CONFIDENCE IN NETWORK META-ANALYSIS

Georgia Salanti & Theodore Papakonstantinou

Institute of Social and Preventive Medicine University of Bern Switzerland



The most critical question raised by patients and clinicians at the point of care is

"what is the drug of choice for the given condition?"

Del Fiol G et al. Clinical questions raised by clinicians at the point of care: a systematic review. JAMA Intern Med. 2014

CINEMA

Leucht S et al. Comparative efficacy and tolerability of 15 antipsychotic drugs in schizophrenia. Lancet 2013



CINEMA







Confidence In Network Meta-Analysis

CINEMA

Leucht S et al. Comparative efficacy and tolerability of 15 antipsychotic drugs in schizophrenia. Lancet 2013



Confidence In Network Meta-Analysis

CINEMA

456 published networks in the medical literature comparing at least 4 medical interventions (March 2015)

(Petropoulou et al. Journal of Clinical Epidemiology 2016, Zarin et al. BMC Medicine 2016)



None of the 456 NMAs published until 3/2015 attempted to evaluate the confidence in NMA results!

OPEN 승 ACCESS Freely available online

PLOS ONE

BMJ 2014:349:a5

A GRA

quality

meta-a

Network me

used to exa

on how to ra

We present

NMA estima

F

Evaluating the Quality of Evidence from a Network Meta-Analysis

Georgia Salanti¹, Cinzia Del Giovane², Anna Chaimani¹, Deborah M. Caldwell³, Julian P. T. Higgins^{3,4}*

1 Department of Hygiene and Epidemiology, University of Ioannina School of Medicine, Ioannina, Greece, 2 Statistics Unit, Department of Clinical and Diagnostic Medicine and Public Health, University of Modena and Reggio Emilia, Modena, Italy, 3 School of Social and Community Medicine, University of Bristol, Bristol, United Kingdom, 4 Centre for Reviews and Dissemination, University of York, York, United Kingdom

Abstract

Systematic reviews that collate data about the relative effects of multiple interventions via network meta-analysis are highly informative for decision-making purposes. A network meta-analysis provides two types of findings for a specific outcome: the relative treatment effect for all pairwise comparisons, and a ranking of the treatments. It is important to consider the confidence with which these two types of results can enable clinicians, policy makers and patients to make informed decisions. We propose an approach to determining confidence in the output of a network meta-analysis. Our proposed approach is based on methodology developed by the Grading of Recommendations Assessment, Development and Evaluation (GRADE) Working Group for pairwise meta-analyses. The suggested framework for evaluating a network metaanalysis acknowledges (i) the key role of indirect comparisons (ii) the contributions of each piece of direct evidence to the network meta-analysis estimates of effect size; (iii) the importance of the transitivity assumption to the validity of network meta-analysis; and (iv) the possibility of disagreement between direct evidence and indirect evidence. We apply our proposed strategy to a systematic review comparing topical antibiotics without steroids for chronically discharging ears with underlying eardrum perforations. The proposed framework can be used to determine confidence in the results from a Brignardello - crelsen - ousvinger / Singh : Anons envessers - eorgen - equal - or

CINEM

a published to very low a and likely to Milo A Puha Working Group

CINCMA framework

Consider the **network estimates**

Study limitations	— , — , — , — ,
Indirectness	Rate each network estimate
Inconsistency (heterogeneity, incoherence)	No concerns
Imprecision	Some concerns
Publication bias	Major concerns

Network estimate	Study limitations	Indirectness	Inconsistency		Imprecision	Publication bias	Confidence
			Heterogeneity	Incoherence			
A vs B	Some concerns	Some concerns	Major concerns	Some concerns	Some concerns	undetected	Very low
A vs C	No concerns	No concerns	No concerns Major Concerns		No concerns	suspected	Low

< > =

My Projects CINeMA Documentation

Methods developed by:

Georgia Salanti Julian Higgins Adriani Nikolakopoulou

Web developer:

Theodore Papakonstantinou

Project supervision: Matthias Egger

Welcome to CINEMA!

0

CINeMA (Confidence in Network Meta-Analysis) is a web application that simplifies the evaluation of confidence in the findings from n analysis.

cinema.ispm.ch

It is based on a framework described in (1) which considers the five GRADE domains: study limitations, indirectness, inconsistenc and publication bias. The framework combines judgments about direct evidence with their statistical contribution to network meta-a enabling evaluation of the credibility of NMA treatment effects.

1. Salanti G, Del Giovane C, Chaimani A, Caldwell DM, Higgins JPT. Evaluating the guality of evidence from a network meta-analysis. 2014;9(7):e99682.

To browse you projects or upload a new one go to My Projects



CINeMA is distributed, in the hope that it will be useful but without any warranty, under the AGPLV3 license



By using CINeMA you accept the following DISCLAIMER

UNIVERSITÄT BERN

C

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.

Number of studies 22

Number of treatment nodes 6

Drimary outcome	Effect of antihypertensives on incidence diabetes mellitus -
Filliary Outcome	proportion of patients who developed diabetes

Measurement Binary

Intervention comparison type pharmacological vs placebo

Lancet 2007; 369: 201–07

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





CINEMA

				Inco	nsistency						
Comparison	Number of Studies	Study Limitations	Imprecision	Heterogeneity	Incoherence	Indirectness	Publication bias	CONFIDENCE			
			Mixed evidence								
ACE vs BBlocker	3	No concerns	No concerns	Some concerns	Some concerns	No concerns	Undetected	MODERATE			
ACE vs CCB	3	No concerns	Some concerns	Some concerns	No concerns	No concerns	Undetected	MODERATE			
ACE vs Diuretic	2		Sen	ni-autom	ated proc	cess		MODERATE			
ACE vs Placebo	3				_			LOW			
ARB vs BBlocker	1		• • • • •		• • • • • •			MODERATE			
ARB vs CCB	1	Expl	icit rules	that class	siry each i	network	meta-	LOW			
ARB vs Diuretic	1		analysi	s effect fo	r each do	omain to		MODERATE			
ARB vs Placebo	2	Noc	oncerns	Some co	ncerns. M	laior coi	ncerns	VERY LOW			
BBlocker vs CCB	5		as desc	ribed in tl		ontation	n	MODERATE			
BBlocker vs Diuretic	2		as uese.					MODERATE			
BBlocker vs Placebo	1							VERY LOW			
CCB vs Diuretic	2		The r	ules can]	be overw	ritten!		MODERATE			
CCB vs Placebo	1	No concerns	Some concerns	Some concerns	No concerns	No concerns	Suspected	LOW			
Diuretic vs Placebo	3	No concerns	No concerns	Some concerns	No concerns	No concerns	Suspected	LOW			
			Indirect eviden	се							
ACE vs ARB		No concerns	Some concerns	Some concerns	No concerns	No concerns	Undetected	LOW			

The aim of the webinar is to explain the methods used in CINeMA and present an alpha version of the web application

pollev.com/gmhbe



STUDY LIMITATION

- Major concerns
- Some concerns
- No concerns





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

	<u>Study name</u>	<u>Risk of Bias</u>
	AASK	LOW
	ALLHAT	LOW
	ALPINE	LOW
	ANBP-2	LOW
	ASCOT	LOW
CCB vs Diuretics:	CAPPP	MODERATE
overall low right of bigg	CHARM	LOW
overall low lisk of blas	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

CINEMA





Comparison

OR from NMA

B vs Placebo

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



Go to:

pollev.com/gmhbe







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



Comparison

ARB

OR from NMA

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









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

	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
Mixed estimate	<u>5</u>													
ACE:BBlocker	32	10	10	8	6	1	0	4	15	6	2	5	2	0
ACE:CCB	10	26	13	11	1	6	0	4	9	1	0	13	6	0
ACE:Diuretic	6	7	57	5	0	2	0	2	1	5	0	12	2	2
ACE:Placebo	5	7	5	56	3	3	0	6	1	0	2	3	8	2
ARB:BBlocker	4	1	0	3	41	21	0	5	19	2	2	2	1	0
ARB:CCB	1	2	1	2	8	67	0	6	8	1	0	2	4	0
ARB:Diuretic	3	2	11	5	10	27	0	8	0	7	0	25	0	2
ARB:Placebo	3	3	2	7	6	15	0	49	0	1	2	2	10	1
BBlocker:CCB	6	4	1	1	11	12	0	0	53	4	2	5	2	0
BBlocker:Diuretic	: 10	1	13	2	5	3	0	2	19	20	2	21	0	2
BBlocker:Placebo	o 10	2	2	14	13	3	0	16	16	4	8	1	11	2
CCB:Diuretic	2	6	11	3	1	3	0	2	7	6	0	56	3	2
CCB:Placebo	2	6	4	12	1	15	0	16	6	0	2	5	28	2
Diuretic:Placebo	0	0	20	20	2	7	0	9	0	5	2	17	11	7
Indirect estimat	es													
ACE:ARB	10	11	8	16	11	20	0	14	1	1	0	7	2	0
													C	NeMA

The contribution matrix

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

	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
Mixed estimate	<u>5</u>													
ACE:BBlocker	32	10	10	8	6	1	0	4	15	6	2	5	2	0
ACE:CCB	10	26	13	11	1	6	0	4	9	1	0	13	6	0
ACE:Diuretic	6	7	57	5	0	2	0	2	1	5	0	12	2	2
ACE:Placebo	5	7	5	56	3	3	0	6	1	0	2	3	8	2
ARB:BBlocker	4	1	0	3	41	21	0	5	19	2	2	2	1	0
ARB:CCB	1	2	1	2	8	67	0	6	8	1	0	2	4	0
ARB:Diuretic	3	2	11	5	10	27	0	8	0	7	0	25	0	2
ARB:Placebo	3	3	2	7	6	15	0	49	0	1	2	2	10	1
BBlocker:CCB	6	4	1	1	11	12	0	0	53	4	2	5	2	0
BBlocker:Diureti	= 10	1	13	2	5	3	0	2	19	20	2	21	0	2
BBlocker:Placebo	b 10	2	2	14	13	3	0	16	16	4	8	1	11	2
CCB:Diuretic	2	6	11	3	1	3	0	2	7	6	0	56	3	2
CCB:Placebo	2	6	4	12	1	15	0	16	6	0	2	5	28	2
Diuretic:Placebo	0	0	20	20	2	7	0	9	0	5	2	17	11	7
Indirect estimat	tes													
ACE:ARB	10	11	8	16	11	20	0	14	1	1	0	7	2	0
													C	NeMA

The contribution matrix

The contribution matrix



The contribution matrix

	ACE: BBlocker	ACE: CCB	ACE: Diuretic	ACE: Placebo	ARB: BBlocker	ARB: CCB	ARB: Diuretic	ARB: Placebo	BBlocker: CCB	BBlocker: Diuretic	BBlocker: Placebo	CCB: Diuretic
<u>Mixed estimates</u>									ACE			
ACE:BBlocker	32	10	10	8	6							5
ACE:CCB	10	26	13	11	1							13
ACE:Diuretic	6	7	57	5	0		Placebo				BBlocker	12
ACE:Placebo	5	7	5	56	3							3
ARB:BBlocker	4	1	0	3	41							2
ARB:CCB	1	2	1	2	8							2
ARB:Diuretic	3	2	11	5	10				×	X		25
ARB:Placebo	3	3	2	7	6							2
BBlocker:CCB	6	4	1	1	11							5
BBlocker:Diuretic	10	1	13	2	5							21
BBlocker:Placebo	10	2	2	14	13		ARB				ССВ	1
CCB:Diuretic	2	6	11	3	1			\sim				56
CCB:Placebo	2	6	4	12	1							5
Diuretic:Placebo	0	0	20	20	2				Diuretic			17
Indirect estimates	<u>i</u>											
ACE:ARB	10	11	8	16	11	20	0	14	1	1	0	7
		10	20	30	40		50	60	70	80	90	100

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

Major concerns Some concerns No concerns

INDIRECTINESS

- Major concerns
- Some concerns
- No concerns

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.ispm.ch

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
 - Could be set using the Minimum Clinically Important Difference

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**

CINeMA

NMA estimated odds ratios for diabetes

Now it is time for....

INCONSISTENCY

HETEROGENEITY

- Major concerns
 - **Some concerns**
- No concerns

INCOHERENCE

- Major concerns
- **Some concerns**
- No concerns

INCONSISTENCY

Heterogeneity disabetweentstudyeen dvariance within af comparison

- 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

CINEMA

Rules implemented in the software

No concerns: Confidence and prediction intervals agree in relation to clinically important effect No concerns: Confidence and prediction intervals agree in relation to clinically important effects Some concerns: Prediction interval extends into clinically important or unimportant effects Major concerns: Prediction interval extends into clinically important effects in both directions Major concerns: Prediction interval extends into clinically important effects in both directions No concerns: Confidence and prediction intervals agree in relation to clinically important effects Some concerns: Prediction interval extends into clinically important or unimportant effect No concerns: Confidence and prediction intervals agree in relation to clinically important effects No concerns: Confidence and prediction intervals agree in relation to clinically important effects

Margin of equivalence

Prediction interval — Confidence interval —

- 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 <u>observed values</u> of τ^2 are can be compared with the <u>expected values</u> 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

Comparison Evidence: mixed	ACE:BBlocker
Between-study heteroge	eneity for each
direct comparison	
l ² :	49.8 %
Estimated τ^2 :	0.019
Reference Values for τ^2	
first quantile:	0.003
median:	0.014
third quantile:	0.061
95% intervals for NMA e	stimate
Confidence interval:	(1.245,1.498)
Prediction interval:	(0.992,1.879)
Prediction interval extend	s into clinically
important or unimportan	nt effects
Heterogeneity judgement	Serious 🜲

Comparison Evidence: mixed	ARB:BBlocker
Between-study heterog direct comparison	eneity for each
l ² : Estimated τ ² :	NA NA
Reference Values for τ^2	
first quantile:	0.003
median:	0.014
third quantile:	0.061
95% intervals for NMA	estimate
Confidence interval:	(1.372,1.657)
Prediction interval:	(1.094,2.077)
Confidence and predictio relation to clinically impor	n intervals agree in tant effect
Heterogeneity judgement	No serious

Comparison Evidence: mixed	BBlocker:CCB					
Between-study heterogeneity for each direct comparison						
l ² :	62.5%					
Estimated τ ² :	0.013					
Reference Values for τ^2						
first quantile:	0.003					
median:	0.014					
third quantile:	0.061					
95% intervals for NMA est	imate					
Confidence interval:	(0.768,0.871)					
Prediction interval:	(0.600,1.115)					
Prediction interval extends into clinically important effects in both directions						
Heterogeneity judgement	/ery Serious 💲					

INCONSISTENCY

Heterogeneity between-study variance within a comparison Incoherence disagreement between different sources of evidence

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

INCONSISTENCY

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

Separate Direct from Indirect Evidence test (node-splitting) : Compare direct and indirect relative treatment effects using a Z-test : one test for each treatment comparisons

Design-by-treatment test X²

: one test for the network

Separate Direct from Indirect Evidence test

ARB

Does the assumption of coherence hold for the entire network?

BBlocker Placebo Diuretic CCBWhite et al. Consistency and inconsistency in network meta-analysis. Res Synth Meth 2012

ACE

Treatment comparisons that take at least 90% of the information from direct evidence have no concerns for incoherence For comparisons with at least 10% of information derived from indirect evidence we use the following rules

Design-by-treatment interaction model

		p-value>0.1	0.01 <p-value<0.1< th=""><th>p-value<0.01</th></p-value<0.1<>	p-value<0.01
	p-value>0.1	No concerns	No concerns	Some concerns
SIDE	0.01 <p-value<0.1< th=""><th>Some concerns</th><th>Some concerns</th><th>Major concerns</th></p-value<0.1<>	Some concerns	Some concerns	Major concerns
	p-value<0.01	Some concerns	Major concerns	Major concerns

Confidence In Network Meta-Analysis

CINeM/

Comparison Evidence: mixed	ACE:BBlocker
Direct contribution:	51.4%
Inconsistency measure	es
Ratio of odds ratios:	0.719(0.533,0.969)
Z statistic:	-2.165
P value:	0.030
Incoherence judgement	
Some concerns 🜲	

Comparison Evidence: mixed	ARB:CCB
Direct contribution:	41.7%
Inconsistency measur	es
Ratio of odds ratios:	1.012(0.709,1.444)
Z statistic:	0.066
P value:	0.948
Incoherence judgement	
No concerns	

Comparison	BBlocker:Placebo
Evidence: mixed	
Direct contribution:	
Inconsistency measur	es
Ratio of odds ratios: 0.524(0.299.0.918)	

Comparison Evidence: mixed	ACE:CCB
Direct contribution:	41.5%
Inconsistency measure	S
Ratio of odds ratios:	1.099(0.810,1.490)
Z statistic:	0.605
P value:	0.545
Incoherence judgement	
No concerns	

Comparison Evidence: mixed	ARB:Diuretic
Direct contribution:	1.0%
Inconsistency measures	
Ratio of odds ratios:	5.247(0.634,43.445)
Z statistic:	1.537
P value:	0.124
Incoherence judgemen	t
No concerns 🗘	

Comparison	CCB:Diuretic	
Evidence: mixed		
Direct contribution:	48.0%	
Inconsistency measu	res	
Ratio of odds ratios: 0.932(0.676.1.286)		

Salanti G. Indirect and mixed-treatment comparison, network, or multiple-treatments metaanalysis. Res Synth Meth 2012

CINEMA

PUBLICATION BIAS

SuspectedUndetected

Comparison Evidence: mixed	ACE:BBlocker	Comparison ACE:CCB Evidence: mixed
Publication bias judgemen		Publication bias judgement Undetected \$
	Suspected	
Comparison Evidence: mixed	ACE:Placebo	ComparisonARB:BBlockerEvidence: mixed
Publication bias judgemen	Undetected 🗘	Publication bias judgement Undetected 🗘
Comparison Evidence: mixed	ARB:Diuretic	Comparison ARB:Placebo Evidence: mixed
Publication bias judgemen	Undetected	Publication bias judgement Undetected \$
Comparison Evidence: mixed	BBlocker:Diuretic	Comparison BBlocker:Placebo Evidence: mixed
Publication bias judgemen	t Undetected 🖨	Publication bias judgement Undetected \$
Comparison Evidence: mixed	CCB:Placebo	ComparisonDiuretic:PlaceboEvidence: mixed
Publication bias judgemen	Undetected 🖨	Publication bias judgement Undetected \$

Now it is time for....

DISCLAIMER

You are welcome to use CINeMA with the understanding that

- it is still under development
 - We will improve the data input module
 - We will fix some known bugs in the calculations
 - For some calculations CINeMA the netmeta package in R, so updates/debugging in netmeta affect CINeMA too
 - Please notify us for any problems you come across <u>cinema.ispm@gmail.com</u>
 - If you use it in a publication you can cite

CINeMA: Confidence in Network Meta-Analysis [Software]. University of Bern 2017. Available from <u>cinema.ispm.ch</u>

