



Department of Clinical Epidemiology and Biostatistics

Journal Club
06/06/2025

A guide and pragmatic considerations for applying GRADE to network meta-analysis

Nattapat Khumtong
6636457



A guide and pragmatic considerations for applying GRADE to network meta-analysis

Ariel Izcovich,¹ Derek K Chu,^{2,3} Reem A Mustafa,^{2,4} Gordon Guyatt,^{2,3}
Romina Brignardello-Petersen²

¹Hospital Alemán de Buenos Aires, Buenos Aires, C1118AAT CABA, Argentina

²Department of Health Research Methods, Evidence, and Impact McMaster University, Hamilton, ON, Canada

³Department of Medicine, McMaster University, Hamilton, ON, Canada

⁴Department of Medicine, University of Kansas Medical Center, Kansas City, MO, USA

Correspondence to: A Izcovich
Ariel.izcovich@gmail.com
(ORCID 0000-0001-9053-4396)

Additional material is published online only. To view please visit the journal online.

Cite this as: *BMJ* 2023;381:e074495
<http://dx.doi.org/10.1136/bmj-2022-074495>

Accepted: 20 May 2023

Assessing the certainty of evidence from network meta-analyses using the GRADE (grading of recommendations, assessment, development, and evaluations) approach requires not only a thorough understanding of the methods but also substantial workload for raters. This article describes how implementing practical strategies (including rule setting and automation) can facilitate efficient application of the GRADE approach to rating certainty of evidence in network meta-analyses while maintaining rigor. This article describes a stepwise strategy for

spreadsheet that incorporates automation of several of the steps described is also provided (<https://www.covid19lnma.com/>).

Network meta-analysis (NMA) allows assessing the comparative effectiveness of multiple interventions by combining direct and indirect evidence in one statistical model, resulting in estimates of effect comparing every pair of interventions included in the network—even if they have not been compared directly in trials.¹ Assessing the certainty of the evidence (also known as quality of the evidence, and confidence in the estimates of effects) from NMAs is crucial for interpreting those estimates and moving from evidence to decision. The GRADE (grading of recommendations, assessment, development, and evaluations) working group has provided guidance for assessing the certainty of the evidence and drawing conclusions from NMAs.²⁻



Outline

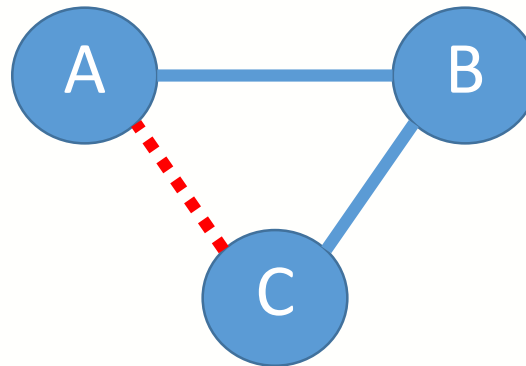
- Introduction
- GRADE for NMA
- Data preparation
- Process for CoE using GRADE for NMA
- Example
- Conclusion



Introduction

- Network meta-analysis (NMA) allows assessing the comparative effectiveness of multiple interventions by combining direct and indirect evidence in one statistical model

>> be able to compare every pair of interventions even if they have not been compared directly in trials





Introduction

- Assessing the **“Certainty of the evidence(CoE)”** from NMAs is crucial for interpreting those estimates and moving from evidence to decision.



GRADE for NMA

- The **GRADE** (grading of recommendations, assessment, development, and evaluations) working group for NMA
- Firstly proposed in 2014
- Certainty of the evidence = how confident we are that the **true effect of an intervention** is close to what the studies found.

GRADE rates evidence for each outcome as:

High (⊕⊕⊕⊕) — Very confident in the effect estimate.

Moderate (⊕⊕⊕○) — Moderately confident; true effect may differ slightly.

Low (⊕⊕○○) — Limited confidence; true effect may differ substantially.

Very Low (⊕○○○) — Very little confidence in the effect estimate.



GRADE for NMA

- Five Domains That Can Lower Certainty

1. **Risk of Bias** — Flaws in study design or execution.
2. **Inconsistency** — Wide variation in results across studies.
3. **Indirectness** — Evidence doesn't directly apply (e.g., different population or intervention).
4. **Imprecision** — Small sample sizes or wide confidence intervals.
5. **Publication Bias** — Missing studies, often with negative or null results.



Data preparation

Adequate assessment of the CoE for every assessed outcome requires

- NMA estimates and 95% CI for every comparison using both relative and absolute estimates of effect for each outcome
- Direct and indirect estimates and 95% CI for each comparison
- Risk-of bias assessments for each outcome
- Forest plots of all direct comparisons
- Network plot for each outcome



Department of Clinical Epidemiology and Biostatistics

Example 1 (from Siemieniuk et. al)¹: Comparison: Systemic corticosteroids;
Outcome: Mortality

Estimates of effect (Direct)

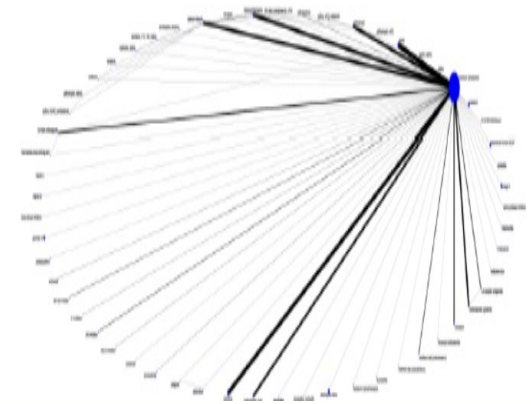
Treatment 1	Treatment 2	Relative estimate			Absolute estimate (per 1000)		
		Point estimate	CI lower limit	CI upper limit	Point estimate	CI lower limit	CI upper limit
Corticosteroids	SOC	0.84	0.64	1.07	-19.06	-42.5	7.6

Estimates of effect (Indirect)

Treatment 1	Treatment 2	Relative estimate			Absolute estimate (per 1000)		
		Point estimate	CI lower limit	CI upper limit	Point estimate	CI lower limit	CI upper limit
Corticosteroids	SOC	0.36	0.16	0.8	-78.8	-107.1	-22.8

Estimates of effect (NMA)

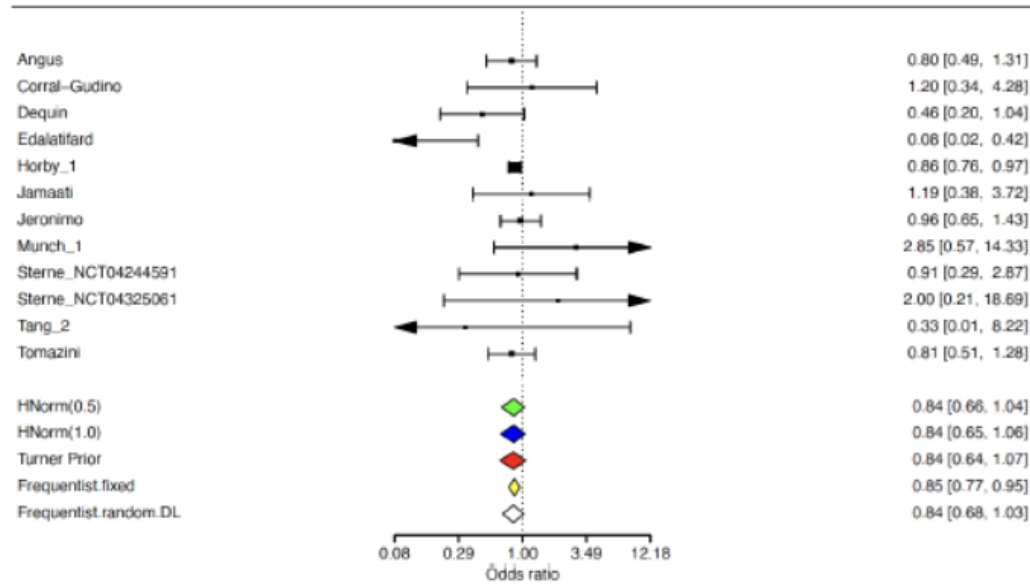
Treatment 1	Treatment 2	Relative estimate			Absolute estimate (per 1000)		
		Point estimate	CI lower limit	CI upper limit	Point estimate	CI lower limit	CI upper limit
Corticosteroids	SOC	0.8	0.65	0.94	-23.48	-40.24	-6.94





Department of Clinical Epidemiology and Biostatistics

Forest plot



Risk of bias assessment

study/studies (synthesis)	standard care/placebo								
29672, 6743	Corral-Gudino	low risk of bias	probably high risk of bias	low risk of bias	low risk of bias	low risk of bias	low risk of bias	low risk of bias	low risk of bias
33061, 7061, 862	Horby_1	low risk of bias	probably high risk of bias	low risk of bias	low risk of bias	low risk of bias	low risk of bias	low risk of bias	low risk of bias
34899	Jamaati	low risk of bias	low risk of bias	low risk of bias	low risk of bias	low risk of bias	low risk of bias	low risk of bias	low risk of bias
33862	Dequin	low risk of bias	low risk of bias	low risk of bias	low risk of bias	low risk of bias	low risk of bias	low risk of bias	low risk of bias
33361	Jeronimo	low risk of bias	probably high risk of bias	low risk of bias	low risk of bias	low risk of bias	low risk of bias	low risk of bias	low risk of bias
33348	Angus	probably high risk of bias	probably high risk of bias	low risk of bias	low risk of bias	low risk of bias	low risk of bias	low risk of bias	low risk of bias
33376_NCT041250	Sterne_NCT04325061	low risk of bias	probably high risk of bias	low risk of bias	low risk of bias	low risk of bias	low risk of bias	low risk of bias	low risk of bias
33376_NCT041250	Sterne_NCT04325061	probably low risk of bias	probably high risk of bias	low risk of bias	low risk of bias	low risk of bias	low risk of bias	low risk of bias	low risk of bias
33269	Edalatfard	probably high risk of bias	probably high risk of bias	probably high risk of bias	low risk of bias	low risk of bias	low risk of bias	low risk of bias	low risk of bias
2	Parashar	probably low risk of bias	probably low risk of bias	low risk of bias	low risk of bias	low risk of bias	low risk of bias	probably low risk of bias	low risk of bias
27649	Tang_2	probably low risk of bias	probably high risk of bias	low risk of bias	low risk of bias	low risk of bias	low risk of bias	low risk of bias	low risk of bias
31163	Jamaati	probably high risk of bias	probably high risk of bias	low risk of bias	low risk of bias	low risk of bias	low risk of bias	low risk of bias	low risk of bias
49239	Munch_1	low risk of bias	low risk of bias	low risk of bias	low risk of bias	low risk of bias	low risk of bias	low risk of bias	low risk of bias

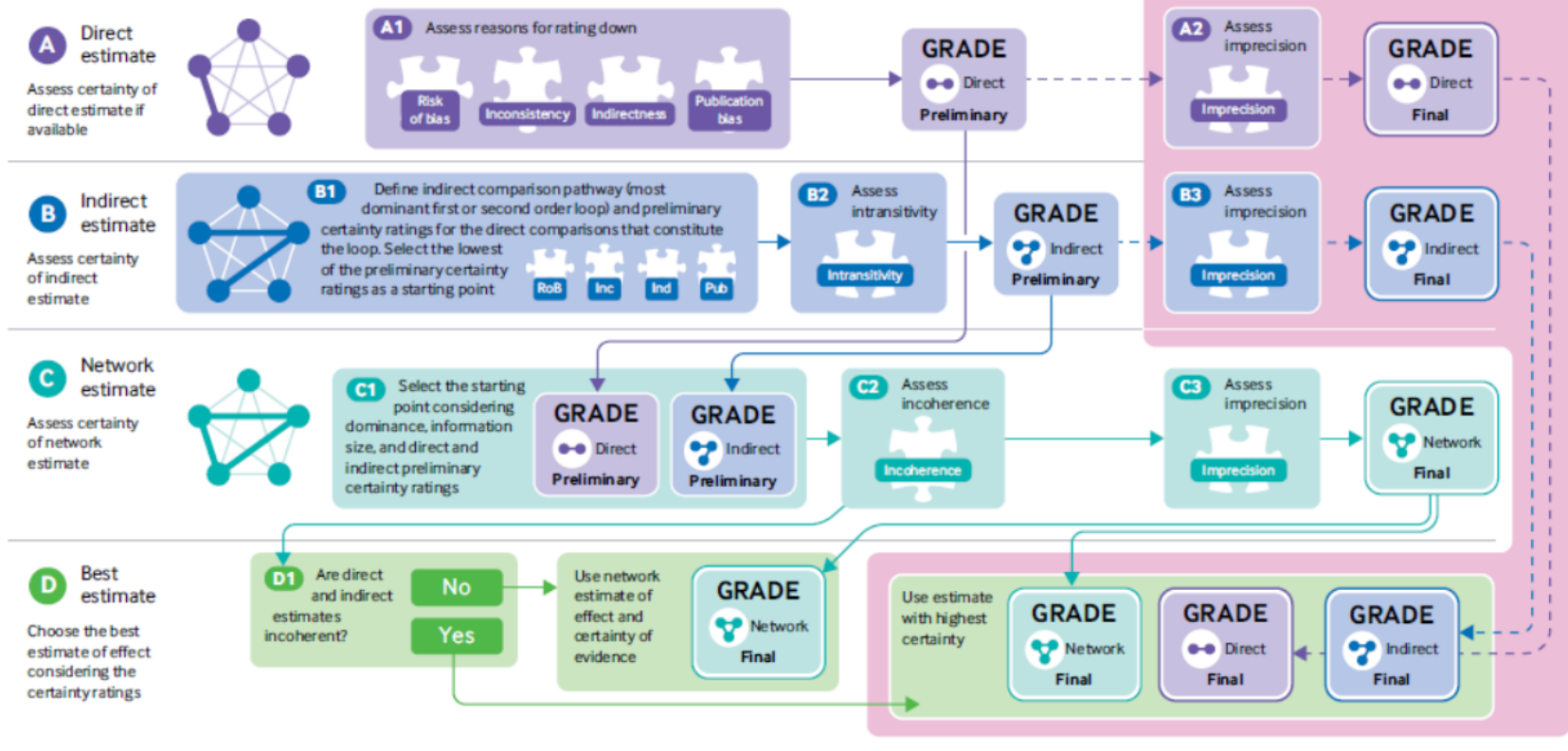
Assessed domains in order from left to right: Randomization, deviations from de intended intervention, missing outcome data, measurement of outcome, selection of the reported results.



Process for assessing CoE for NMA

Assessing the certainty of evidence for one comparison in a network meta-analysis

This diagram explains the process of assessing a single comparison within a network meta-analysis.
For a summary of the process for assessing certainty in a whole network, please see supplementary table 1





Process for assessing CoE for NMA

Direct estimate

Assessing the certainty of evidence for one comparison in a network meta-analysis

This diagram explains the process of assessing a single comparison within a network meta-analysis.
For a summary of the process for assessing certainty in a whole network, please see supplementary table 1



GRADE Rating Levels for Each Domain

- Not serious (0) : the issue is negligible or well-managed.
- Serious (-1 level) : The issue likely affects confidence in the results.
- Very serious (-2 level) : The issue severely limits confidence in the results.



Process for assessing CoE for NMA

Direct estimate

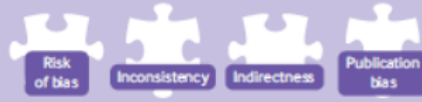
Assessing the certainty of evidence for one comparison in a network meta-analysis

This diagram explains the process of assessing a single comparison within a network meta-analysis.
For a summary of the process for assessing certainty in a whole network, please see supplementary table 1

A Direct estimate
Assess certainty of direct estimate if available



A1 Assess reasons for rating down



GRADE
Direct
Preliminary

Only necessary if direct and indirect estimates prove incoherent

A2 Assess imprecision
Imprecision

GRADE
Direct
Final

RESET	Comparison	Threshold per 1000 (small)
Nr.		Threshold per 1000 (Large)
	Treatment 1	Treatment 2
1	A	Plac
2	B	Plac
3	C	Plac
4	D	Plac
5	A	B
6	A	C
7	A	D
8	B	C
9	B	D
10	C	D

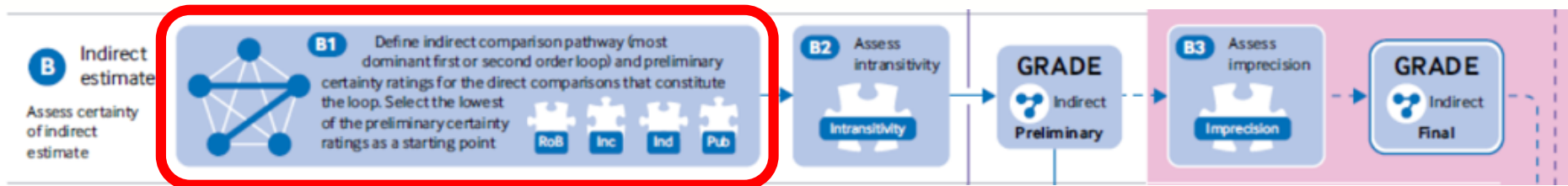
GRADE					
Risk of bias	Inconsistency	Indirectness	Publication bias	Rating for NMA	
Not serious	Not serious	Not serious	Not serious	HIGH	
Serious	Not serious	Not serious	Not serious	MODERATE	
Serious	Not serious	Not serious	Not serious	MODERATE	
Not serious	Not serious	Not serious	Not serious	HIGH	
Not serious	Not serious	Not serious	Not serious	HIGH	
Not serious	Not serious	Not serious	Not serious	HIGH	

Imprecision	Final rating
Extremely serious	VERY LOW
Serious	LOW
Very serious	VERY LOW
Extremely serious	VERY LOW
Extremely serious	VERY LOW
	NA
	NA
Very serious	LOW
	NA
	NA



Process for assessing CoE for NMA

Indirect estimate

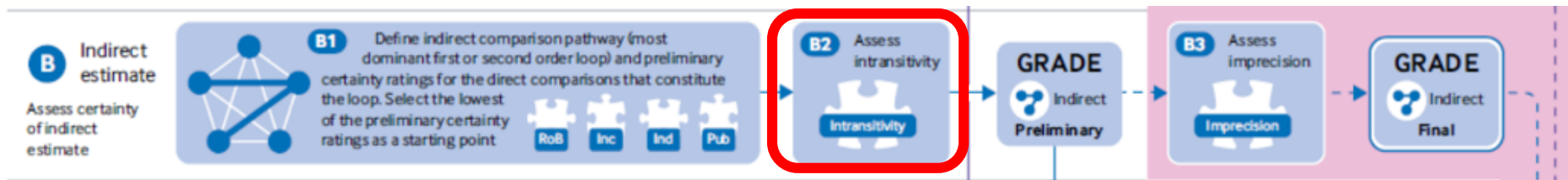


- The certainty of the evidence of indirect comparisons is based on the certainty ratings of the direct comparisons on which those indirect estimates are calculated.
- The initial certainty of the indirect estimate **is the lowest of the preliminary certainty ratings of the direct comparisons** that constitute the most dominant first order loops



Process for assessing CoE for NMA

Indirect estimate

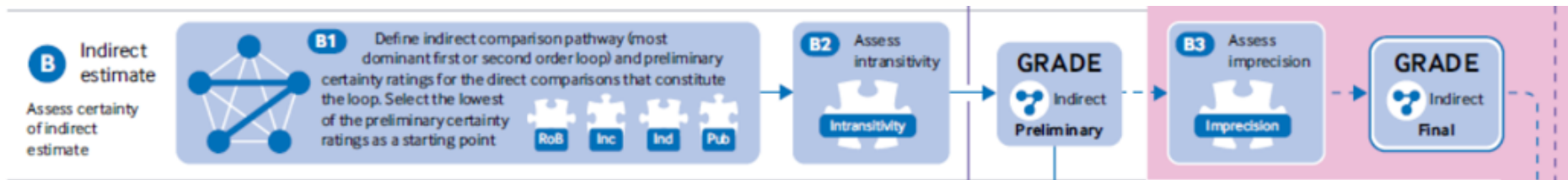


- The transitivity assumption must not be violated
- Can be assessed using checklists, or the ICEMAN instrument.



Process for assessing CoE for NMA

Indirect estimate

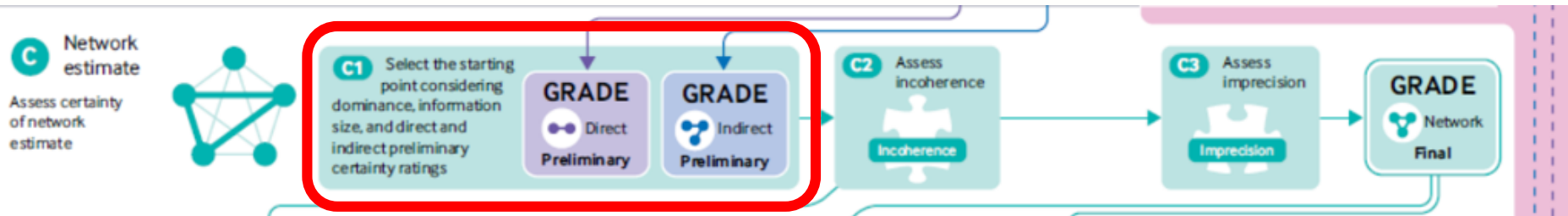


RESET	Comparison	Threshold per 1000 (small)	Threshold per 1000 (Large)					
Nr.					GRADE			
	Treatment 1	Treatment 2		Start rating	Intransitivity	Rating for NMA	Imprecision	Final rating
1	A	Plac						NA
2	B	Plac		MODERATE	Not serious	MODERATE	Very serious	VERY LOW
3	C	Plac		MODERATE	Not serious	MODERATE	Very serious	VERY LOW
4	D	Plac						NA
5	A	B						NA
6	A	C		HIGH	Not serious	HIGH	Extremely serious	VERY LOW
7	A	D		HIGH	Not serious	HIGH	Extremely serious	VERY LOW
8	B	C		MODERATE	Not serious	MODERATE	Very serious	VERY LOW
9	B	D		MODERATE	Not serious	MODERATE	Extremely serious	VERY LOW
10	C	D		MODERATE	Not serious	MODERATE	Extremely serious	VERY LOW



Process for assessing CoE for NMA

Network estimate

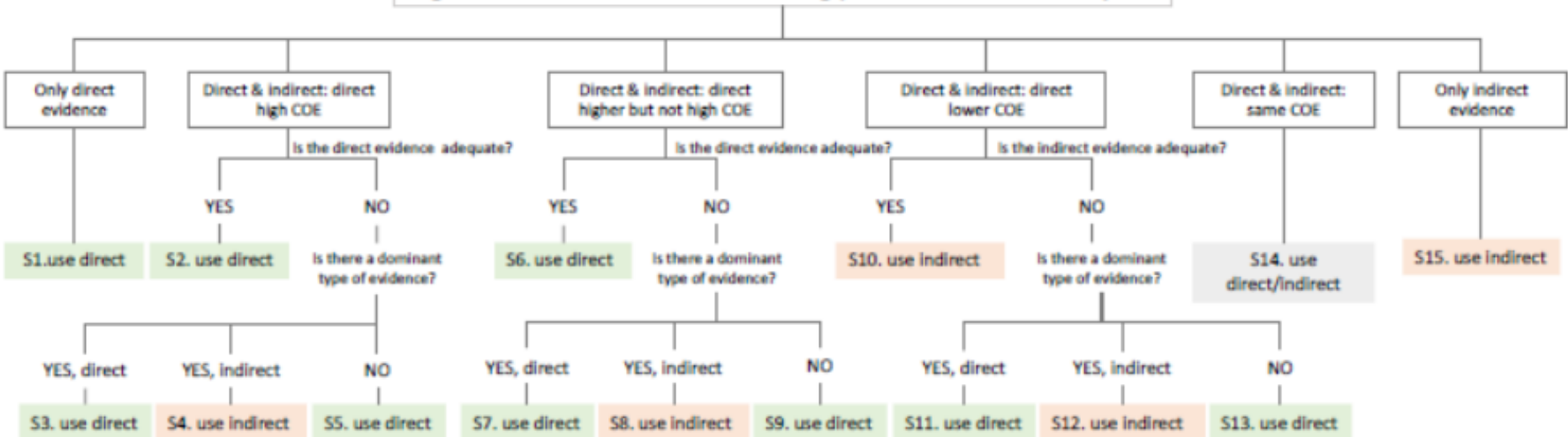


- The initial certainty of network estimate is based on the information size, result from direct or indirect preliminary certainty rating in step A and B



Department of Clinical Epidemiology and Biostatistics

Algorithm to determine the starting point of NMA certainty*



S1. Only direct evidence; use direct
 S2. Direct & indirect: direct high COE; adequate; use direct
 S3. High COE of direct; inadequate; dominant direct, use direct
 S4. High COE of direct, inadequate, dominant indirect, use indirect
 S5. High COE of direct, inadequate, undominant, use direct
 S6. Direct higher but not high COE; adequate; use direct
 S7. Direct higher but not high COE; inadequate; dominant direct; use direct
 S8. Direct higher but not high COE; inadequate; dominant indirect; use indirect
 S9. Direct higher but not high COE; inadequate; undominant; use direct
 S10. Direct lower COE; adequate indirect; use indirect
 S11. Direct lower COE; in adequate indirect; dominant direct; use direct

S12. Direct lower COE; in adequate indirect; dominant indirect; use indirect
 S13. Direct lower COE; in adequate indirect; undominant; use direct
 S14. Same COE to direct and indirect; use direct or indirect
 S15. Only indirect; use indirect

Summary:

In S1-S3: the starting point of NMA certainty is based on the direct evidence, and thus, not relevant to indirect evidence;
 In S4-S14: the starting point of NMA certainty should consider both the indirect evidence and the direct evidence;
 In S15: only consider indirect evidence.

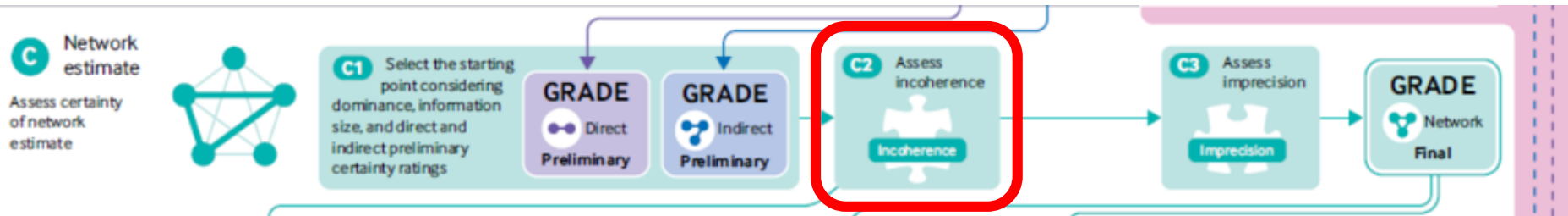
COE: Certainty of the evidence

*The starting point NMA certainty not including incoherence and imprecision



Process for assessing CoE for NMA

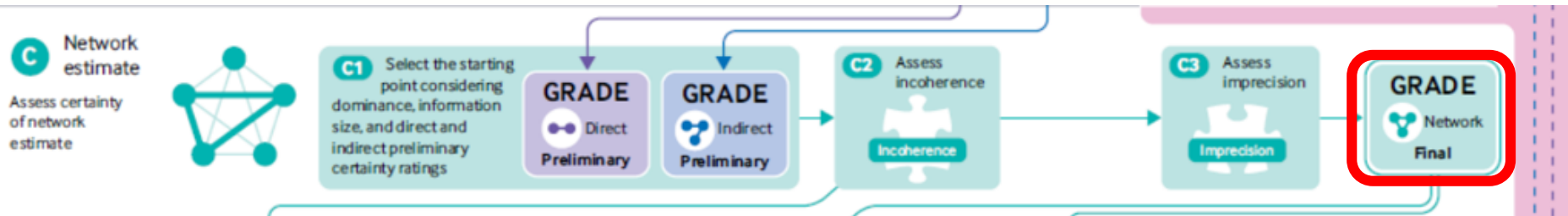
Network estimate



- When results from direct and indirect estimate diverge beyond what chance can explain and reviewers fail to identify an explanation, reviewers should consider rating down the certainty of the NMA estimate for incoherence.



Process for assessing CoE for NMA Network estimate

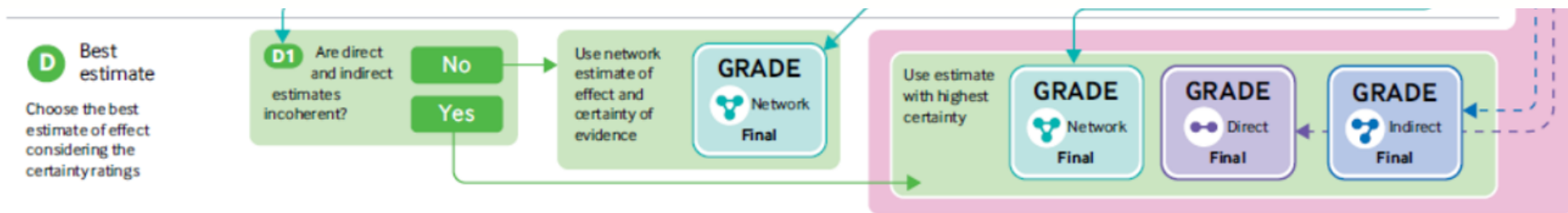


RESET	Comparison		Threshold per 1000 (small)							
Nr.			Threshold per 1000 (Large)							
	Treatment 1	Treatment 2	GRADE	Starting rating	Incoherence	Imprecision	Final rating	Best estimate	Reasons for downgrading auto-calculator	
1	A	Plac	Starting evidence Direct	HIGH	Not serious	Extremely serious	VERY LOW	NMA	Imprecisionx3	
2	B	Plac	Direct	MODERATE	Not serious	Serious	LOW	NMA	RoB, Imprecision	
3	C	Plac	Indirect	MODERATE	Not serious	Very serious	VERY LOW	NMA	RoB, Imprecisionx2	
4	D	Plac	Direct	HIGH	Not serious	Extremely serious	VERY LOW	NMA	Imprecisionx3	
5	A	B	Direct	HIGH	Not serious	Extremely serious	VERY LOW	NMA	Imprecisionx3	
6	A	C	Indirect	HIGH	Not serious	Extremely serious	VERY LOW	NMA	Imprecisionx3	
7	A	D	Indirect	HIGH	Not serious	Extremely serious	VERY LOW	NMA	Imprecisionx3	
8	B	C	Direct	HIGH	Not serious	Very serious	LOW	NMA	Imprecisionx2	
9	B	D	Indirect	MODERATE	Not serious	Extremely serious	VERY LOW	NMA	RoB, Imprecisionx3	
10	C	D	Indirect	MODERATE	Not serious	Extremely serious	VERY LOW	NMA	RoB, Imprecisionx3	



Process for assessing CoE for NMA

Best estimate



RESET	Comparison	Threshold per 1000 (small)	Threshold per 1000 (Large)	GRADE	Starting rating	Incoherence	Imprecision	Final rating	Best estimate	Reasons for downgrading auto-calculated
Nr.	Treatment 1	Treatment 2		Starting evidence						
1	A	Plac		Direct	HIGH	Not serious	Extremely serious	VERY LOW	NMA	Imprecisionx3
2	B	Plac		Direct	MODERATE	Not serious	Serious	LOW	NMA	RoB, Imprecision
3	C	Plac		Indirect	MODERATE	Not serious	Very serious	VERY LOW	NMA	RoB, Imprecisionx2
4	D	Plac		Direct	HIGH	Not serious	Extremely serious	VERY LOW	NMA	Imprecisionx3
5	A	B		Direct	HIGH	Not serious	Extremely serious	VERY LOW	NMA	Imprecisionx3
6	A	C		Indirect	HIGH	Not serious	Extremely serious	VERY LOW	NMA	Imprecisionx3
7	A	D		Indirect	HIGH	Not serious	Extremely serious	VERY LOW	NMA	Imprecisionx3
8	B	C		Direct	HIGH	Not serious	Very serious	LOW	NMA	Imprecisionx2
9	B	D		Indirect	MODERATE	Not serious	Extremely serious	VERY LOW	NMA	RoB, Imprecisionx3
10	C	D		Indirect	MODERATE	Not serious	Extremely serious	VERY LOW	NMA	RoB, Imprecisionx3



Department of Clinical Epidemiology and Biostatistics

Assessing the certainty of evidence for one comparison in a network meta-analysis

This diagram explains the process of assessing a single comparison within a network meta-analysis.

For a summary of the process for assessing certainty in a whole network, please see supplementary table 1



Comparison	Risk-of-bias assessment	Inconsistency assessment	Indirectness assessment	Publication bias assessment	Direct certainty rating to inform the network estimate*	Imprecision assessment*	Final certainty rating of direct estimate*
	Step A1 (necessary for rating the network estimate)					Step A2 (optional, not used for rating the network estimate)	
Direct estimate on systemic corticosteroids v standard of care for covid-19. Outcome: Mortality	Serious methodological limitations	No serious inconsistency	No serious indirectness	Publication bias undetected	Moderate certainty ⊕⊕⊕○	Serious imprecision	Low certainty ⊕⊕○○
	Rate down one level	Do not rate down	Do not rate down	Do not rate down		Rate down one level	



Department of Clinical Epidemiology and Biostatistics

B Indirect estimate
Assess certainty of indirect estimate



Comparison	First order loop	Start rating*			Intransitivity assessment	Direct certainty rating to inform the network estimate*	Imprecision assessment*	Final certainty rating of indirect estimate*
	Step B1 (necessary for rating the network estimate)				Step B2 (necessary for rating the network estimate)		Step B3 (optional, not used for rating the network estimate)	
Indirect estimate on systemic corticosteroids v standard of care for covid-19. Outcome: Mortality	Systemic corticosteroids → IL6 antagonist → standard of care	Systemic corticosteroids v IL6 antagonist Direct preliminary certainty rating	IL6 antagonist v standard of care Direct preliminary certainty rating	Lowest certainty of the first loop components	No serious intransitivity	Moderate certainty ⊕⊕⊕○	Very serious imprecision	Very low certainty ⊕○○○
								
		Moderate certainty ⊕⊕⊕○	High certainty ⊕⊕⊕⊕	Moderate certainty ⊕⊕⊕○	Do not rate down		Rate down two levels	



Department of Clinical Epidemiology and Biostatistics

C Network estimate
Assess certainty of network estimate



C1 Select the starting point considering dominance, information size, and direct and indirect preliminary certainty ratings

GRADE
Direct
Preliminary

GRADE
Indirect
Preliminary

C2 Assess incoherence
Incoherence

C3 Assess imprecision
Imprecision

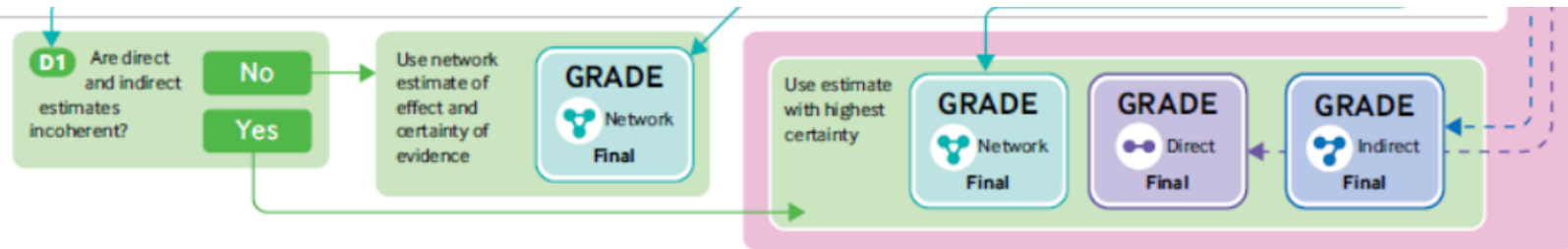
GRADE
Network
Final

Comparison	Start rating			Incoherence	Preliminary certainty rating for NMA estimate*	Imprecision assessment*	Final certainty rating of direct estimate*
	Step C1			Step C2		Step C3	
NMA estimate on systemic corticosteroids v standard of care for covid-19. Outcome: Mortality	Direct preliminary rating to inform NMA	Indirect preliminary rating to inform NMA	Direct estimate dominant	No serious incoherence	Moderate certainty ⊕⊕⊕○	No serious imprecision	Moderate certainty ⊕⊕⊕○
	Moderate certainty ⊕⊕⊕○	Moderate certainty ⊕⊕⊕○	Moderate certainty ⊕⊕⊕○	Do not rate down		Do not rate down	



Department of Clinical Epidemiology and Biostatistics

D Best estimate
Choose the best estimate of effect considering the certainty ratings



Comparison	Final certainty ratings			Best estimate and certainty rating*	
	Step D				
Systemic corticosteroids v standard of care for covid-19. Outcome: Mortality	Final direct rating†	Final indirect rating†	Final NMA rating	Moderate certainty ⊕⊕⊕○	NMA estimate selected as direct and indirect did not provide higher certainty
					
	Low certainty ⊕⊕○○	Very low certainty ⊕○○○	Moderate certainty ⊕⊕⊕○		



Conclusion

- The GRADE working group has provided guidance on how to draw conclusion from an NMA using contextualised frameworks.
- Interventions placed in higher categories are likely to be more effective than interventions placed in lower categories.
- Use of these frameworks can be especially useful in the context of large NMAs.



Department of Clinical Epidemiology and Biostatistics

Thank you

For your attention

Question and answer