Cochrane Sweden

Methods seminar: GRADE for uncertainty assessments

14 November 2024

Martin Ringsten
Cochrane Sweden

Trusted evidence.
Informed decisions.
Better health.

Cochrane Sweden

GRADE Working Group

Open research collaboration

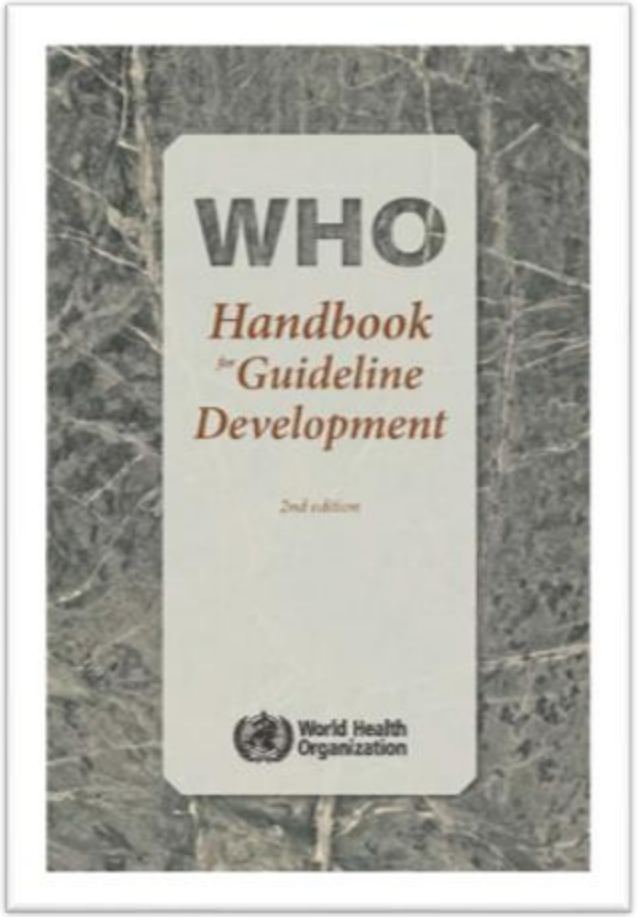
- 20 years ago
- 16 centers, 8 networks (Nordic)
- 20+ topic groups

Three methods/frameworks:

- **GRADE for "uncertainty of evidence"**
- GRADE for recommendations (GRADE Evidence-to-Decision)
- GRADE for adapting recommendations (GRADE Adolopment)

Cochrane Sweden

Use of GRADE for uncertainty



Cochrane Sweden

Domains for rating uncertainty

- Today's talk – introductory and conceptual level
  - 40+ official guidance articles
- Many domains exist elsewhere, different terminology
- Bring all uncertainty factors together to **one judgement**
- Aims to be possible to be applied them in a similar way across studies and fields

Cochrane Sweden

Domains for rating uncertainty

5 factors that can **lower** certainty

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

3 factors can **increase** certainty

1. Large magnitude of effect
2. Opposing plausible residual bias or confounding
3. Dose-response gradient

Risk of bias – primary studies and systematic reviews

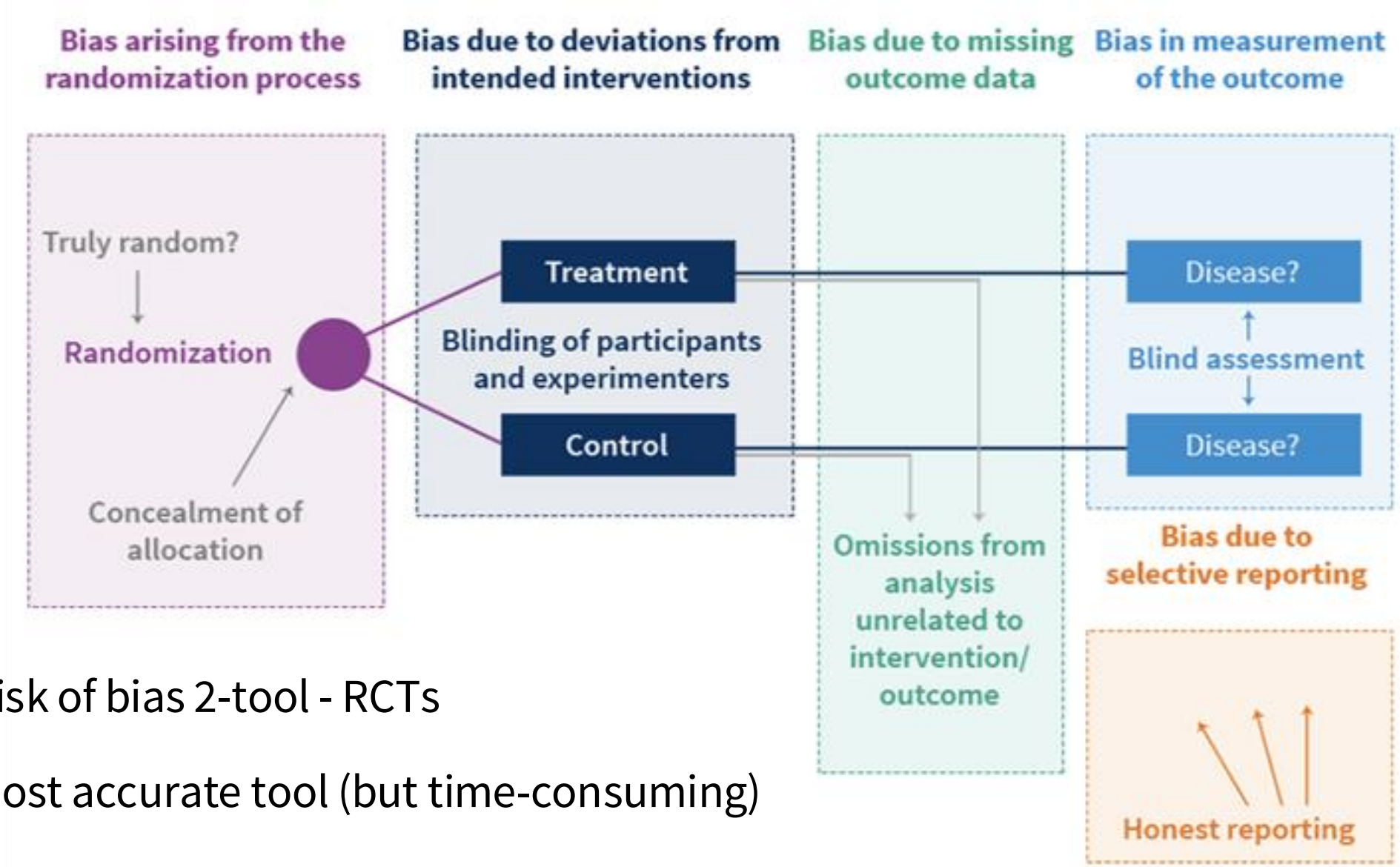
Are the research studies well done.

- Limitations in study design and/or execution that can deviate the results in one direction

Most commonly used tools:

- Risk of Bias 2- tool - RCTs
- ROBINS-I – Observational studies
- ROBINS-E - Exposure

Risk of bias – randomised studies



- Risk of bias 2-tool - RCTs
- Most accurate tool (but time-consuming)

Risk of bias – non-randomised studies

ROBINS-I

Based in the target trial framework

Switch the randomisation domain to:

- Bias due to confounding
- Bias in selection of participants into the study
- Bias in classification of interventions

Same/similar as for randomised trials:

- Bias due to deviations from intended interventions
- Bias due to missing data
- Bias in measurement of outcomes
- Bias in selection of the reported result

JAMA Guide to Statistics and Methods

December 12, 2022

Target Trial Emulation

A Framework for Causal Inference From Observation-al Data

Miguel A. Hernán, MD, DrPH[1]; Wei Wang, PhD[2]; David E. Leaf, MD, MMSc[3]

» Author Affiliations

JAMA. 2022;328(24):2446-2447. doi:10.1001/jama.2022.21383

RESEARCH METHODS AND REPORTING

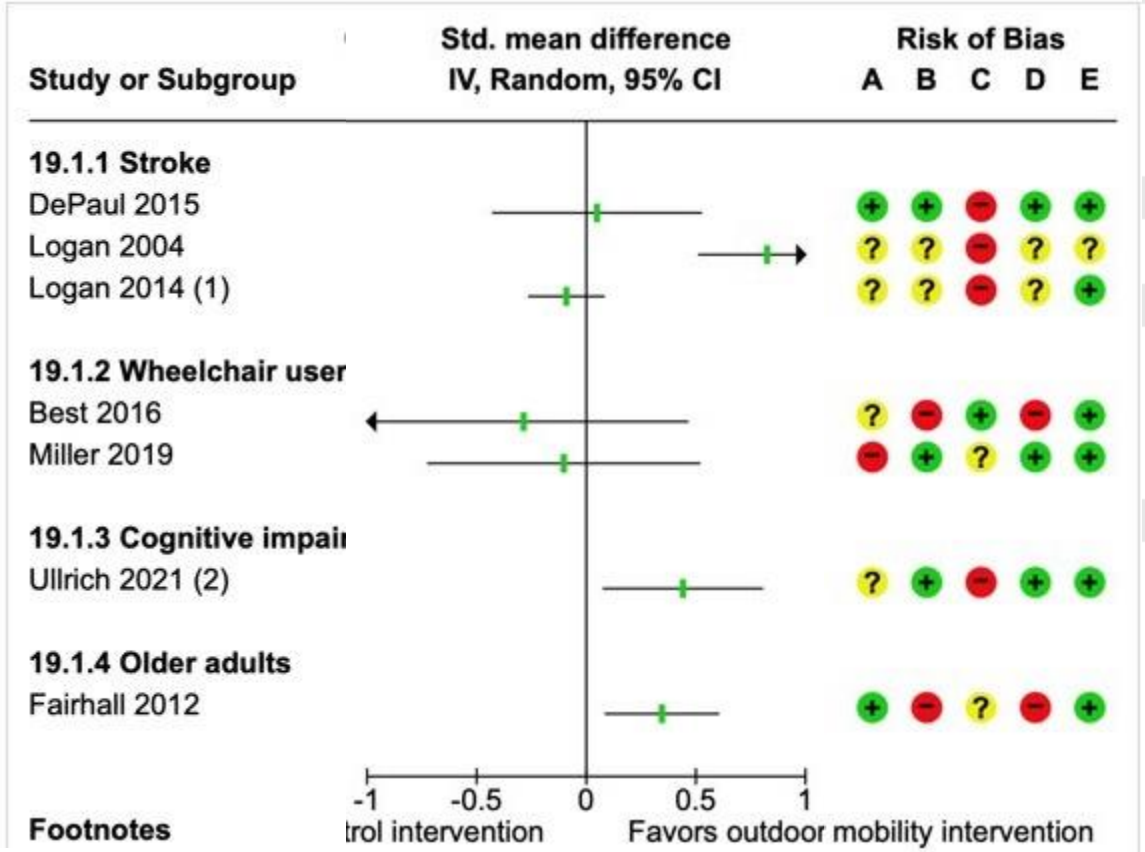
ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions

Jonathan AC Sterne,[1] Miguel A Hernán,[2] Barnaby C Reeves,[3] Jelena Savović,[1,4] Nancy D Berkman,[5] Meera Viswanathan,[6] David Henry,[7] Douglas G Altman,[8] Mohammed T Ansari,[9] Isabelle Boutron,[10] James R Carpenter,[11] An-Wen Chan,[12] Rachel Churchill,[13] Jonathan J Deeks,[14] Asbjørn Hróbjartsson,[15] Jamie Kirkham,[16] Peter Jüni,[17] Yoon K Loke,[18] Theresa D Pigott,[19] Craig R Ramsay,[20] Deborah Regidor,[21] Hannah R Rothstein,[22] Lakhbir Sandhu,[23] Pasqualina L Santaguida,[24] Holger J Schünemann,[25] Beverly Shea,[26] Ian Shrier,[27] Peter Tugwell,[28] Lucy Turner,[29] Jeffrey C Valentine,[30] Hugh Waddington,[31] Elizabeth Waters,[32] George A Wells,[33] Penny F Whiting,[34] Julian PT Higgins[35]

Risk of bias – example

- Step 1: effectiveness in acutal outdoor mobility
- Is this enough?

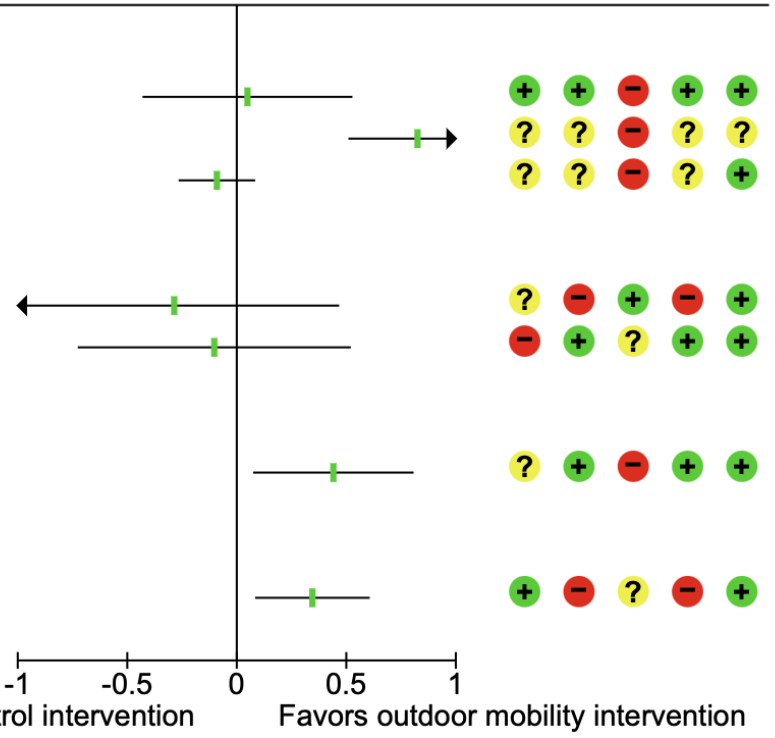
- Adding risk of bias to the equation
- Would anyone make the same choice in using/implementing these interventions?



Cochrane Sweden

Adding details to <6 months (condition and RoB)

Study or Subgroup	Outdoor mobility intervention Mean	SD	Total	Control intervention Mean	SD	Total	Std. mean difference IV, Random, 95% CI
19.1.1 Stroke							
DePaul 2015	59.08	21.41	33	58.03	21.11	34	0.05 [-0.43 , 0.53]
Logan 2004	37	32.6	86	14	21.5	82	0.83 [0.51 , 1.14]
Logan 2014 (1)	1	1	263	1.1	1.2	241	-0.09 [-0.27 , 0.08]
19.1.2 Wheelchair users							
Best 2016	47.6	24.6	16	53.7	13.8	12	-0.29 [-1.04 , 0.47]
Miller 2019	30.2	20.9	18	32.6	24.4	22	-0.10 [-0.73 , 0.52]
19.1.3 Cognitive impairments							
Ullrich 2021 (2)	32.5	10.079979	55	27.7	11.415673	63	0.44 [0.07 , 0.81]
19.1.4 Older adults							
Fairhall 2012	35.5	16.1	111	30.3	13.9	117	0.35 [0.08 , 0.61]



Footnotes

Cochrane Sweden

Inconsistency – systematic reviews (and primary studies?)

- Do the studies agree with each other, or show widely different estimates?
- Can any variation be confidently explaned, or is the variation unexplainable?
- **Your primary studies results will be compared and summarised together with other studies**

Cochrane Sweden

Inconsistency – example

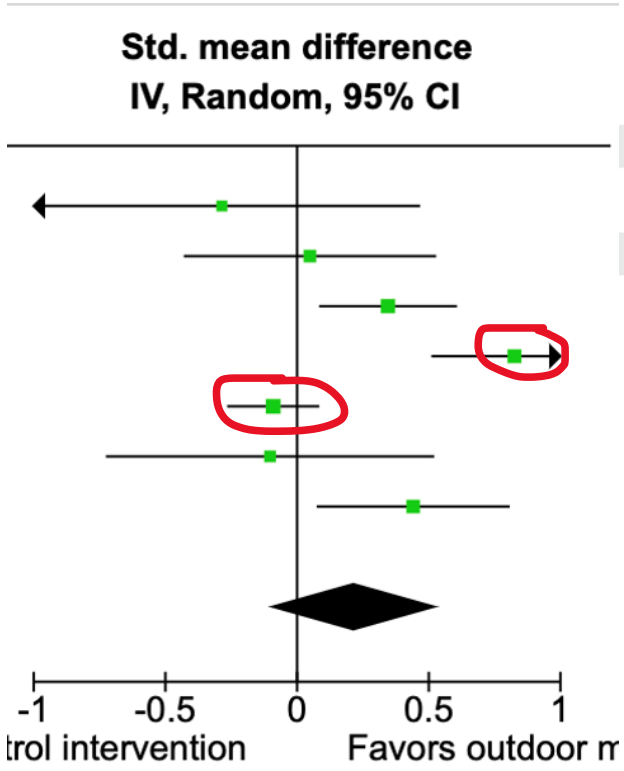
Are results consistent across studies?

Study or Subgroup	Outdoor mobility intervention Mean	SD	Total	Control intervention Mean	SD	Tot
Best 2016	47.6	24.6	16	53.7	13.8	
DePaul 2015	59.08	21.41	33	58.03	21.11	
Fairhall 2012	35.5	16.1	111	30.3	13.9	
Logan 2004	37	32.6	86	14	21.5	
Logan 2014 (1)	1	1	263	1.1	1.2	
Miller 2019	30.2	20.9	18	32.6	24.4	
Ullrich 2021 (2)	32.5	10.079979	55	27.7	11.415673	
Total (95% CI)			**582**			

Heterogeneity: Tau² = 0.12; Chi² = 31.32, df = 6 (P < 0.0001); I² = 81%

Test for overall effect: Z = 1.37 (P = 0.17)

Test for subgroup differences: Not applicable



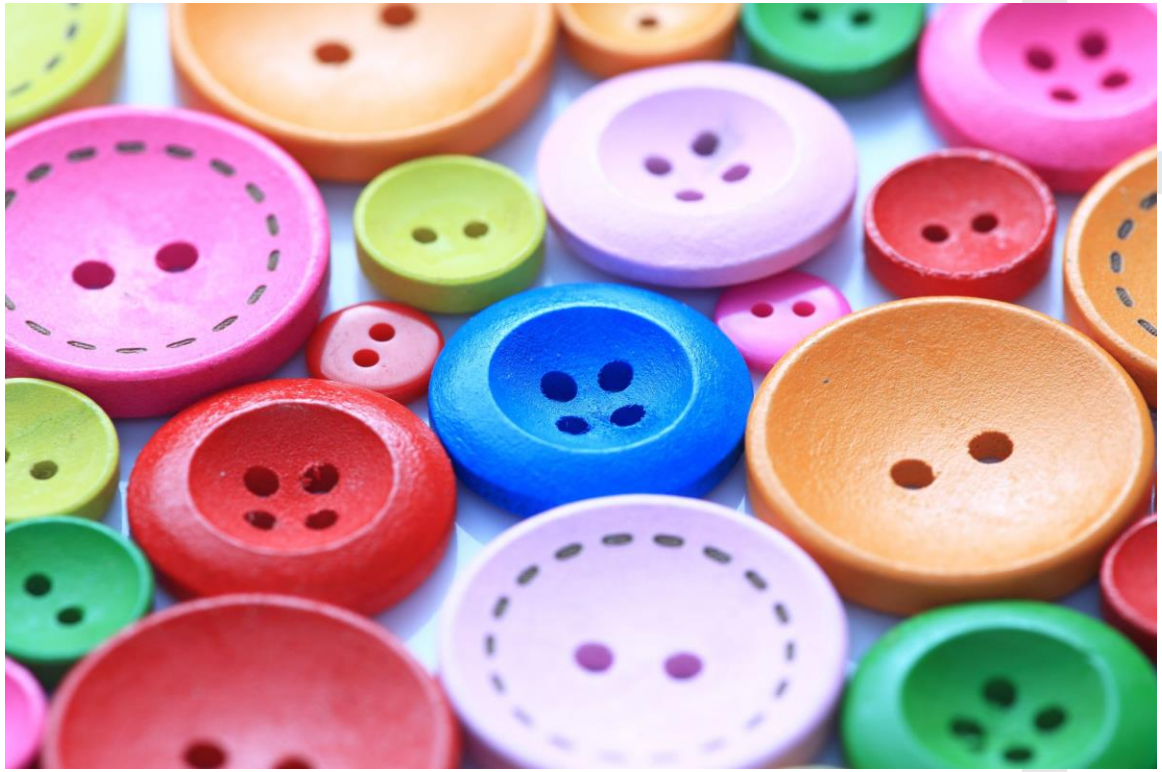
Indirectness – primary studies and systematic reviews

How directly do the results relate to the actual question that you want to answer?

Would the effects be different in studies compared to effects in the real world?

- Population
- Intervention / exposure
- Outcomes

Your primary studies will be assessed how close they are to real world context



Indirectness – example

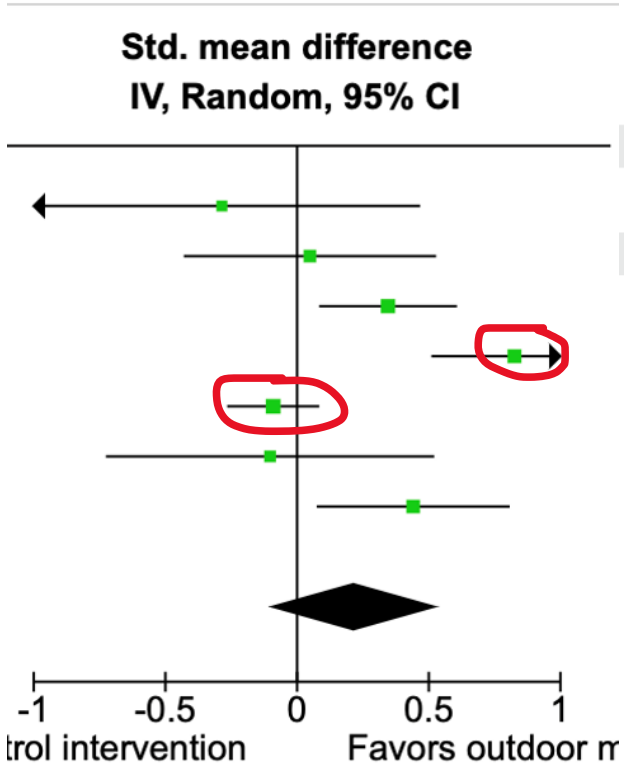
How would these interventions work in the Swedish setting?

Study or Subgroup	Outdoor mobility intervention Mean	SD	Total	Control intervention Mean	SD	Tot
Best 2016	47.6	24.6	16	53.7	13.8	
DePaul 2015	59.08	21.41	33	58.03	21.11	
Fairhall 2012	35.5	16.1	111	30.3	13.9	
Logan 2004	37	32.6	86	14	21.5	
Logan 2014 (1)	1	1	263	1.1	1.2	
Miller 2019	30.2	20.9	18	32.6	24.4	
Ullrich 2021 (2)	32.5	10.079979	55	27.7	11.415673	
Total (95% CI)			**582**			

Heterogeneity: Tau² = 0.12; Chi² = 31.32, df = 6 (P < 0.0001); I² = 81%

Test for overall effect: Z = 1.37 (P = 0.17)

Test for subgroup differences: Not applicable



Cochrane Sweden

Imprecision – example

Is the effect size precise? Do we have sufficient amount of information?

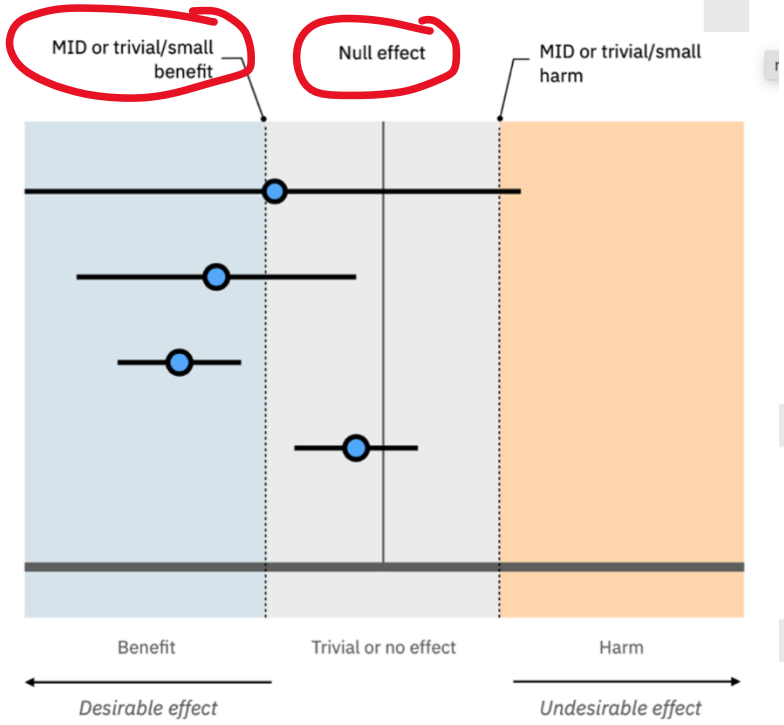
Random error:

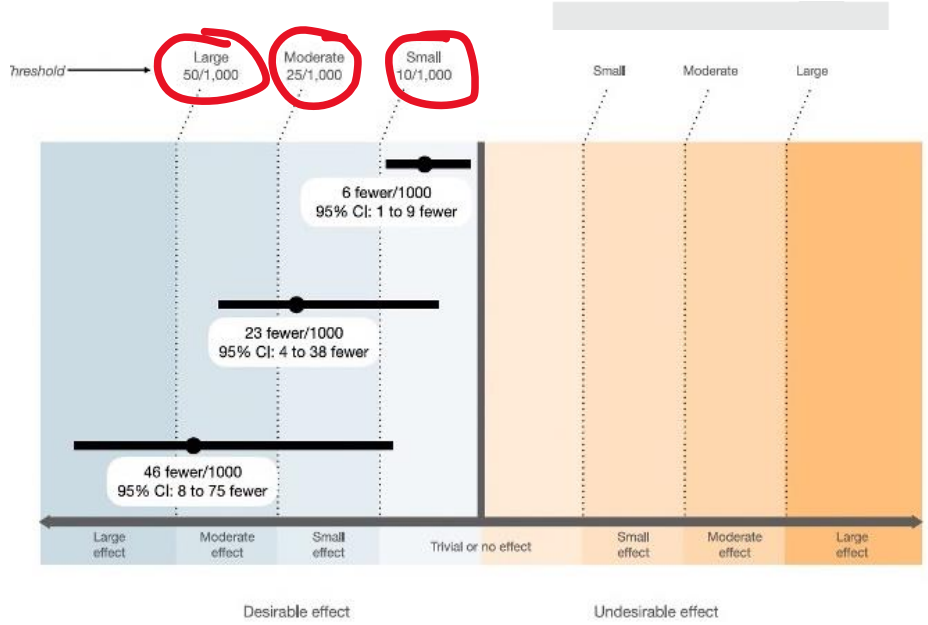
- Sample size
- Number of events / effect size

GRADE approach:

- Confidence intervals and thresholds for effect size

Your study will be assessed how precise the effects are, and if any important effect thresholds are crossed





Cochrane Sweden

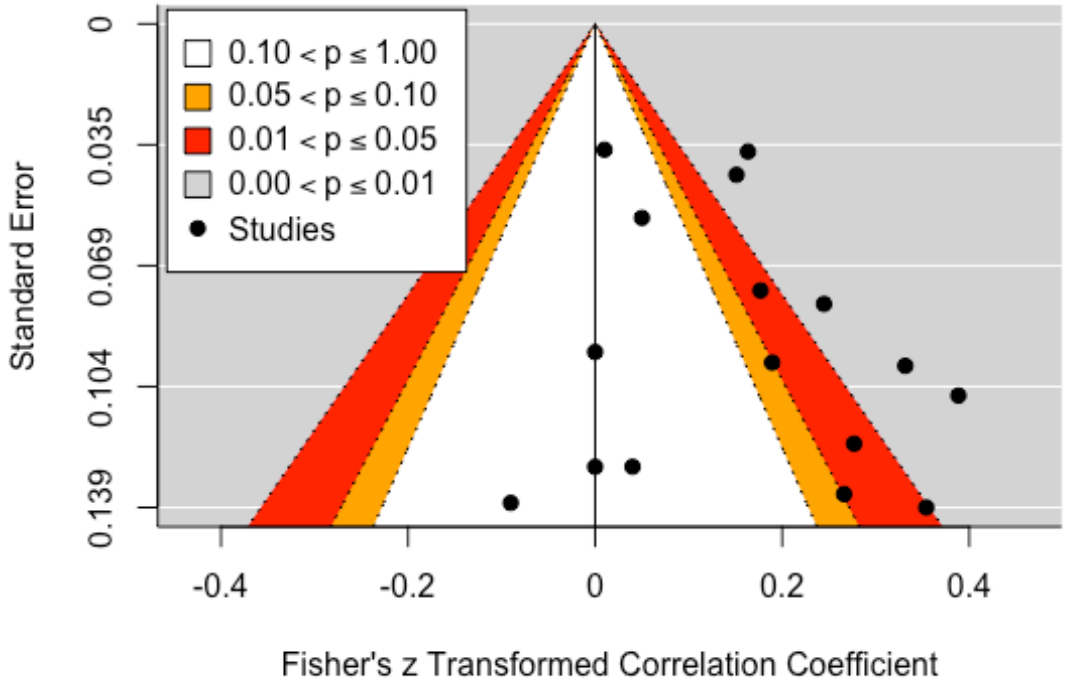
Publication/dissemination bias – systematic reviews

Have all results been reported?

Are there indications that some outcomes, or studies results have not been reported?

- Complete publications and outcomes (harms?)
- Primary studies will be assessed if they leave do not report certain results
- Funnel plots, protocols, ROB-ME tool

You study will be assessed if your outcomes are reported.

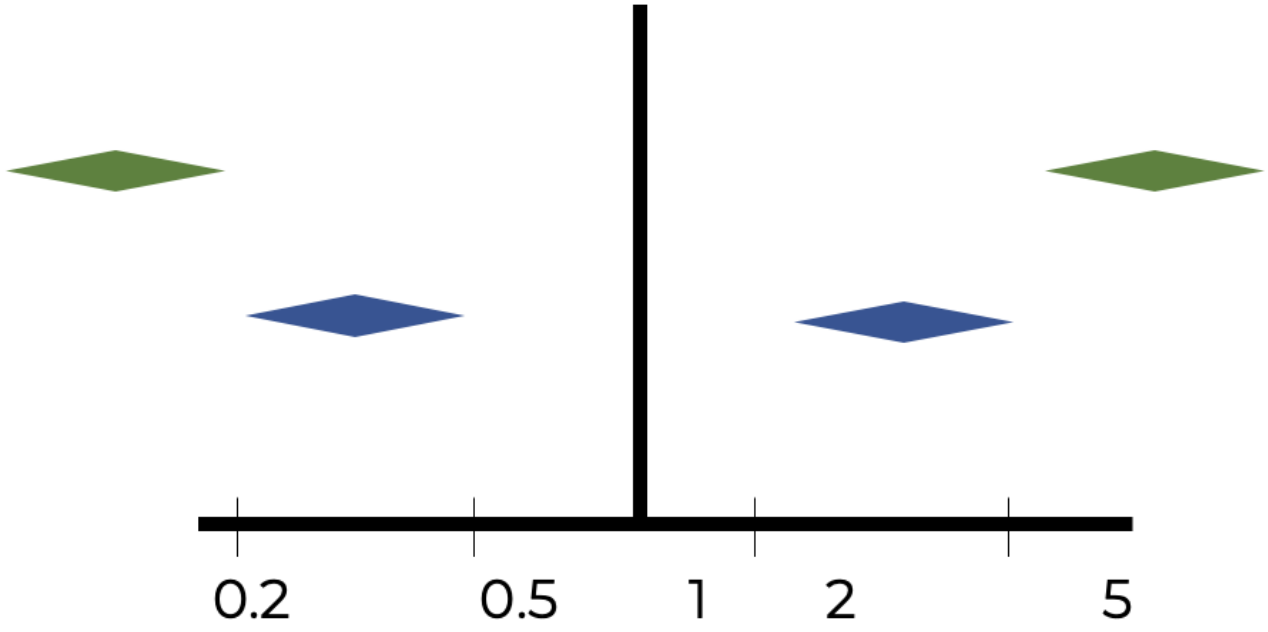


Special cases - Large effects

- Just used in observational studies
- Is the effects so large and extremely robust to any other uncertainty?

A way to rate up previous uncertainty from mainly confounding and selection bias

Your (observational) study will be assessed for how large the effects are



Special cases - Large effects - example

- Motorcycle helmets for preventing head injury
- Odds ratio 0.31

Analysis 1.3. Comparison 1 Motorcycle helmet versus no helmet, Outcome 3 Head Injury (adjusted).

Review: Helmets for preventing injury in motorcycle riders

Comparison: 1 Motorcycle helmet versus no helmet

Outcome: 3 Head Injury (adjusted)

Study or subgroup	log [Adjusted Odds Ratio] (SE)	Adjusted Odds Ratio IV,Random,95% CI	Weight	Adjusted Odds Ratio IV,Random,95% CI
1 Case-control studies				
Gabella 1995	-0.8796 (0.342)		8.9 %	0.41 [0.21, 0.81]
Tsai 1995	-1.3471 (0.3089)		10.9 %	0.26 [0.14, 0.48]
Subtotal (95% CI)			**19.8 %**	**0.32 [0.20, 0.51]**
Heterogeneity: $Tau^2 = 0.00$; $Chi^2 = 1.03$, df = 1 (P = 0.31); $I^2 = 3\%$				
Test for overall effect: Z = 4.89 (P < 0.00001)				
2 Cross-sectional studies				
Christian 2003	-1.4697 (0.2547)		15.9 %	0.23 [0.14, 0.38]
Romano 1991	-1.335 (0.2057)		23.9 %	0.26 [0.18, 0.39]
Rowland 1996	-1.1314 (0.2233)		20.4 %	0.32 [0.21, 0.50]
Sauter 2005	-0.8439 (0.2263)		19.9 %	0.43 [0.28, 0.67]
Subtotal (95% CI)			**80.2 %**	**0.30 [0.24, 0.39]**
Heterogeneity: $Tau^2 = 0.02$; $Chi^2 = 4.11$, df = 3 (P = 0.25); $I^2 = 27\%$				
Test for overall effect: Z = 8.98 (P < 0.00001)				
Total (95% CI)			**100.0 %**	**0.31 [0.25, 0.38]**
Heterogeneity: $Tau^2 = 0.00$; $Chi^2 = 5.18$, df = 5 (P = 0.39); $I^2 = 3\%$				
Test for overall effect: Z = 11.41 (P < 0.00001)				

0.01 0.1 1 10 100

Favours treatment Favours control

Cochrane Sweden

Special cases – Dose-response relationships

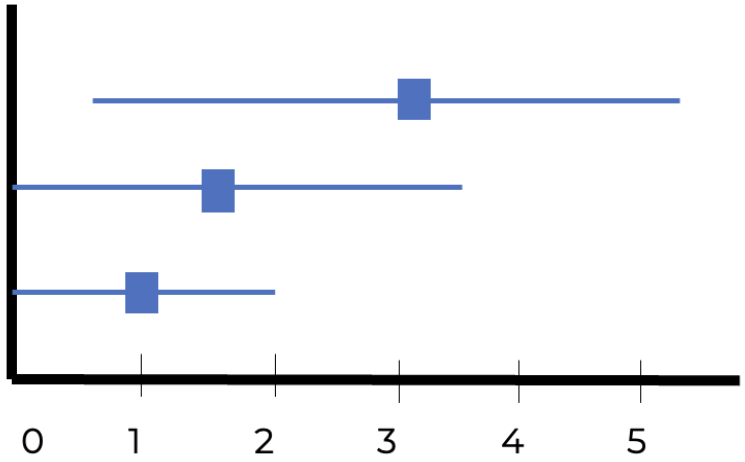
Just used in observational studies

Can we identify any exposure-effect relationship.

Strengthens any claims of effects.

But needs to be assessed for credibility.

Your (observational) study will be explored if any dose-response relatioships are presented



Start with assessing credibility

1. Appropriate analytical approach
2. Likelihood of residual confounding
3. Likelihood of ecological bias
4. Consistency across studies
5. Support by indirect evidence

Cochrane Sweden

Special cases – Plausible opposing residual confounding

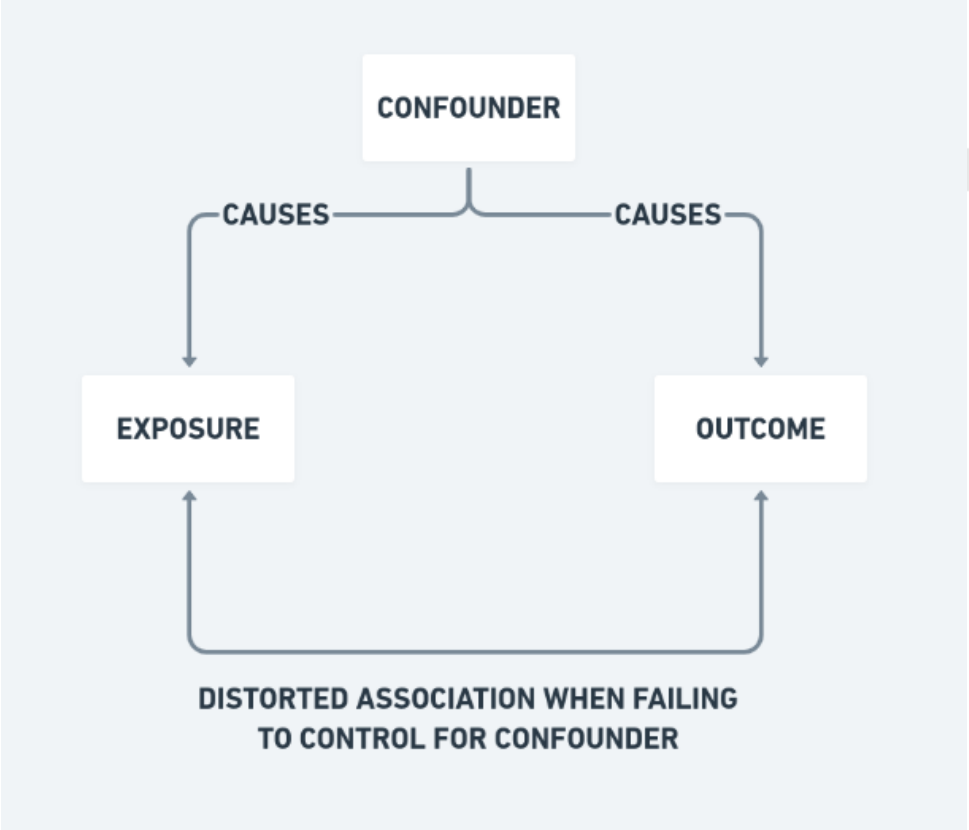
Just used in observational studies.

If you strongly believe you have confounding, which have not been accounted for, that:

1. Would reduce the effect found and there is still an effect
2. Would increase the effect found and there is still no effect

Stronger evidence.

Your (observational) study will be assessed if residual confounding is present and the direction of this confounding.



Dealing with all the ratings

Table: GRADE's approach to rating certainty/quality of evidence (aka confidence in effect estimates)

For each outcome based on a systematic review and across outcomes (lowest quality across the outcomes critical for decision making)

1. Establish initial level of certainty		2. Consider lowering or raising level of certainty		3. Final level of certainty rating
Study design	***Initial certainty in an estimate of effect***	***Reasons for considering lowering or raising certainty***		***Certainty in an estimate of effect across those considerations***
		⬇ Lower if	**⬆ Higher if***	
Randomized trials or studies evaluated with ROBINS ➔	**High certainty**	**Risk of Bias** **Inconsistency** **Indirectness** **Imprecision** **Publication bias**	**Large effect** **Dose response** **All plausible confounding & bias** • would reduce a demonstrated effect **or** • would suggest a spurious effect if no effect was observed	**High** ⊕⊕⊕⊕
				Moderate ⊕⊕⊕◯
Observational studies not using ROBINS ➔	**Low certainty**			**Low** ⊕⊕◯◯
				Very low ⊕◯◯◯

*upgrading criteria are usually applicable to observational studies only.

Cochrane Sweden

What do these ratings mean?

High certainty: We are very confident that the true effect lies close to that of the estimate of the effect (or within a range or beyond a certain threshold)

Moderate certainty: We are moderately confident in the effect estimate: The true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different

Low certainty: Our confidence in the effect estimate is limited: The true effect may be substantially different from the estimate of the effect

Very low certainty: We have very little confidence in the effect estimate: The true effect is likely to be substantially different from the estimate of effect

Cochrane Sweden

Used in GRADE EtD

- GRADE Evidence to decision frameworks incorporate this uncertainty.
- Also many other factors for making reliable and transparent decisions and receommendations.

Cochrane Sweden

GRADE in other areas

- Interventions
- Prognosis
- Network meta-analysis
- Tests (diagnosis)
- Values and preferences
- Cost
- Preclinical animal research
- Environmental health exposures

- https://training.cochrane.org/online-learning/cochrane-methodology/grade-approach/jce-series

Cochrane Sweden

Trusted evidence.
Informed decisions.
Better health.

Cochrane Sweden

Time to reflect, talk and answer questions to the best of my ability