Probability bounds analysis as a way to open up for semi-automatic quantification of bias terms in RoB-adjusted evidence synthesis

Ullrika Sahlin, Lund University, Sweden

UNITED STATES ENVIRONMENTAL PROTECTION AGENCY
WASHINGTON, D.C. 20460

FEB 26 1992

OFFICE OF
THE ADMINISTRATOR

MEMORANDUM

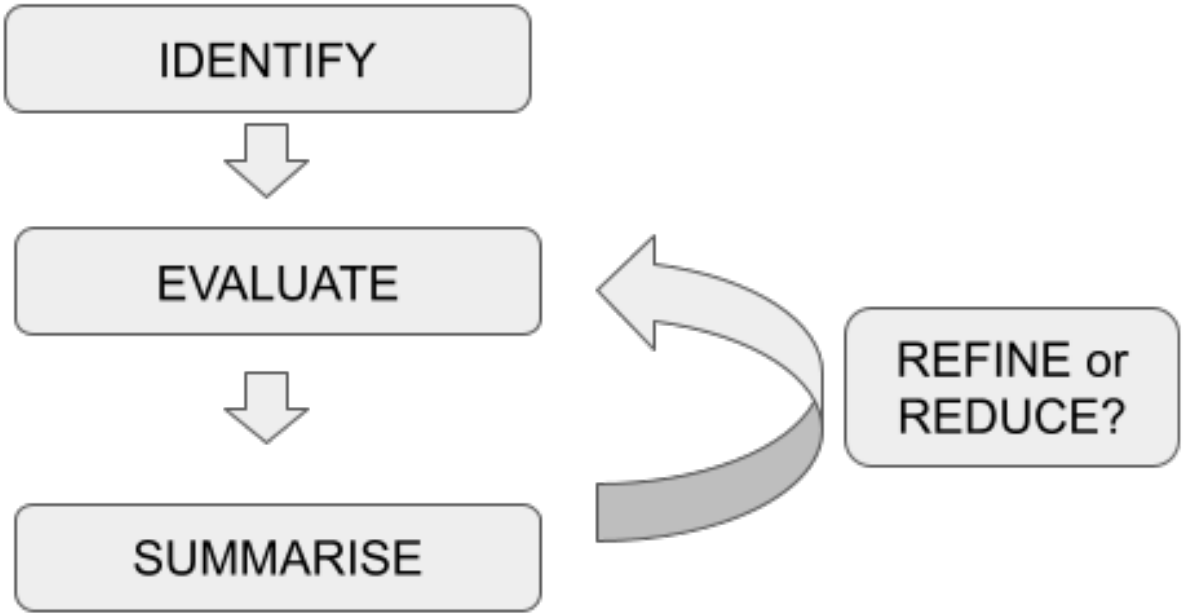
SUBJECT: Guidance on Risk Characterization for Risk Managers and Risk Assessors

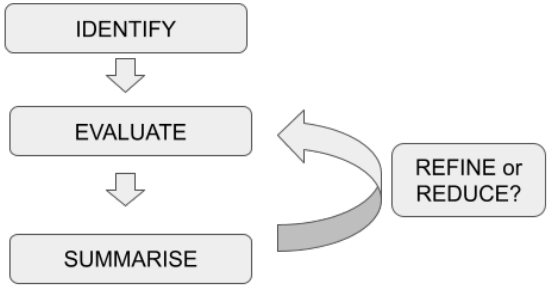
FROM: F. Henry Habicht II
Deputy Administrator

TO: Assistant Administrators
Regional Administrators

2. Regarding risk characterization, key scientific information on data and methods (e.g., use of animal or human data for extrapolating from high to low doses, use of pharmacokinetics data) must be highlighted. We also expect a statement of confidence in the assessment that identifies all major uncertainties along with comment on their influence on the assessment, consistent with guidance in the attached Appendix.

Uncertainty analysis





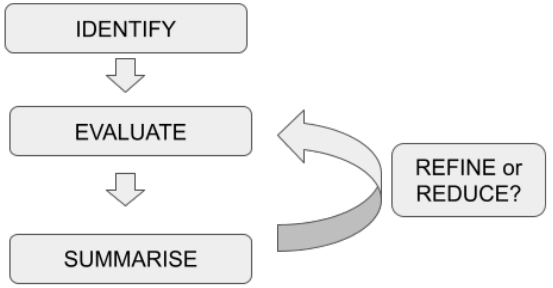
IDENTIFY

Sources of uncertainty in INPUTS & METHODOLOGIES

My favourite short list (van der Bles et al. 2019):

1. Variability within a sampled population or repeated measures leading to, for example, statistical margins-of-error
2. Computational or systematic inadequacies of measurement
3. Limited knowledge and ignorance about underlying processes
4. Expert disagreement

EVALUATE



Data, Evidence and Modelling considered together with Sources of Uncertainty

CONCLUSION REACHED BY:

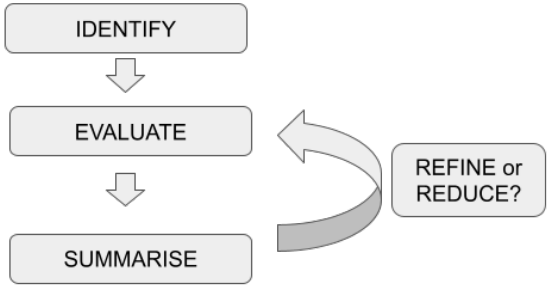
A rule

An individual judgement

A group judgement

A group judgement that confirms the rule

SUMMARISE



WHAT to be uncertain about:

- **Facts** - categorical variables that are (at least theoretically) directly verifiable
- **Numbers** - continuous variables that describe the world. They may, at least in principle, be directly observable, or they may be theoretical constructs which are used as parameters within a model of the world.
- **Scientific hypotheses** - theories about how the world works, expressed as structural models of the relationship between variables. Scientific models and hypotheses are, like parameters, not directly observable 'things', but working assumptions.

Is an agent carcinogenic or not

The dose at which there is a health effect

A possible mechanism for the effect

The form of a dose–response relationship

Van der Bles et al. 2019

Practices to communicate uncertainty occur at two levels:

- Direct uncertainty about the fact, number or scientific hypothesis. This can be communicated either in absolute quantitative terms, say a probability distribution or confidence interval, or expressed relative to alternatives, such as likelihood ratios, or given an approximate quantitative form, verbal summary and so on.

QuaNtitative QuaLitative

- Indirect uncertainty in terms of the **quality of the underlying knowledge** that forms a basis for any claims about the fact, number or hypothesis. This will generally be communicated as a list of caveats about the underlying sources of evidence, possibly amalgamated into a qualitative or ordered categorical scale.

QuaLitative

Van der Bles et al 2019

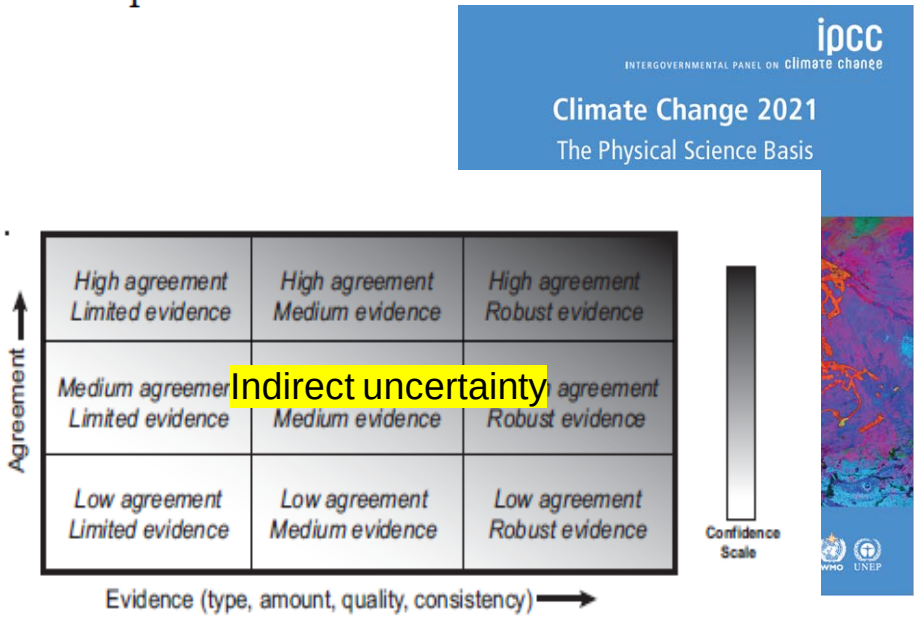
Example of levels of uncertainty communication:

A.1.6 It is *virtually certain* that the global upper ocean (0–700 m) has warmed since the 1970s and *extremely likely* that human influence is the main driver. It is *virtually certain* that human-caused CO_2 emissions are the main driver of current global acidification of the surface open ocean. There is *high confidence* that oxygen levels have dropped in many upper ocean regions since the mid-20th century, and *medium confidence* that human influence contributed to this drop.
{2.3, 3.5, 3.6, 5.3, 9.2, TS.2.4}

Table 1. Likelihood Scale

Term*	Likelihood of the Outcome
Virtually certain	99-100% probability
Very likely	90-100% probability
Likely	66-100% probability
About as likely as not	33 to 66% probability
Unlikely	0-33% probability
Very unlikely	0-10% probability
Exceptionally unlikely	0-1% probability

Direct uncertainty



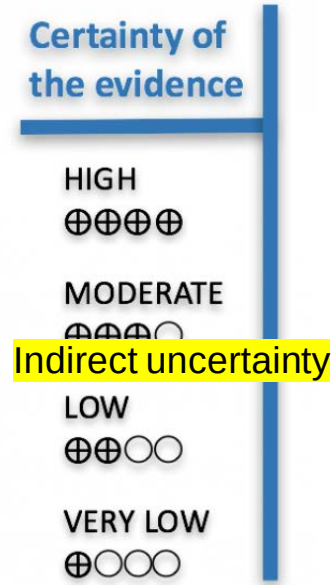
Indirect uncertainty

Example of levels of uncertainty communication:

Direct uncertainty

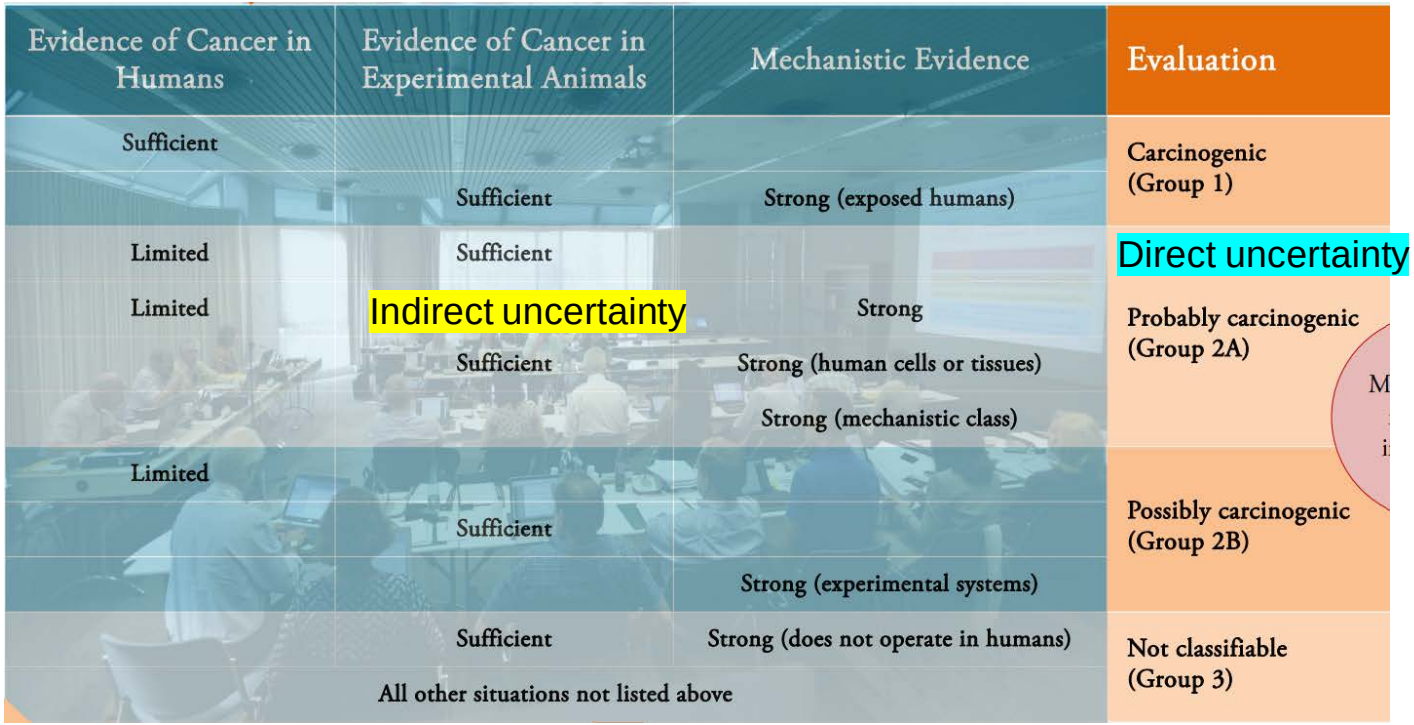
Photobiomodulation compared to Placebo for Fractures

Outcomes	Anticipated absolute effects* (95% CI)		Relative effect (95% CI)	№ of participants (studies)	Certainty of the evidence (GRADE)
	Risk with Placebo	**Risk with LLIT**			
Pain intensity (VAS scale; 0 to 10) Follow-up range: 1 day to 2 weeks	The mean pain intensity was **4.15** points	The mean pain intensity in the intervention group was 1.19 points higher (range, 0.61 to 1.77 higher)	-	106 (2 RCTs)	⨁◯◯◯ VERY LOW [a,b,c]
Radiographic signs of bone healing (Absent fracture line) Follow-up: after 2 weeks of treatment	1.000 per 1.000	**1000 per 1.000** (930 to 1.000)	**RR 1.00** (0.93 to 1.08)	50 (1 RCT)	⨁⨁◯◯ LOW [a,c]
Radiographic signs of bone healing (Callus formation) Follow-up: after 2 weeks of treatment	40 per 1.000	**13 per 1.000** (0 to 312)	**RR 0.33** (0.01 to 7.81)	50 (1 RCT)	⨁⨁◯◯ LOW [a,c]



Cochrane

Example of levels of uncertainty communication:

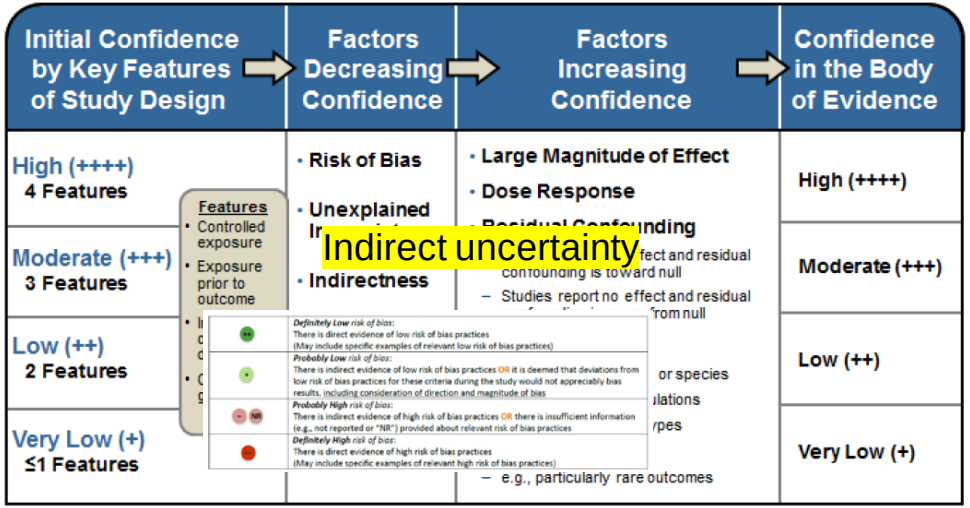


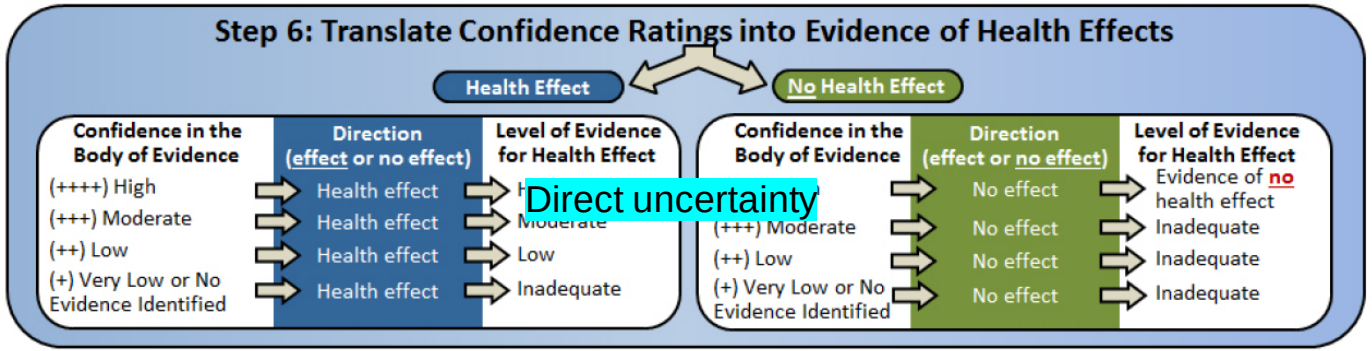
Evidence of Cancer in Humans	Evidence of Cancer in Experimental Animals	Mechanistic Evidence	Evaluation
Sufficient			Carcinogenic (Group 1)
	Sufficient	Strong (exposed humans)	
Limited	Sufficient		Probably carcinogenic (Group 2A)
Limited		Strong	
	Sufficient	Strong (human cells or tissues)	
		Strong (mechanistic class)	
Limited			Possibly carcinogenic (Group 2B)
	Sufficient		
		Strong (experimental systems)	
	Sufficient	Strong (does not operate in humans)	Not classifiable (Group 3)
All other situations not listed above			

Indirect uncertainty

Direct uncertainty

Example of levels of uncertainty communication:





Integration

QuaLIitative

QuaNtitative

QuaLitative

- Summarise side by side

  e.g. Cochrane, IPPC

Indirect uncertainty

Direct uncertainty

- Let indirect uncertainty guide how to communicate direct uncertainty

  e.g. IPPC

Indirect uncertainty

Direct uncertainty

- Integrate indirect uncertainty into the direct uncertainty, communicate direct uncertainty

  e.g. IARC, OHAT

Indirect uncertainty

Options to consider RoB in individual streams

1. Remove studies with a high risk of bias and conduct the analysis with the best available evidence (i.e., high quality studies)
2. Evaluate using sensitivity analysis the influence of including studies of lower quality in the meta-analysis
3. Include all (or a selection of) studies, but adjust for bias

Quantitative bias modelling (bias-adjusted meta-analysis) Turner et al. 2009, Lash et al. 2014

Bias modelling supports synthesis of different type of streams

Internal Validity ("Risk of Bias")

Risk of Bias Question	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	Study 14	Study 15	Study 16	Study 17	Study 18	Study 19
Randomization	+	–	++	++	–	++	+	+	++	–	–	–	+	+	+	–	–	+	++
Allocation concealment	–	–	–	–	–	–	+	–	–	–	–	–	–	–	–	–	–	–	–
Confounding (design/analysis)	++	+	++	++	++	+	++	++	++	++	+	++	++	+	–	–	–	–	++
Unintended exposure	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
Identical experimental conditions	++	++	+	+	++	++	++	++	++	+	++	+	++	++	++	++	++	++	++
Adhere to protocol	+	+	+	+	–	+	+	+	+	+	+	+	+	+	+	+	+	+	+
Blinding of researchers during study	–	–	–	–	–	–	+	–	–	–	–	–	–	–	–	–	–	–	–
Missing outcome data	–	+	++	++	--	–	+	–	–	+	--	–	–	+	++	+	++	+	++
Assessment of confounding variables	+	+	++	++	++	–	+	+	++	++	+	+	+	++	++	–	+	+	++
Exposure characterization	++	–	+	+	–	–	+	+	–	–	–	+	+	+	+	+	+	–	+
Outcome assessment	+	+	+	+	+	+	++	+	+	–	++	+	+	+	+	+	+	+	+
Blinding of outcome assessors	+	+	+	+	++	+	+	+	+	+	+	+	--	+	++	+	+	+	+
Outcome reporting	+	+	+	++	--	+	+	+	+	–	+	+	--	+	+	+	++	–	+

Key:

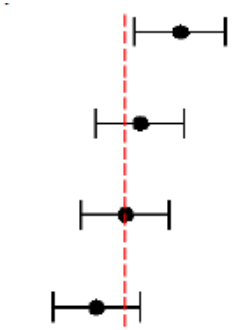
Definitely low risk of bias ++

Probably low risk of bias +

Probably high risk of bias –

Definitely high risk of bias --

Quantitative evidence synthesis

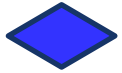


$$q_i = \frac{\sigma^2_{het}}{\sigma^2_{het} + \sigma^2_{bias,i}}$$

"Quality weight"

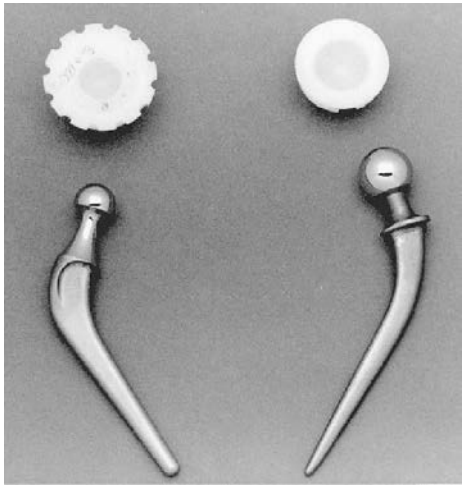
Not biased adjusted effect estimate

Biased adjusted effect estimate



Example: Evidence synthesis for comparison of revision rates

- The Swedish Hip **Registry** provides non-randomized data submitted from all hospitals in Sweden from 1979, with record linkage to further procedures and death. Nine-year follow-up results are used for around 30 000 Charnley and Stanmore prostheses.
- A U.K. Randomized Controlled Trial (**RCT**) randomized around 400 patients to Charnley or Stanmore and reported a mean follow-up of 6.5 years.
- A **Case Series** of around 1200 patients in a single hospital with a mean follow-up of 8 years.



Charnley and Stanmore hip replacement

Spiegelhalter and Best, 2003 using information from a NICE review

Example: Evidence synthesis for comparison of revision rates

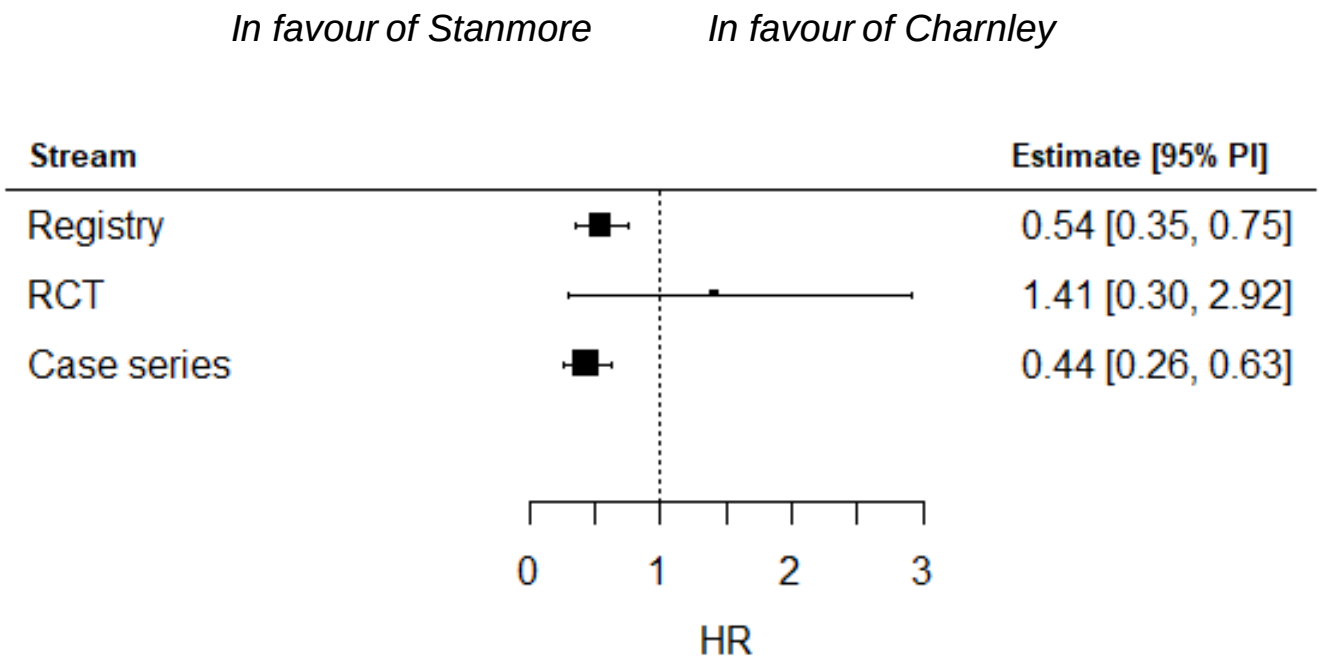
Table IV. Summary of evidence on revision hazards for Charnley and Stanmore prostheses: hazard ratios < 1 are in favour of Stanmore.

Source	Charnley: Number of patients	Charnley: Revision rate	Stanmore: Number of patients	Stanmore: Revision rate	Estimated hazard ratio (HR)	Estimated hazard ratio (95% int.)
					Fixed-effects model	
Registry	28 525	5.9%	865	3.2%	0.55	(0.37–0.77)
RCT	200	3.5%	213	4.0%	1.34	(0.45–3.46)
Case series	208	16.0%	982	7.0%	0.44	(0.28–0.66)
					Common-effect model	
					0.52	(0.39–0.67)
Quality weights [registry, RCT, case series]					*Random-effects model*	
				[1, 1, 1]	0.54	(0.37–0.78)
				[0.5, 1, 0.2]	0.61	(0.36–0.98)
				[0.1, 1, 0.05]	0.82	(0.36–1.67)

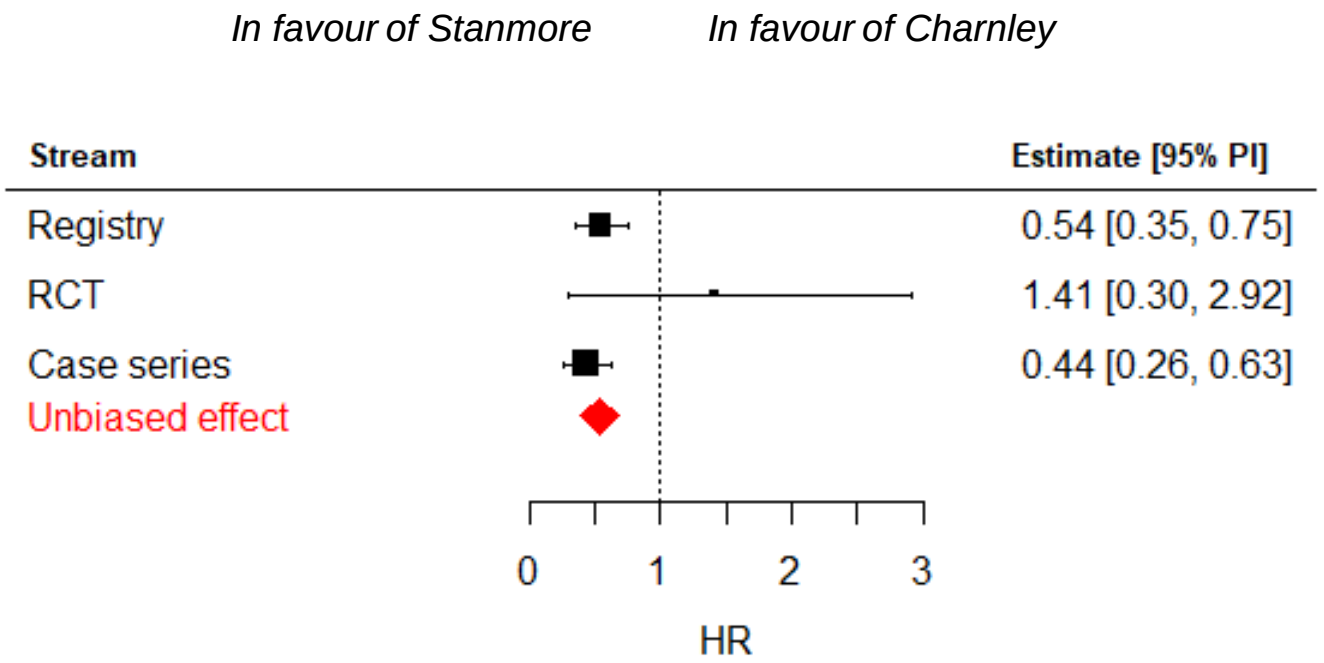
$$q_i = \frac{\sigma^2_{het}}{\sigma^2_{het} + \sigma^2_{bias,i}}$$

Spiegelhalter and Best, 2003

Example: Quantitative evidence synthesis



Example: Quantitative evidence synthesis



Example: Quantitative evidence synthesis

Formulate a hierarchical model combining all streams

$$logHR_i \sim N(logHR, \sigma^2_{het} + \sigma^2_{bias,i})$$

Make judgements or assumptions about magnitude of heterogeneity, e.g. as a prior distribution on σ^2_{het} and values and relations between the bias factors $q_i = \frac{\sigma^2_{het}}{\sigma^2_{het} + \sigma^2_{bias,i}}$

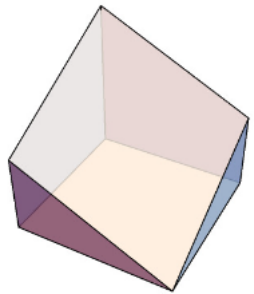
Find bounds on uncertainty in quantities of interest by optimisation under constraints defined by the polyhedron

$$q_{Case} \leq q_{Registry} \leq q_{RCT}$$

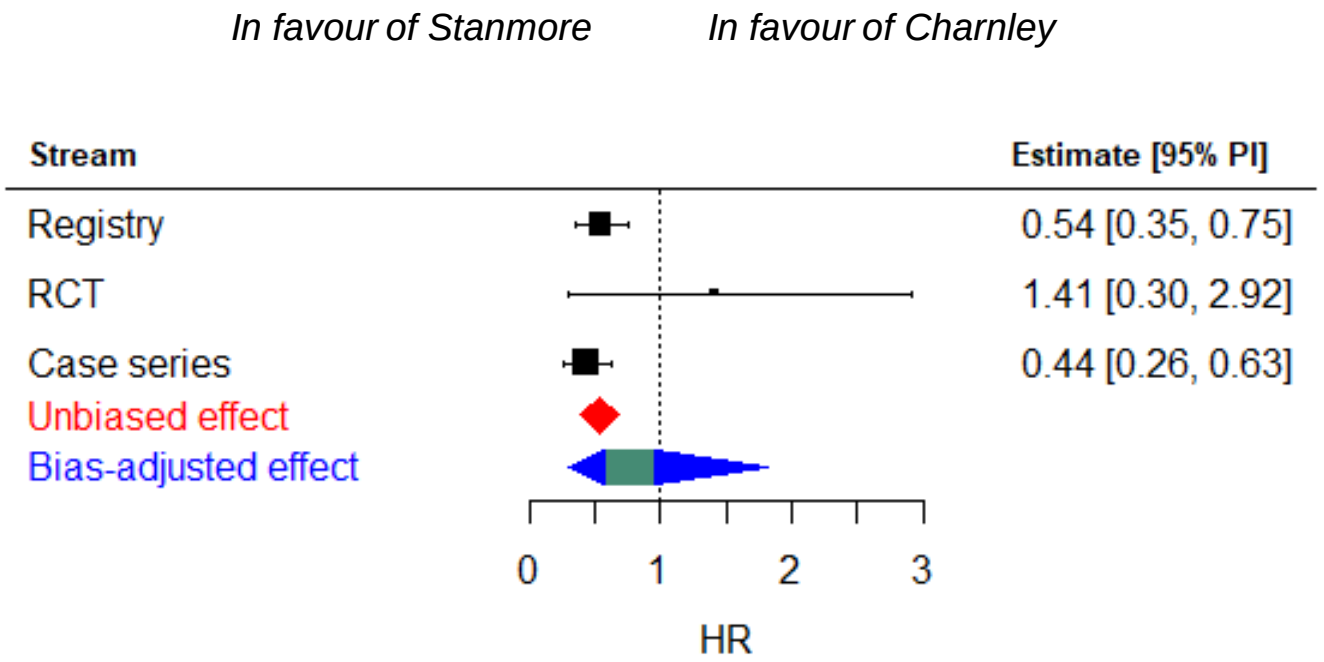
$$0.5 \leq q_{RCT} \leq 1$$

$$0.1 \leq q_{Registry}$$

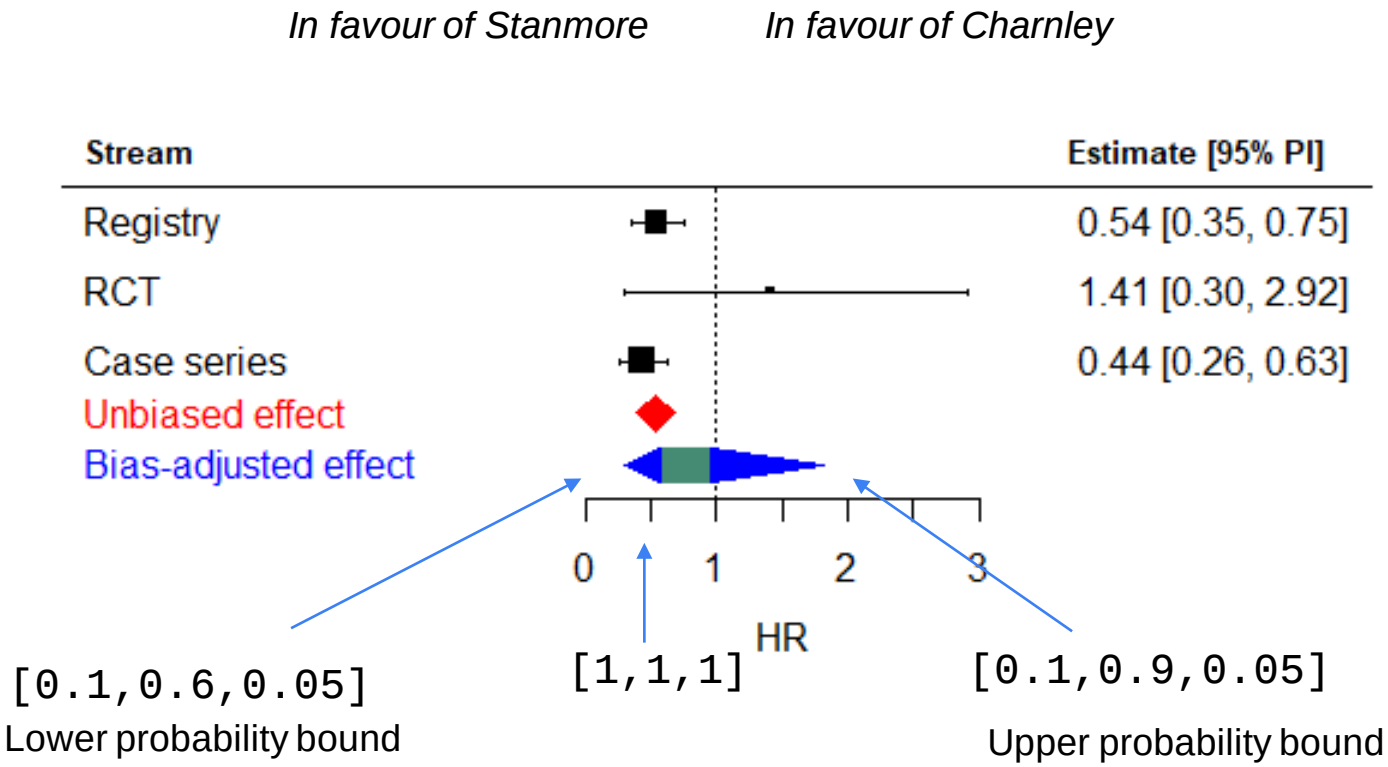
$$0.05 \leq q_{Case}$$



Example: Quantitative evidence synthesis



Example: Quantitative evidence synthesis



Example 2. Bias adjustment over different domains

Cochrane Library

Trusted evidence.
Informed decisions.
Better health.

Title Abstract

Cochrane Reviews | Trials | Clinical Answers | About

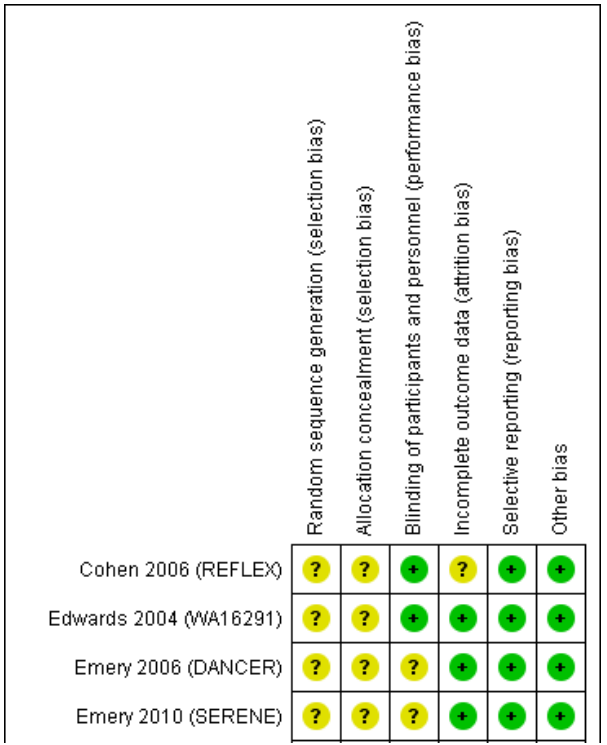
Cochrane Database of Systematic Reviews | Review - Intervention

Rituximab for rheumatoid arthritis

Maria Angeles Lopez-Olivo, Matxalen Amezaga Urruela, Lynda McGahan, Eduardo N Pollono, Maria E Suarez-Almazor Authors' declarations of interest

Version published: 20 January 2015 Version history

https://doi.org/10.1002/14651858.CD007356.pub2



Example 2. Bias adjustment over different domains

Cochrane Library

Trusted evidence.
Informed decisions.
Better health.

Title Abstract

Cochrane Reviews | Trials | Clinical Answers | About

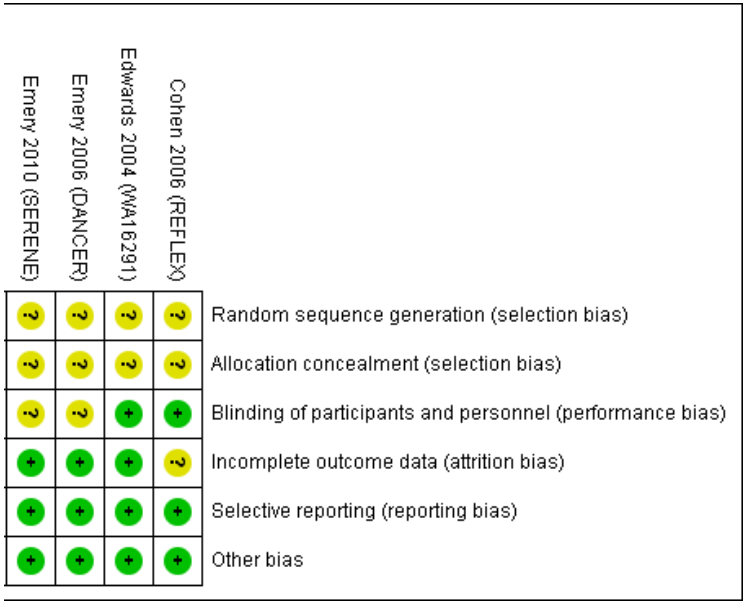
Cochrane Database of Systematic Reviews | Review - Intervention

Rituximab for rheumatoid arthritis

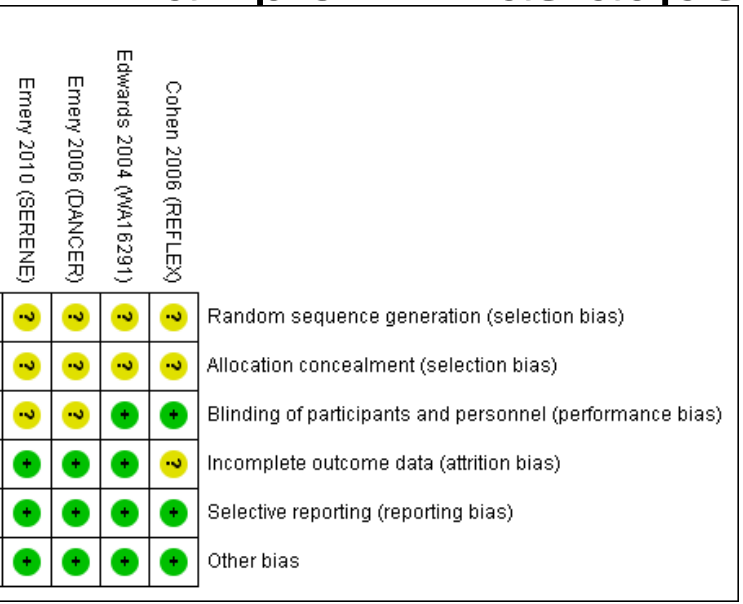
Maria Angeles Lopez-Olivo, Matxalen Amezaga Urruela, Lynda McGahan, Eduardo N Pollono, Maria E Suarez-Almazor Authors' declarations of interest

Version published: 20 January 2015 Version history

https://doi.org/10.1002/14651858.CD007356.pub2

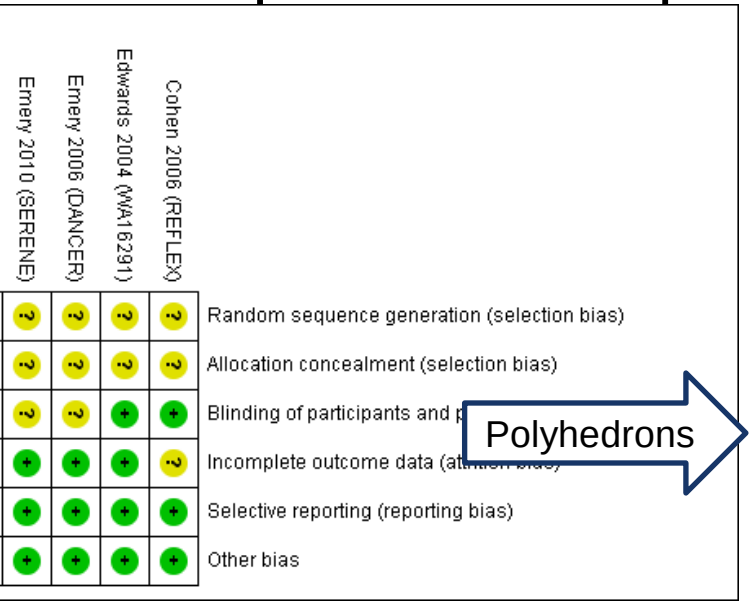


Example 2. Bias adjustment over different domains

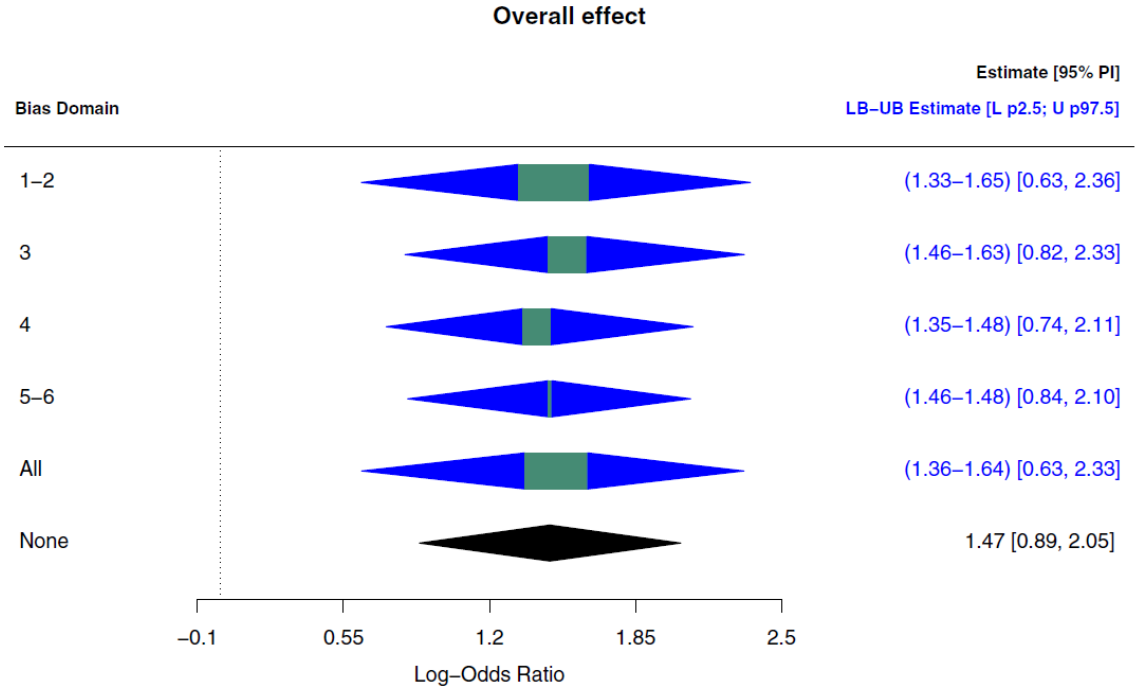


Raizes-Cruz et al., 2020

Example 2. Bias adjustment over different domains



Polyhedrons



Raizes-Cruz et al., 2020

Summary

- Bias modelling make it possible to integrate indirect uncertainty into direct uncertainty, quantitatively
- It requires a statistical model for evidence synthesis
- It requires judgements on variances and bias factors
- I have presented a way to use information about risk of bias in bias adjusted quantitative evidence synthesis
- Comments and suggestions are welcome!

Thank you!

References

- Raices Cruz, I., Troffaes, M. C., Lindström, J., & Sahlin, U. (2022). A robust Bayesian bias-adjusted random effects model for consideration of uncertainty about bias terms in evidence synthesis. Statistics in Medicine, 41(17), 3365-3379
- Spiegelhalter, D. J., & Best, N. G. (2003). Bayesian approaches to multiple sources of evidence and uncertainty in complex cost-effectiveness modelling. *Statistics in medicine*, *22*(23), 3687-3709.
- Van Der Bles, A. M., Van Der Linden, S., Freeman, A. L., Mitchell, J., Galvao, A. B., Zaval, L., & Spiegelhalter, D. J. (2019). Communicating uncertainty about facts, numbers and science. *Royal Society open science*, *6*(5), 181870.

References

- Turner RM, Spiegelhalter DJ, Smith GCS, Thompson SG. Bias modelling in evidence synthesis. J Royal Stat Soc Ser A (Stat Soc). 2009;1:21-47. doi:10.1111/j.1467-985X.2008.00547.x
- Lash TL, Fox MP, MacLehose RF, Maldonado G, McCandless LC, Greenland S. Good practices for quantitative bias analysis. Int J Epidemiol. 2014;43(6):1969-1985. doi:10.1093/ije/dyu149
- NICE Appraisal Group. The effectiveness and cost effectiveness of different prostheses for primary total hip replacement. Technical Report, http://www.nice.org.uk 2000.