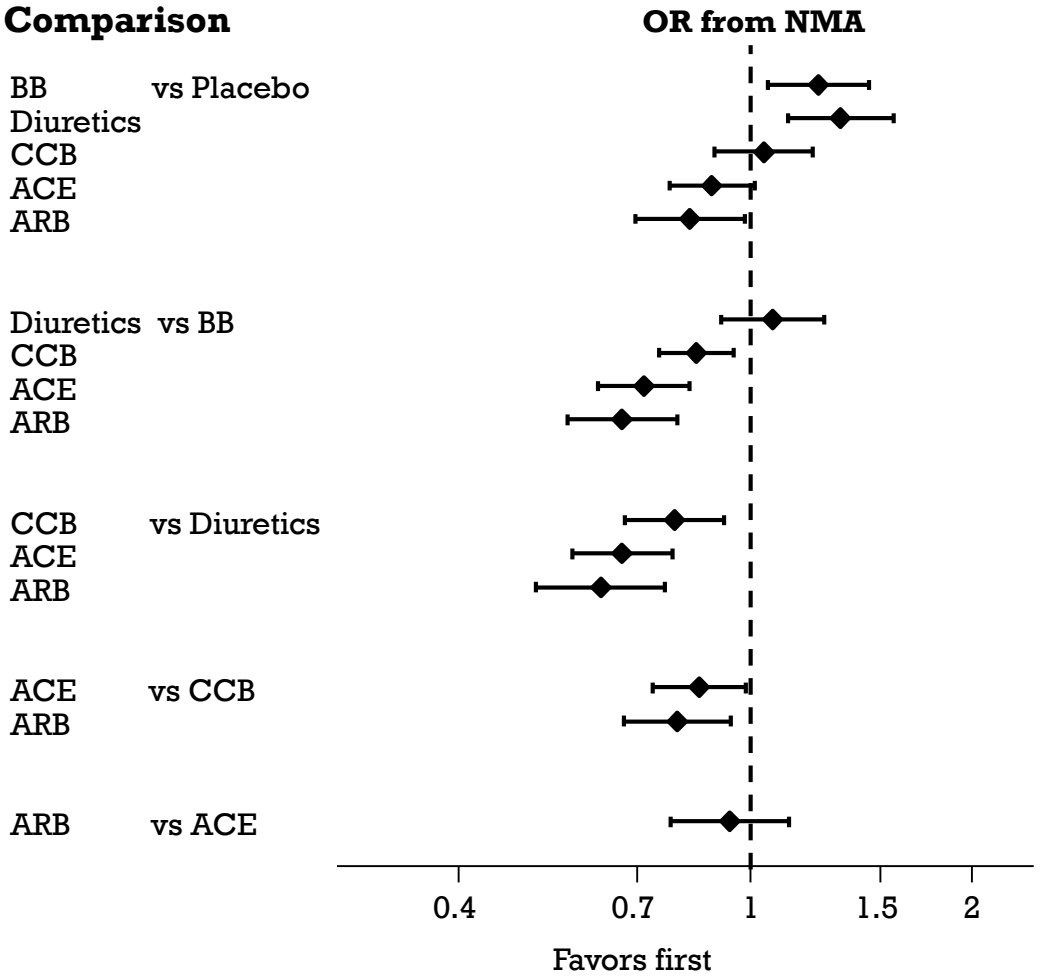
Plot direct comparison in green

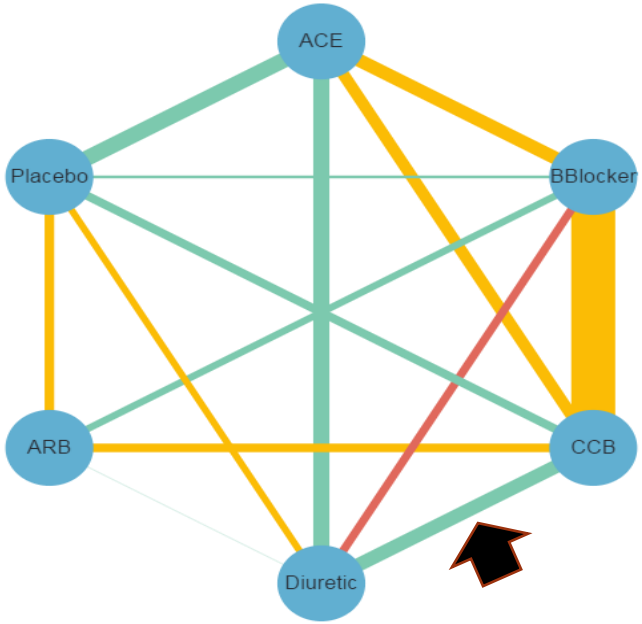
CCB vs Diuretics: overall low risk of bias

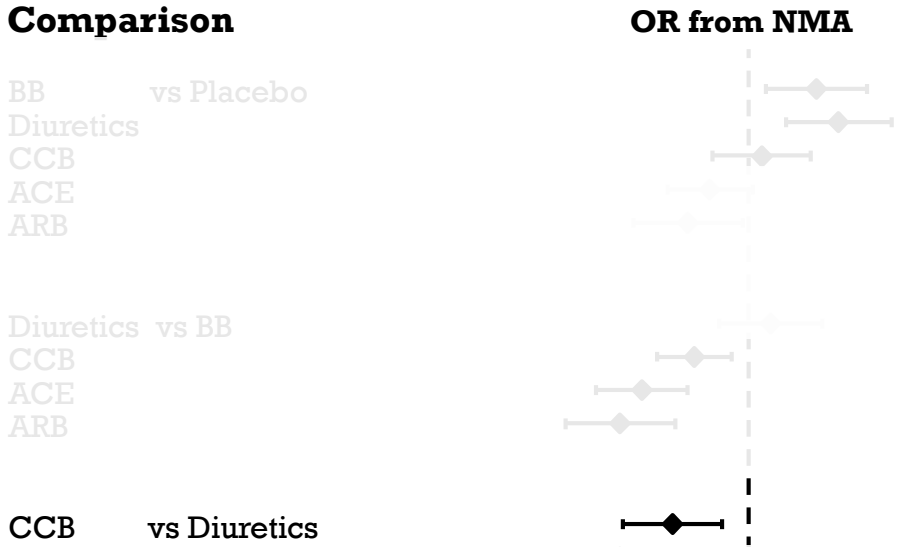
Form risk of bias judgements for each study. Consider selection, performance, attrition, detection and reporting bias

Study name	Risk of Bias
AASK	LOW
ALLHAT	LOW
ALPINE	LOW
ANBP-2	LOW
ASCOT	LOW
CAPPP	MODERATE
CHARM	LOW
DREAM	LOW
EWPHE	MODERATE
FEVER	LOW
HAPPHY	HIGH
HOPE	LOW
INSIGHT	LOW
INVEST	LOW
LIFE	LOW
MRC	LOW
NORDIL	LOW
PEACE	LOW
SCOPE	MODERATE
SHEP	LOW
STOP-2	MODERATE
VALUE	MODERATE


ACE
Placebo
BBlocker
ARB
CCB
Diuretic

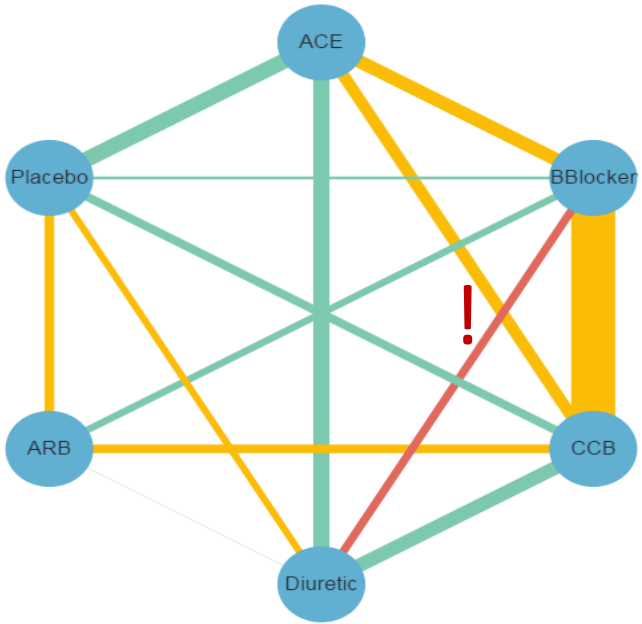

Comparison
OR from NMA
BB vs Placebo
Diuretics
CCB
ACE
ARB
Diuretics vs BB
CCB
ACE
ARB
CCB vs Diuretics
ACE
ARB
ACE vs CCB
ARB
ARB vs ACE
0.4
0.7
1
1.5
2
Favors first



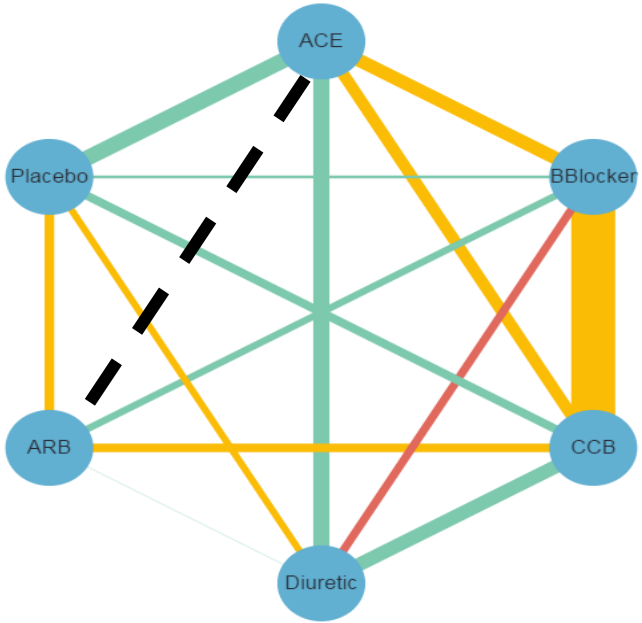


What is your judgement about within-study bias for this (<u>mixed</u>) OR between CCB vs Diuretics estimated in network meta-analysis?

- ❑ Major concerns
- ❑ Some concerns
- ❑ No concerns



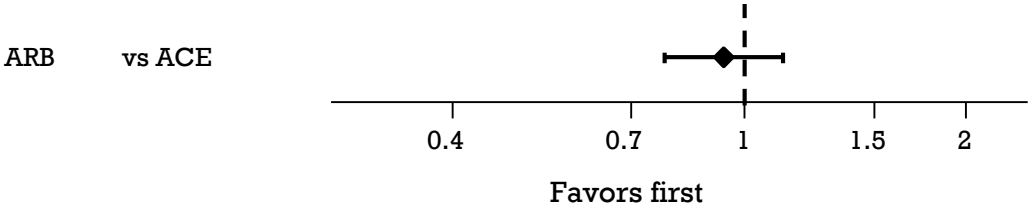
Studies with high risk of bias contribute to the estimation of the OR CCB vs Diuretics!



Comparison **OR from NMA**

What is your judgement about study limitations for this (indirect) OR for ACE vs ARB estimated in NMA?

- ❑ Major concerns
- ❑ Some concerns
- ❑ No concerns



An indirect or mixed treatment effect is a combination of the available direct treatment effects

The contribution matrix

	Study 1	Study 2	Study 3	Study 4	Study 5	Study 6	Study 7	Study 8	Study 9	Study 10	Study 11	Study 12	Study 13	.....
Mixed estimates														
ACE:BBlocker	10	9	0	4	4	25	2	3	0	2	4	2	1	4
ACE:CCB	9	23	0	4	4	8	2	3	0	5	0	2	4	4
ACE:Diuretic	3	28	0	21	0	5	0	4	2	1	5	3	5	0
ACE:Placebo	2	6	0	4	0	3	2	23	1	5	0	15	0	0
ARB:BBlocker	2	0	0	0	5	3	6	2	0	1	2	1	0	5
ARB:CCB	1	3	0	0	4	0	7	2	0	5	0	1	2	4
ARB:Diuretic	1	12	1	4	0	1	10	2	2	0	6	1	8	0
ARB:Placebo	1	3	0	0	0	2	29	3	1	5	1	2	1	0
BBlocker:CCB	6	5	0	0	19	4	0	0	0	2	3	0	2	19
BBlocker:Diuretic	3	14	0	7	5	7	1	0	1	1	17	0	8	5
BBlocker:Placebo	4	3	0	0	4	8	5	7	2	8	4	4	1	4
CCB:Diuretic	2	30	0	6	3	1	1	0	1	4	6	0	20	3
CCB:Placebo	3	9	0	0	3	2	5	6	2	20	1	4	4	3
Diuretic:Placebo	0	12	0	7	0	1	2	6	7	6	3	4	5	0
Indirect estimates														
ACE:ARB	4	8	0	3	0	7	11	7	0	0	1	5	1	0

Papakonstantinou T, Nikolakopoulou A, Rücker G *et al.* Estimating the contribution of studies in network meta-analysis: paths, flows and streams [version 1]. *F1000Research* 2018, 7:610

An indirect or mixed treatment effect is a combination of the available direct treatment effects

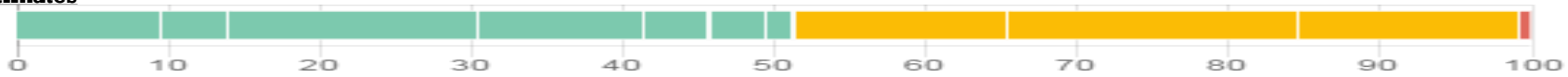
The contribution matrix

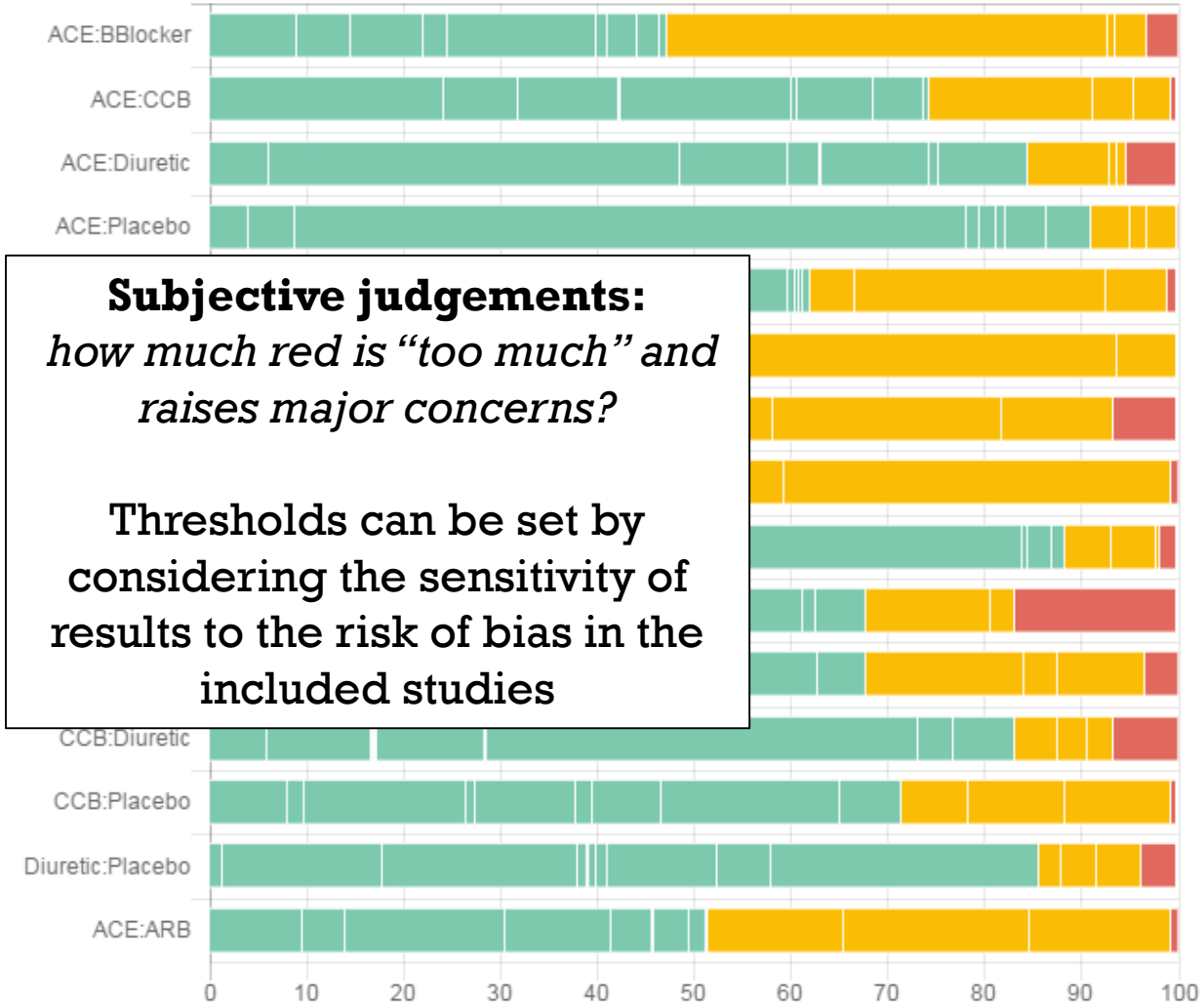
	Study 1	Study 2	Study 3	Study 4	Study 5	Study 6	Study 7	Study 8	Study 9	Study 10	Study 11	Study 12	Study 13	.....
Mixed estimates														
ACE:BBlocker	10	9	0	4	4	25	2	3	0	2	4	2	1	4
ACE:CCB	9	23	0	4	4	8	2	3	0	5	0	2	4	4
ACE:Diuretic	3	28	0	21	0	5	0	4	2	1	5	3	5	0
ACE:Placebo	2	6	0	4	0	3	2	23	1	5	0	15	0	0
ARB:BBlocker	2	0	0	0	5	3	6	2	0	1	2	1	0	5
ARB:CCB	1	3	0	0	4	0	7	2	0	5	0	1	2	4
ARB:Diuretic	1	12	1	4	0	1	10	2	2	0	6	1	8	0
ARB:Placebo	1	3	0	0	0	2	29	3	1	5	1	2	1	0
BBlocker:CCB	6	5	0	0	19	4	0	0	0	2	3	0	2	19
BBlocker:Diuretic	3	14	0	7	5	7	1	0	1	1	17	0	8	5
BBlocker:Placebo	4	3	0	0	4	8	5	7	2	8	4	4	1	4
CCB:Diuretic	2	30	0	6	3	1	1	0	1	4	6	0	20	3
CCB:Placebo	3	9	0	0	3	2	5	6	2	20	1	4	4	3
Diuretic:Placebo	0	12	0	7	0	1	2	6	7	6	3	4	5	0
Indirect estimates														
ACE:ARB	4	8	3											

An indirect or mixed treatment effect is a combination of the available direct treatment effects

The contribution matrix

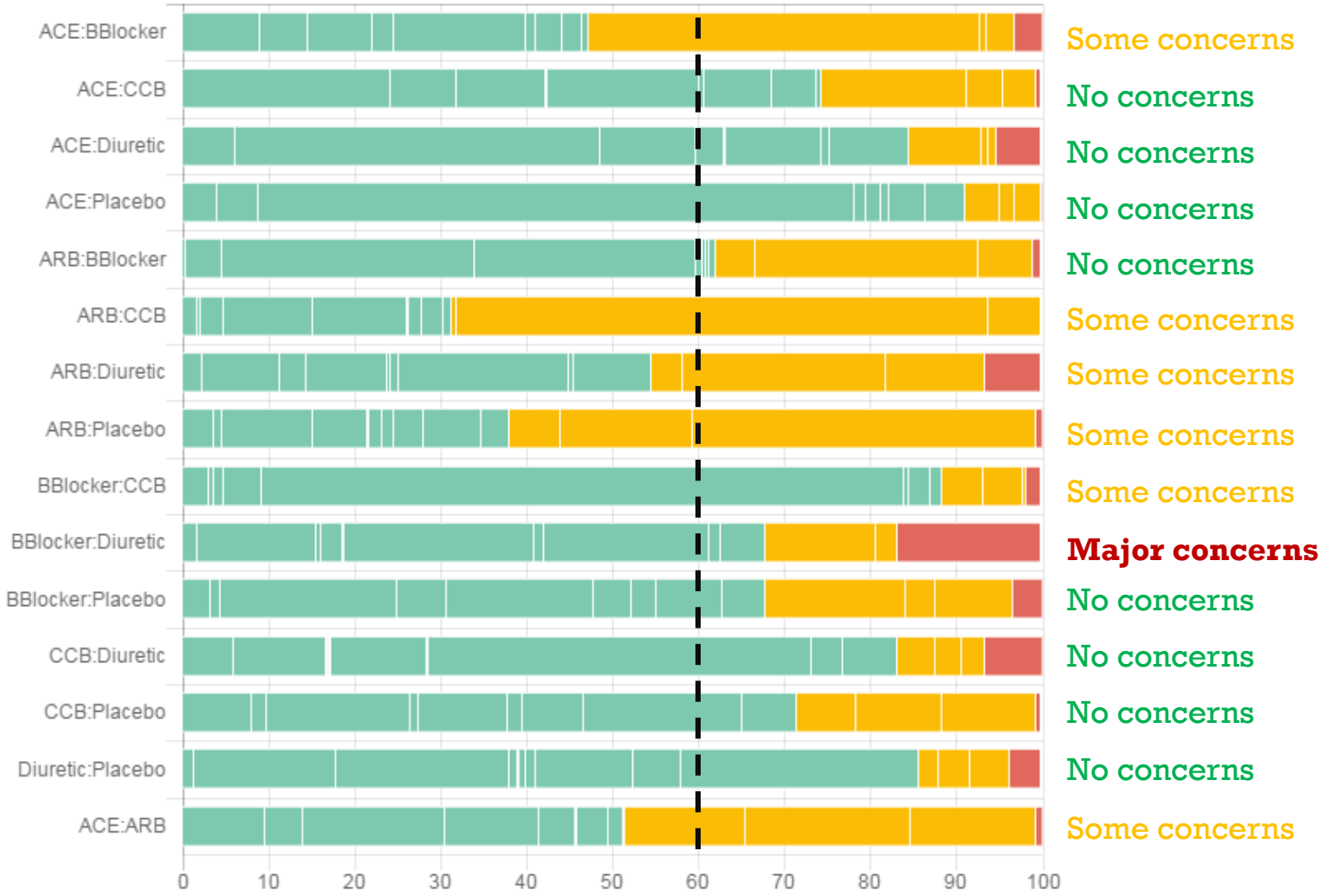
	Study 1	Study 2	Study 3	Study 4	Study 5	Study 6	Study 7	Study 8	Study 9	Study 10	Study 11	Study 12	Study 13	.....
Mixed estimates														
ACE:BBlocker	10	9	0	4	4	25	2	3	0	2	4	2	1	4
ACE:CCB	9	23	0	4	4	8	2	3	0	5	0	2	4	4
ACE:Diuretic	3	28	0	21	0	5	0	4	2	1	5	3	5	0
ACE:Placebo	2	6	0	4	0	3	2	23	1	5	0	15	0	0
ARB:BBlocker	2	0	0	0	5	3	6	2	0	1	2	1	0	5
ARB:CCB	1	3	0	0	4	0	7	2	0	5	0	1	2	4
ARB:Diuretic	1	12	1	4	0	1	10	2	2	0	6	1	8	0
ARB:Placebo	1	3	0	0	0	2	29	3	1	5	1	2	1	0
BBlocker:CCB	6	5	0	0	19	4	0	0	0	2	3	0	2	19
BBlocker:Diuretic	3	14	0	7	5	7	1	0	1	1	17	0	8	5
BBlocker:Placebo	4	3	0	0	4	8	5	7	2	8	4	4	1	4
CCB:Diuretic	2	30	0	6	3	1	1	0	1	4	6	0	20	3
CCB:Placebo	3	9	0	0	3	2	5	6	2	20	1	4	4	3
Diuretic:Placebo	0	12	0	7	0	1	2	6	7	6	3	4	5	0
Indirect estimates														
ACE:ARB														





Subjective judgements:
how much red is "too much" and raises major concerns?

Thresholds can be set by considering the sensitivity of results to the risk of bias in the included studies


ACE:BBlocker
ACE:CCB
ACE:Diuretic
ACE:Placebo
ARB:BBlocker
ARB:CCB
ARB:Diuretic
ARB:Placebo
BBlocker:CCB
BBlocker:Diuretic
BBlocker:Placebo
CCB:Diuretic
CCB:Placebo
Diuretic:Placebo
ACE:ARB
0
10
20
30
40
50
60
70
80
90
100
Some concerns
No concerns
No concerns
No concerns
No concerns
Some concerns
Some concerns
Some concerns
Some concerns
Major concerns
No concerns
No concerns
No concerns
No concerns
Some concerns

INDIRECTNESS

- ❑ Major concerns
- ❑ Some concerns
- ❑ No concerns

The idea is to evaluate the confidence intervals and the prediction intervals against the spectrum of values relevant to decision-making.

INDIRECTNESS

- Considerations similar to those in a pairwise meta-analysis
- **How relevant is the study PICO and setting to the research question?**
- **Score each study at 3 levels**
  - Low indirectness to the research question
  - Moderate indirectness to the research question
  - High indirectness to the research question
- Then study-level judgements are summarized within pairwise comparisons and across the network using the contribution matrix exactly as with the Risk of Bias.
- This also addresses the condition of transitivity!
  - If the studies across comparisons have differences in important characteristics (e.g. effect modifiers) compared to the target population, then the transitivity assumption is challenged

Now it is time for….

CINeMA

IMPRECISION

- ❑ Major concerns
- ❑ Some concerns
- ❑ No concerns

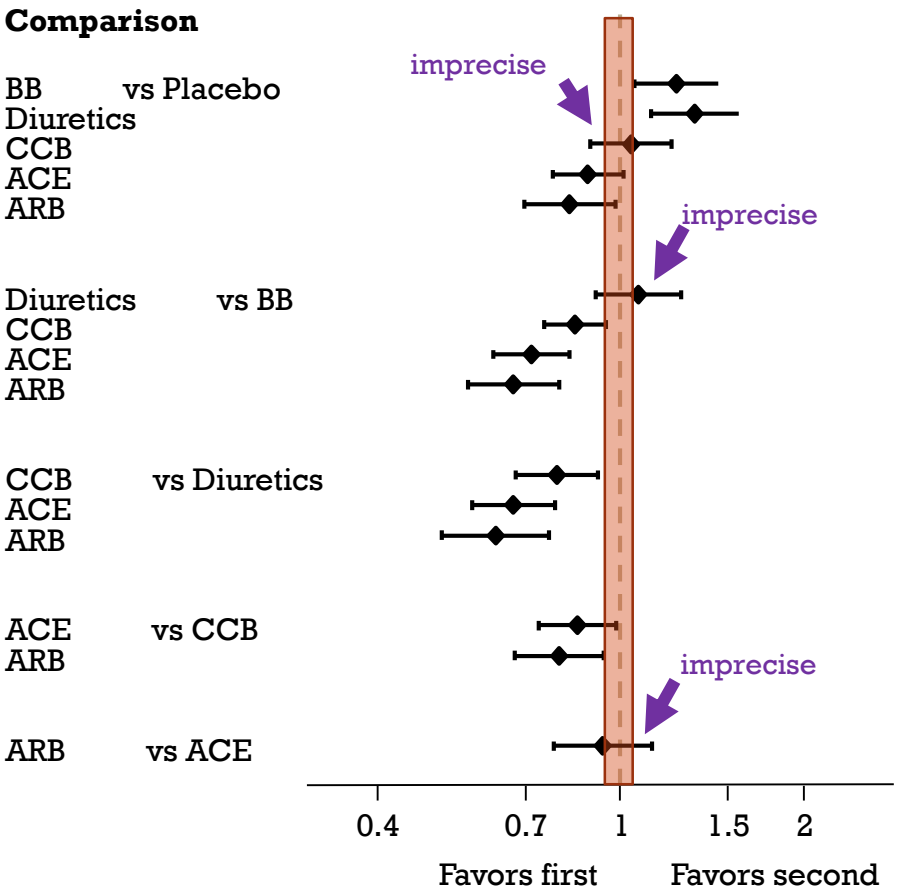
IMPRECISION

- Traditional GRADE considers, among others, the total sample size available and compares it with the Optimal Information Size
- The sample size in a NMA relative effect makes little sense (as studies in the network contribute direct and indirect information!)
- Imprecision relates to the width of the 95% confidence interval:

Does the 95% CI include values that lead to different clinical decisions?

- Set a ”margin of equivalence”
  - the range of relative treatment effect around the no-effect line that do not signify important differences between the interventions

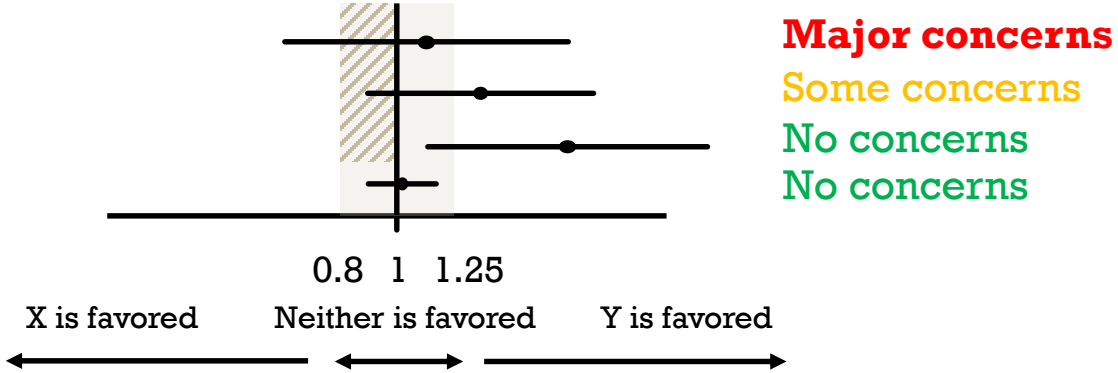
NMA estimated odds ratios for diabetes



Imprecision:
Confidence intervals include values that lead into different clinical decisions

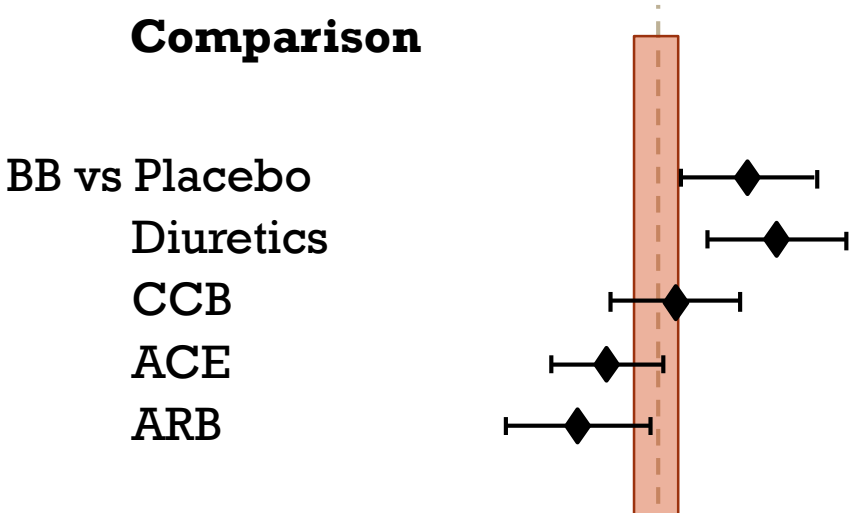
Margin of equivalence:
OR=1.05 in either direction
Imprecision when the confidence interval **crosses both 0.95 and 1.05**

IMPRECISION



Compare the 95% confidence interval with a **subset of the range of equivalence,** the range between the no effect line and the edge of the range of equivalence that is in the direction opposite to the observed point estimate.

NMA estimated odds ratios for diabetes

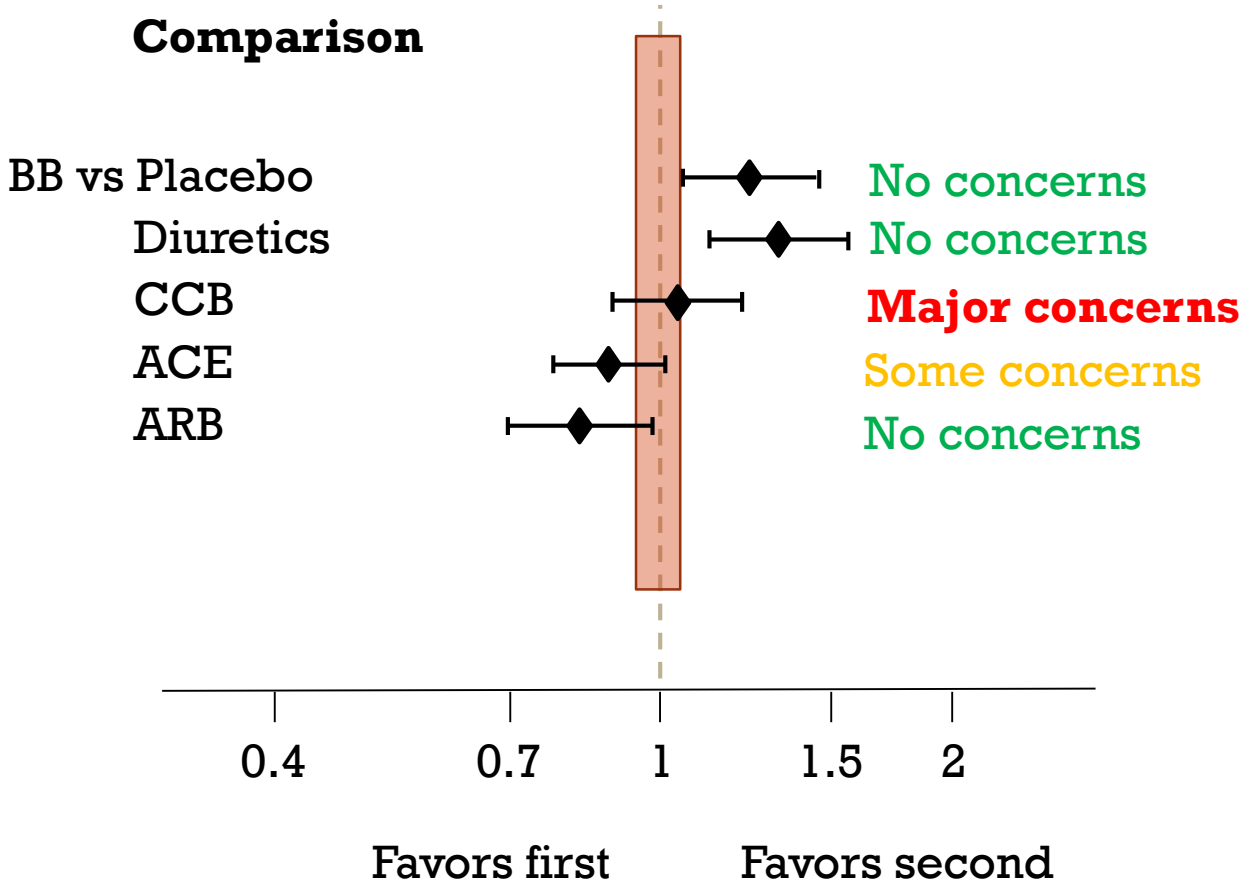


For which comparison do you have major concerns about imprecision?

a) BB vs CCB
b) BB vs ACE
c) BB vs ARB

Favors first Favors second

NMA estimated odds ratios for diabetes



Now it is time for….

CINeMA

VARIABILITY BEYOND CHANCE

Heterogeneity

between-study variance within a comparison

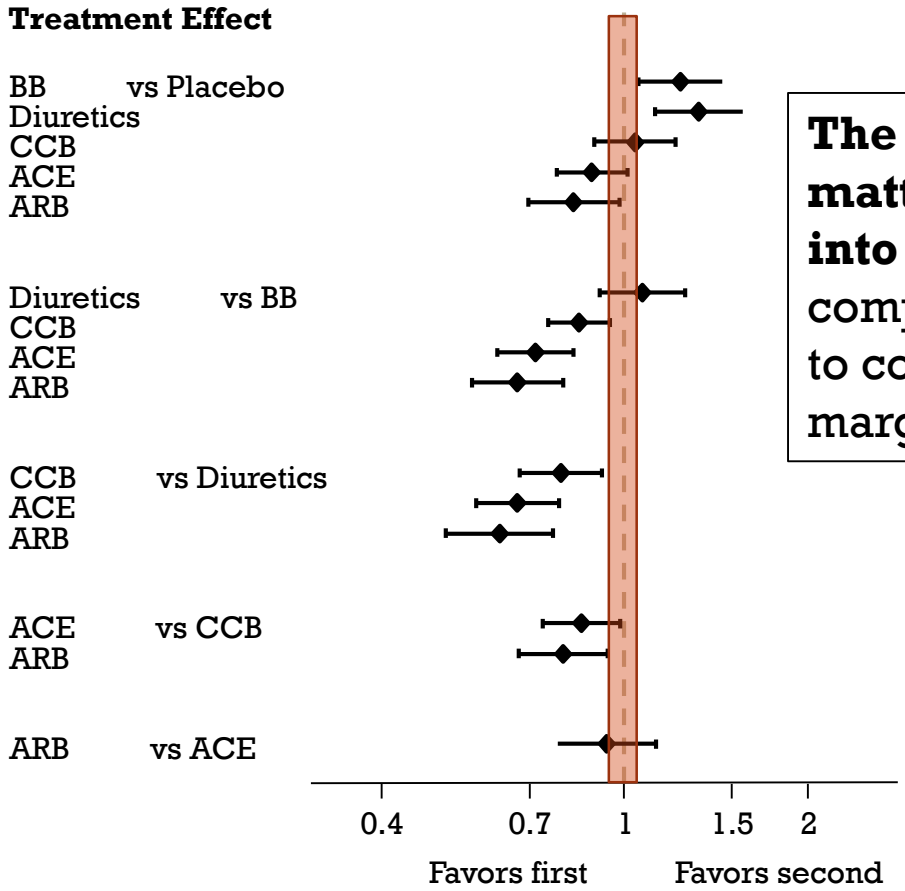
- ❑ Major concerns
- ❑ Some concerns
- ❑ No concerns

- ❑ Major concerns
- ❑ Some concerns
- ❑ No concerns

HETEROGENEITY

- The major driver in judging heterogeneity is whether it impacts on clinical decisions
- Heterogeneity is represented by the **predictive intervals**: the intervals within which we expect to find the true effect size of a new study
- They are extensions of the confidence intervals

HETEROGENEITY



The amount of heterogeneity matters only when it leads into different conclusions: compare prediction intervals to confidence intervals and the margin of equivalence.

HETEROGENEITY

Prediction interval:
Where is the true effect in a new study?

Heterogeneity changes conclusions!

Treatment Effect

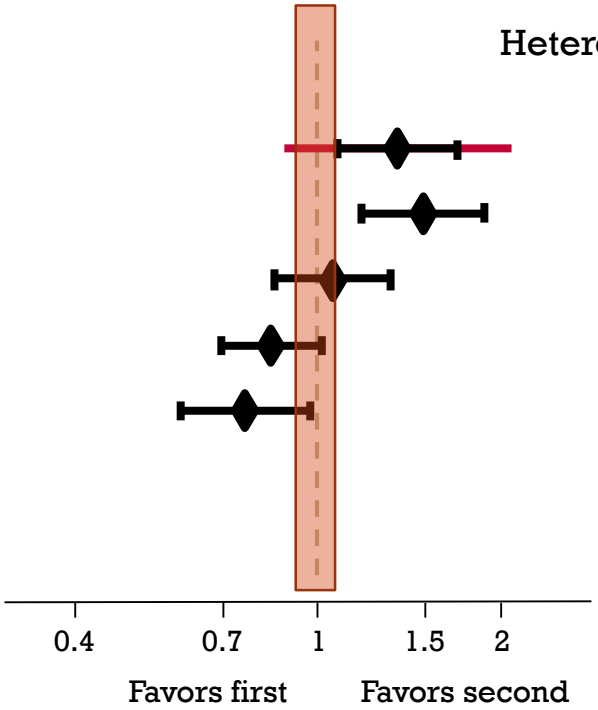
BB vs Placebo

Diuretics

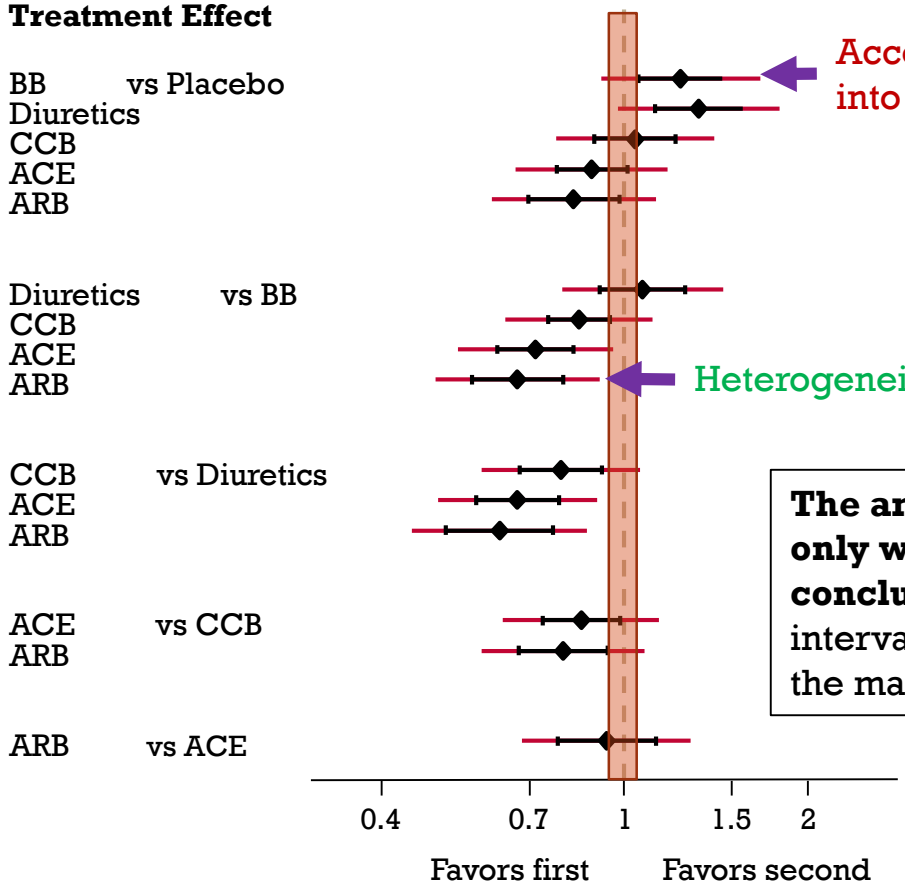
CCB

ACE

ARB



HETEROGENEITY



Accounting for heterogeneity leads into different clinical decisions!

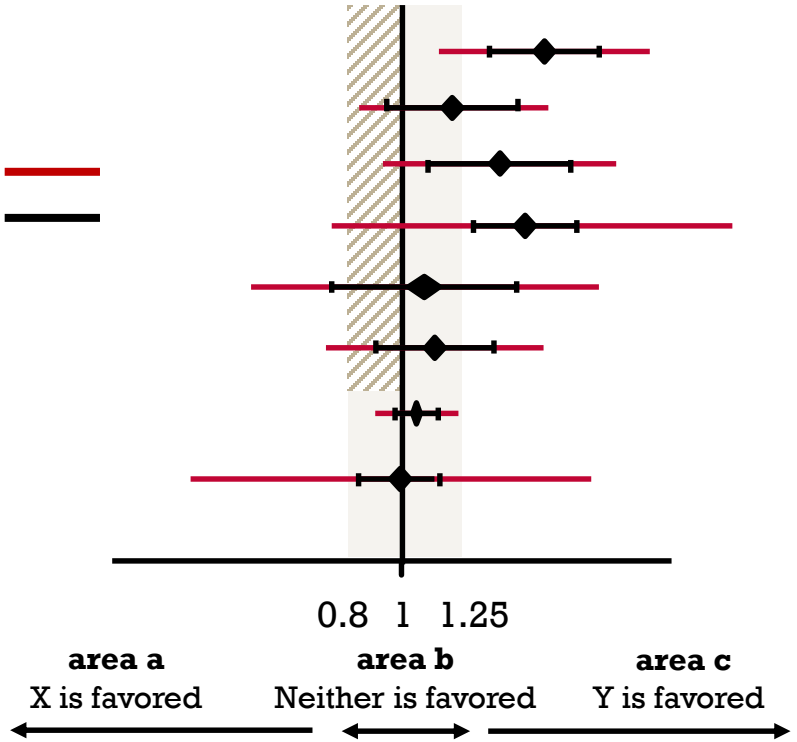
Heterogeneity does not changes conclusions!

The amount of heterogeneity matters only when it leads into different conclusions: compare prediction intervals to confidence intervals and the margin of equivalence.

HETEROGENEITY

Rules implemented in the software

Prediction interval
Confidence interval



No concerns
No concerns
Some concerns
Major concerns
No concerns
Some concerns
No concerns
Major concerns

HETEROGENEITY

- The major driver or our decisions is whether the heterogeneity impacts on clinical decisions
- Heterogeneity is represented by the **predictive intervals**: the intervals within which we expect to find the true effect size of a new study
- They are extensions of the confidence intervals
- Pairwise meta-analysis heterogeneity variances τ^2 can be estimated
  - But their estimation makes sense when you have enough studies
  - The observed values of τ^2 can be compared with the expected values from empirical evidence *(Turner et al Int J Epidemiol. 2012, Rhodes et al. J Clin Epidemiol. 2015)*
  - The expected values depend on the nature of the outcome and the treatments being compared

VARIABILITY BEYOND CHANCE

Heterogeneity
between-study variance within a comparison

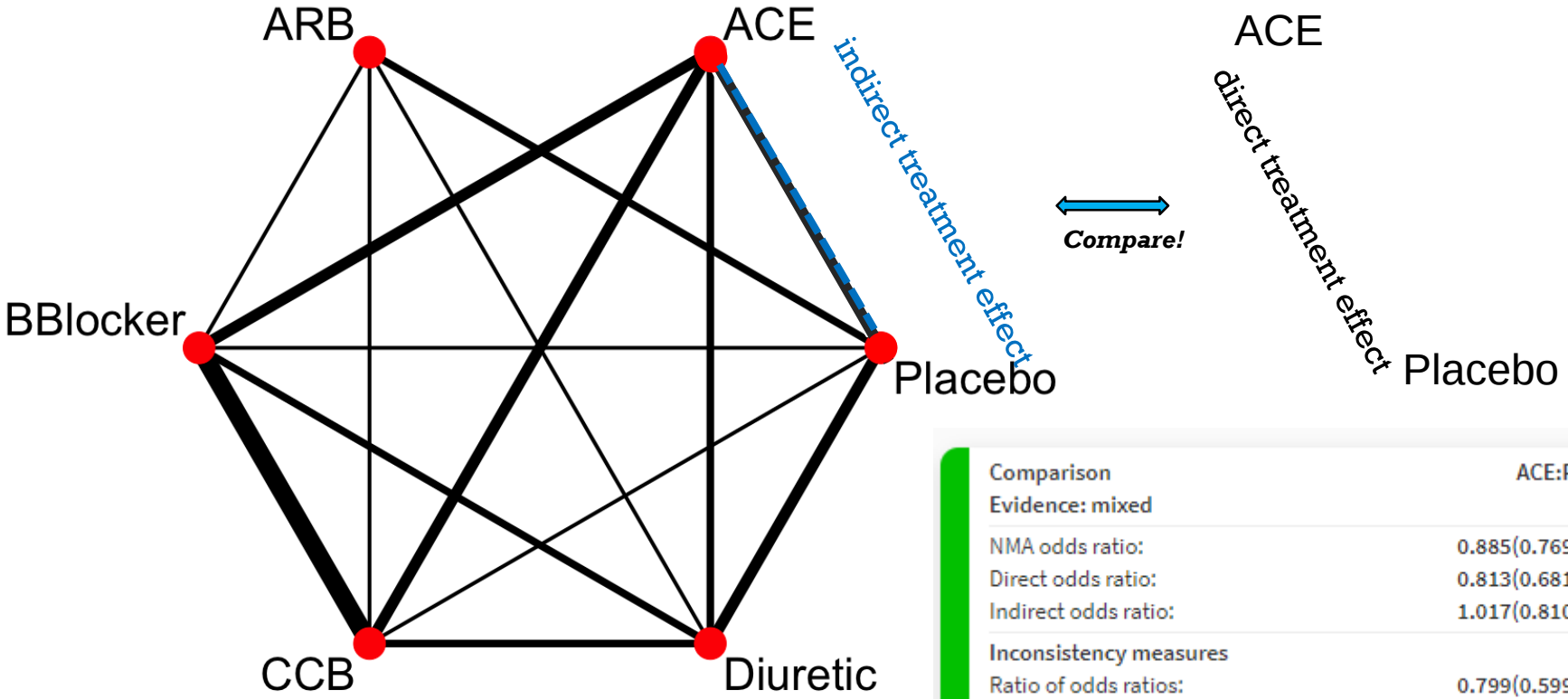
We consider prediction intervals for the **impact of heterogeneity** in clinical decision making

Incoherence
disagreement between different sources of evidence

We consider **how serious is the disagreement** between direct and indirect evidence with respect to clinical decision making

INCOHERENCE

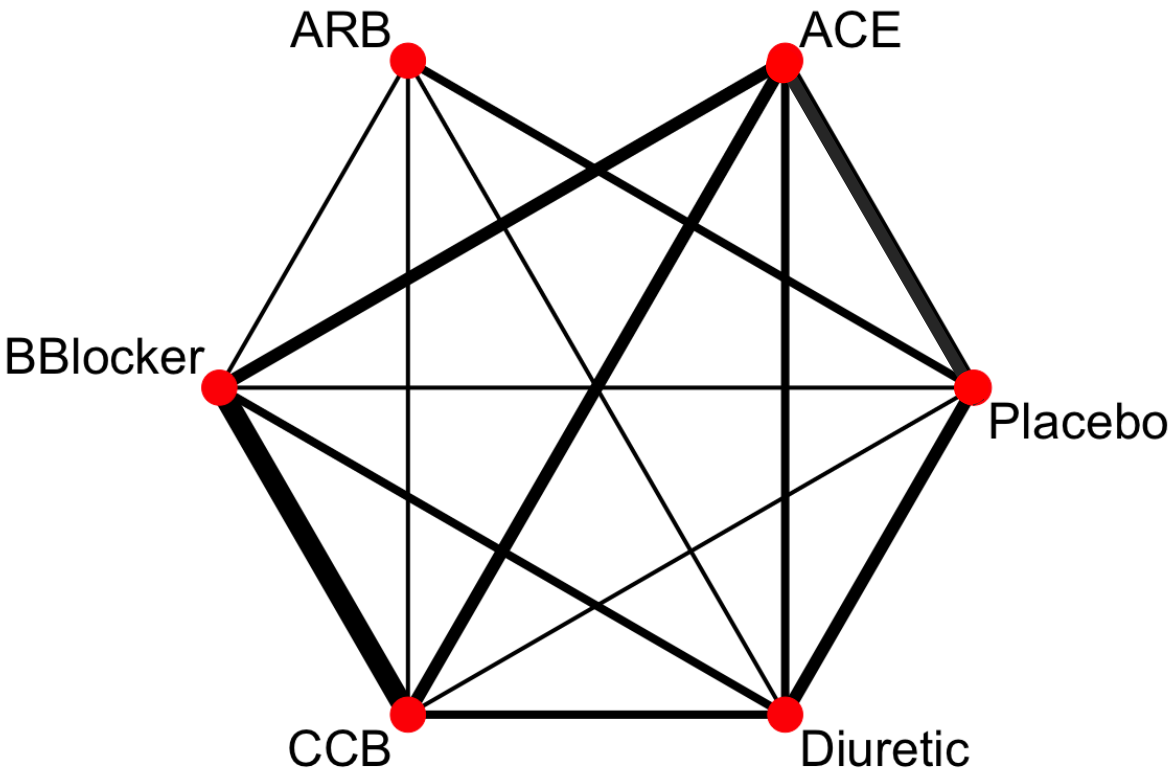
Separate Indirect from Direct Evidence test



Comparison	ACE:Placebo
Evidence: mixed	
NMA odds ratio:	0.885(0.769,1.017)
Direct odds ratio:	0.813(0.681,0.971)
Indirect odds ratio:	1.017(0.810,1.277)
Inconsistency measures	
Ratio of odds ratios:	0.799(0.599,1.067)
P value:	0.128

Dias et al. Checking consistency in mixed treatment comparison meta-analysis Stat Med 2010

INCOHERENCE



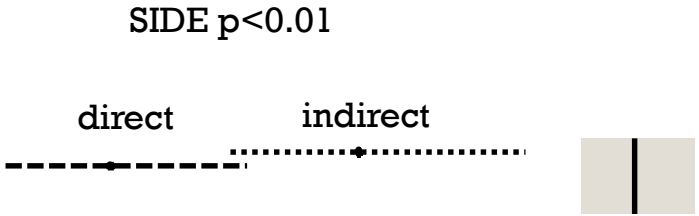
Design-by-treatment X^2 test

Does the assumption of coherence hold for the entire network?

χ^2 =19.325 (13 df)
P-value=0.113

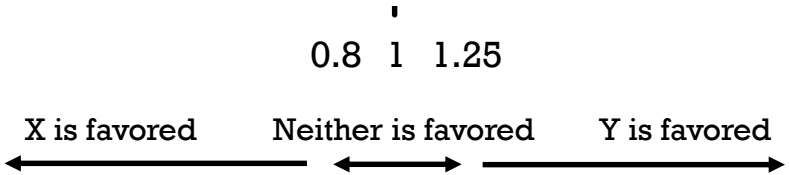
White et al. Consistency and inconsistency in network meta-analysis. Res Synth Meth 2012

INCOHERENCE

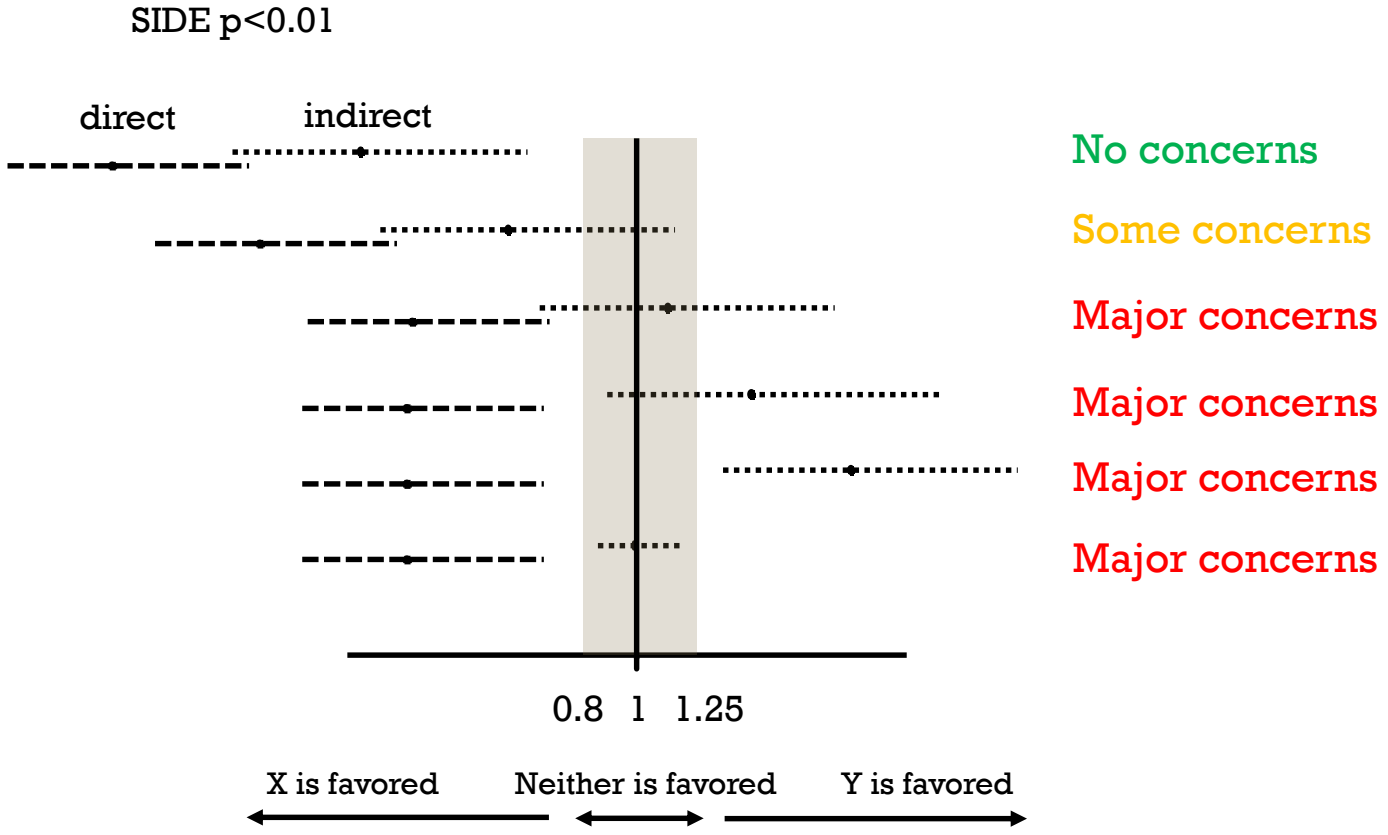


What is your judgement about incoherence for this estimate (SIDE test p-value <0.01)?

- ❑ Major concerns
- ❑ Some concerns
- ❑ No concerns



INCOHERENCE



INCOHERENCE

Comparisons with both direct and indirect evidence:
SIDE test p-value

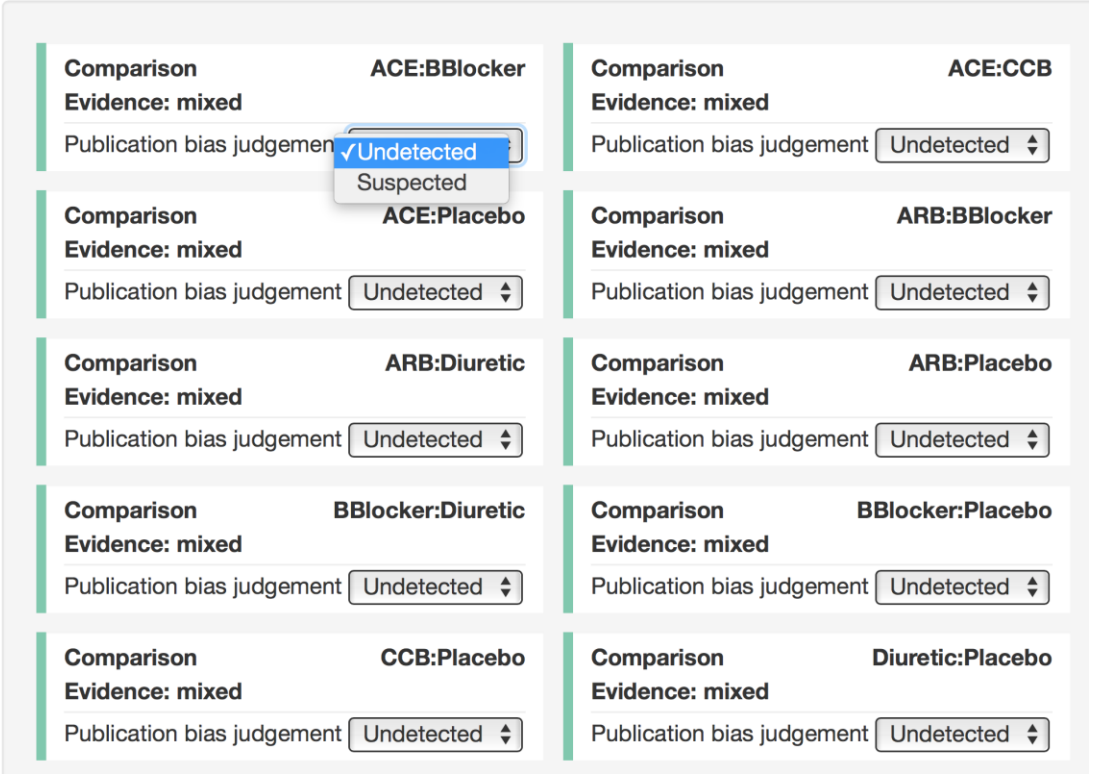
1. '*No concerns*' if p-value>0.10.
2. if p-value<0.10, check confidence interval overlaps and boundaries crossed.

Comparisons with only direct or indirect evidence:
design-by-treatment interaction test

1. '*Major concerns*' if p-value<0.05 or test is not estimable
2. '*Some concerns*' if 0.05<p-value<0.10
3. '*No concerns*' otherwise

REPORTING BIAS

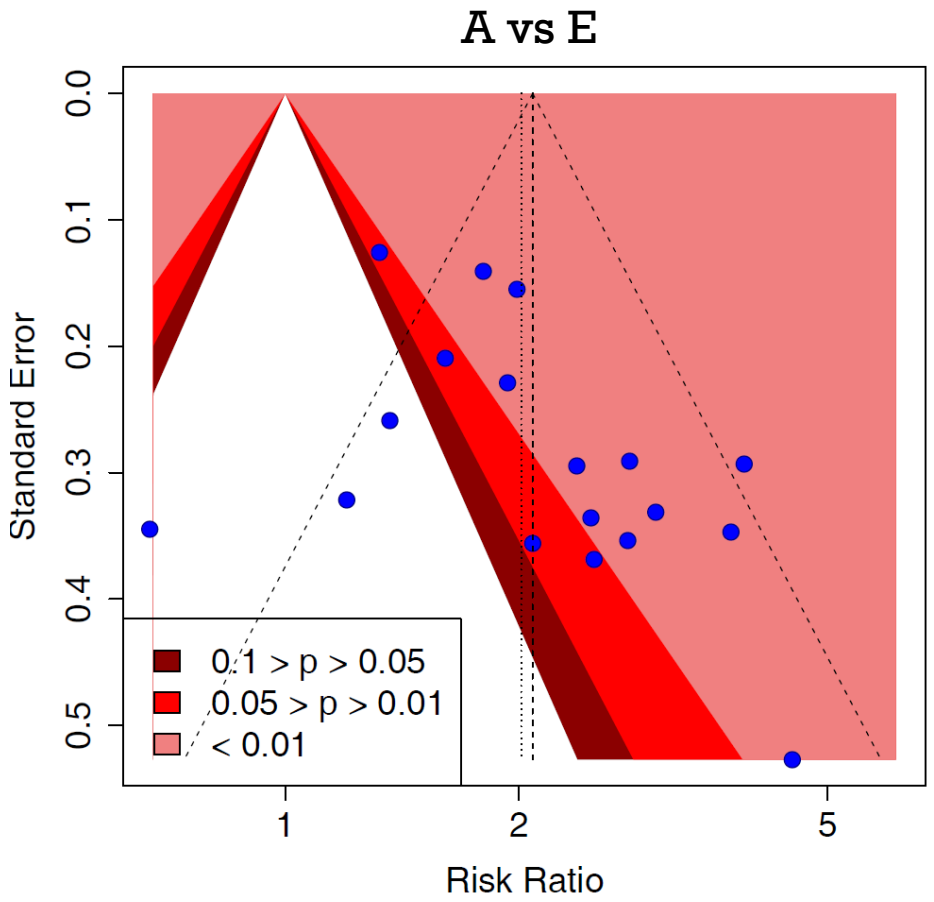
- ❑ Suspected
- ❑ Undetected



REPORTING BIAS

WORK IN PROGRESS

comparison	slope	p-value	interpretation
A vs E	0.21	0.05	"Small studies give smaller effect for 1st intervention"
A vs C	0.02	0	"Small studies give smaller effect for 1st intervention"
A vs D	0.19	0	"Small studies give smaller effect for 1st intervention"
A vs F	0.14	0.05	"Small studies give smaller effect for 1st intervention"
A vs G	-0.35	0.01	"Small studies give larger effect for 1st intervention"



REPORTING BIAS

WORK IN PROGRESS

Framework for assessing risk of bias due to missing results in a synthesis (Cochrane Handbook)

1. Select syntheses to assess for risk of bias due to missing results.
2. Define which results are eligible for inclusion in each synthesis.
3. Record whether any of the studies identified are missing from each synthesis because results known (or presumed) to have been generated by study investigators are unavailable: the '*known unknowns*'.
4. Consider whether each synthesis is likely to be biased because of the missing results in the studies identified.
5. Consider whether results from additional studies are likely to be missing from each synthesis: the '*unknown unknowns*'.
6. Reach an overall judgement about risk of bias due to missing results in each.

Now it is time for….

CINeMA

New updates funded by Cochrane

Reporting bias functionalities already mentioned

Already implemented

Update rules in judging imprecision, heterogeneity incoherence	Previous "rules" were too strict, we now consider one boundary and the null, effect when judging results.
Improve help with importing data	Prompting questions about nature of data
Facilitate scale up	Use ISPM's servers already provided
Sensitivity analysis for low RoB	User has to choose to exclude high or high and unclear studies and league table will be produced. If networks are disconnected the feature will be disabled
Presentations of results	League table and forest plot
Full report	Generate a PDF document for the entire process, including all graphs and tables and the final table and judgements
Final judgement	Choose the domains to downgrade by, and link them to the final confidence judgement (Each interim judgement is currently is marked as 'no concerns', 'some concerns' or 'major concerns', and these should be clickable to choose whether you want to downgrade by one or two levels)
Question mark buttons	To link the process steps with the documentation
Save past projects	Import export project
Technical testing	The system needs technical testing e.g. use weird data and see what it gives, testing with very large or disconnected networks etc. We will come up with 10 integration tests (tests that check entire functionality)

Questions ?