United States Environmental Protection Agency

Session 1: Panel Discussion Overview of GRADE

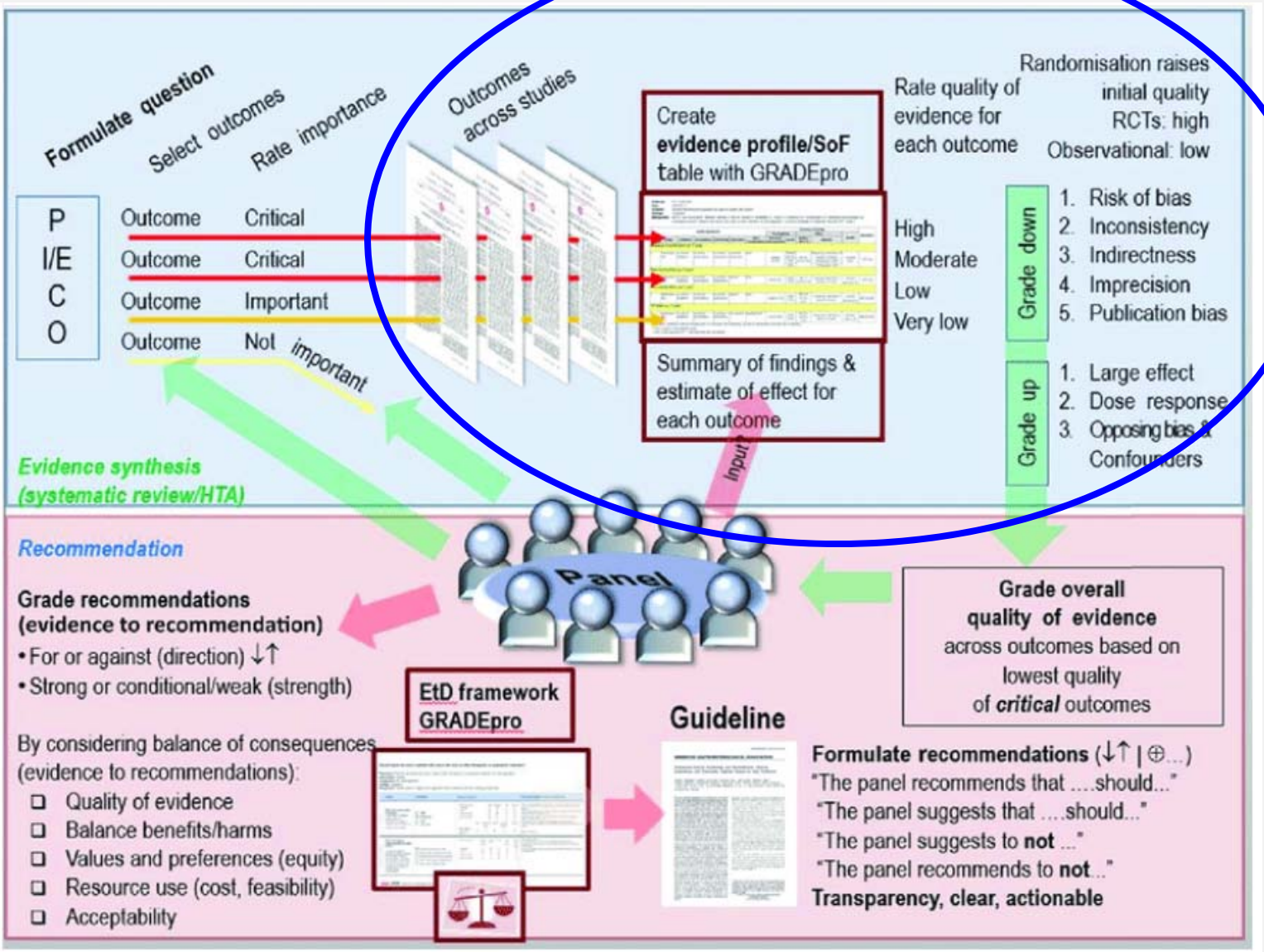
Kris Thayer, National Center for Environmental Assessment (NCEA) Integrated Risk Information System (IRIS) Division Director

NAS Workshop "Evidence integration in Chemical Assessments: Challenges Faced in Developing and Communicating Human Health Effect Conclusions"
June 3-4, 2019 (Washington DC)

GRADE

- Structured frameworks to (a) document certainty in evidence across study judgements, i.e., weight of evidence, strength of evidence, evidence synthesis/integration; and (b) evidence to decision making
- Widely used (100+ organizations from 19 countries)
- GRADE Working Group develops guidance and conducts research
  - Publications (guidance and method research), handbook, software application (GRADEpro Guideline Development Tool), bi-annual meetings, use of case examples to address methodological challenges
  - GRADE Working Group has open and free membership www.gradeworkingroup.org
  - GRADE Environmental Health Project Group established 2015 (Morgan et al., Environ Int. 2016 Jul-Aug;92-93:611-6)

United States Environmental Protection Agency



The process of assessing the quality of evidence with GRADE approach. EtD = Evidence to Decision; GRADE = Grading of Recommendations, Assessment, Development, and Evaluation; HTA = Health technology assessment; RCT = randomised controlled trials; SoF = Summary of Findings.

EPA United States Environmental Protection Agency

PRISMA-P (Preferred Reporting Items for Systematic review and Meta-Analysis Protocols) 2015 checklist: recommended items to address in a systematic review protocol

Section and topic	Item No	Checklist item
INTRODUCTION		
Rationale	6	Describe the rationale for the review in the context of what is already known
Objectives	7	Provide an explicit statement of the question(s) the review will address with reference to participants, interventions, comparators, and outcomes (PICO)
METHODS		
Eligibility criteria	8	Specify the study characteristics (such as PICO, study design, setting, time frame) and report characteristics (such as years considered, language, publication status) to be used as criteria for eligibility for the review
Information sources	9	Describe all intended information sources (such as electronic databases, contact with study authors, trial registers or other grey literature sources) with planned dates of coverage
Search strategy	10	Present draft of search strategy to be used for at least one electronic database, including planned limits, such that it could be repeated
Study records:		
Data management	11a	Describe the mechanism(s) that will be used to manage records and data throughout the review
Selection process	11b	State the process that will be used for selecting studies (such as two independent reviewers) through each phase of the review (that is, screening, eligibility and inclusion in meta-analysis)
Data collection process	11c	Describe planned method of extracting data from reports (such as piloting forms, done independently, in duplicate), any processes for obtaining and confirming data from investigators
Data items	12	List and define all variables for which data will be sought (such as PICO items, funding sources), any pre-planned data assumptions and simplifications
Outcomes and prioritization	13	List and define all outcomes for which data will be sought, including prioritization of main and additional outcomes, with rationale
Risk of bias in individual studies	14	Describe anticipated methods for assessing risk of bias of individual studies, including whether this will be done at the outcome or study level, or both; state how this information will be used in data synthesis
Data synthesis	15a	Describe criteria under which study data will be quantitatively synthesised
	15b	If data are appropriate for quantitative synthesis, describe planned summary measures, methods of handling data and methods of combining data from studies, including any planned exploration of consistency (such as I^2, Kendall's τ)
	15c	Describe any proposed additional analyses (such as sensitivity or subgroup analyses, meta-regression)
	15d	If quantitative synthesis is not appropriate, describe the type of summary planned
Meta-bias(es)	16	Specify any planned assessment of meta-bias(es) (such as publication bias across studies, selective reporting within studies)
Confidence in cumulative evidence	17	Describe how the strength of the body of evidence will be assessed (such as GRADE)

From: Shamseer L, Moher D, Clarke M, Ghersi D, Liberati A, Petticrew M, Shekelle P, Stewart L, PRISMA-P Group. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015: elaboration and explanation. BMJ. 2015 Jan 2;349(jan02 1):g7647.

Certainty in the Evidence: How Confident in the Research

- Are the research studies well done? Risk of bias
- Are the results consistent across studies ? Inconsistency
- How directly do the results relate to the question? Indirectness
- Is the association precise - due to random error? Imprecision
- Are these all of the studies that have been conducted? Publication Bias
- Is there anything else that makes us particularly certain? Large associations, worst case scenario predictors still allows strong conclusions, dose-response gradient

United States Environmental Protection Agency

JECH Online First, published on October 14, 2010 as 10.1136/jech.2010.119933

Essay

The GRADE approach and Bradford Hill's criteria for causation

Holger Schünemann,[1] Suzanne Hill,[2] Gordon Guyatt,[1] Elie A Akl,[3] Faruque Ahmed[4]

[1]Departments of Clinical Epidemiology & Biostatistics and of Medicine, McMaster University Health Sciences Centre, Hamilton, Ontario, Canada
[2]World Health Organization, Geneva, Switzerland
[3]State University of New York, Buffalo, New York, USA
[4]Centers for Disease Control and Prevention, Atlanta, Georgia, USA

Correspondence to
Holger J Schünemann, Department of Clinical Epidemiology & Biostatistics, McMaster University Health Sciences Centre, Room 2C10B, 1200 Main Street West, Hamilton, ON L8N 3Z5, Canada; schuneh@mcmaster.ca

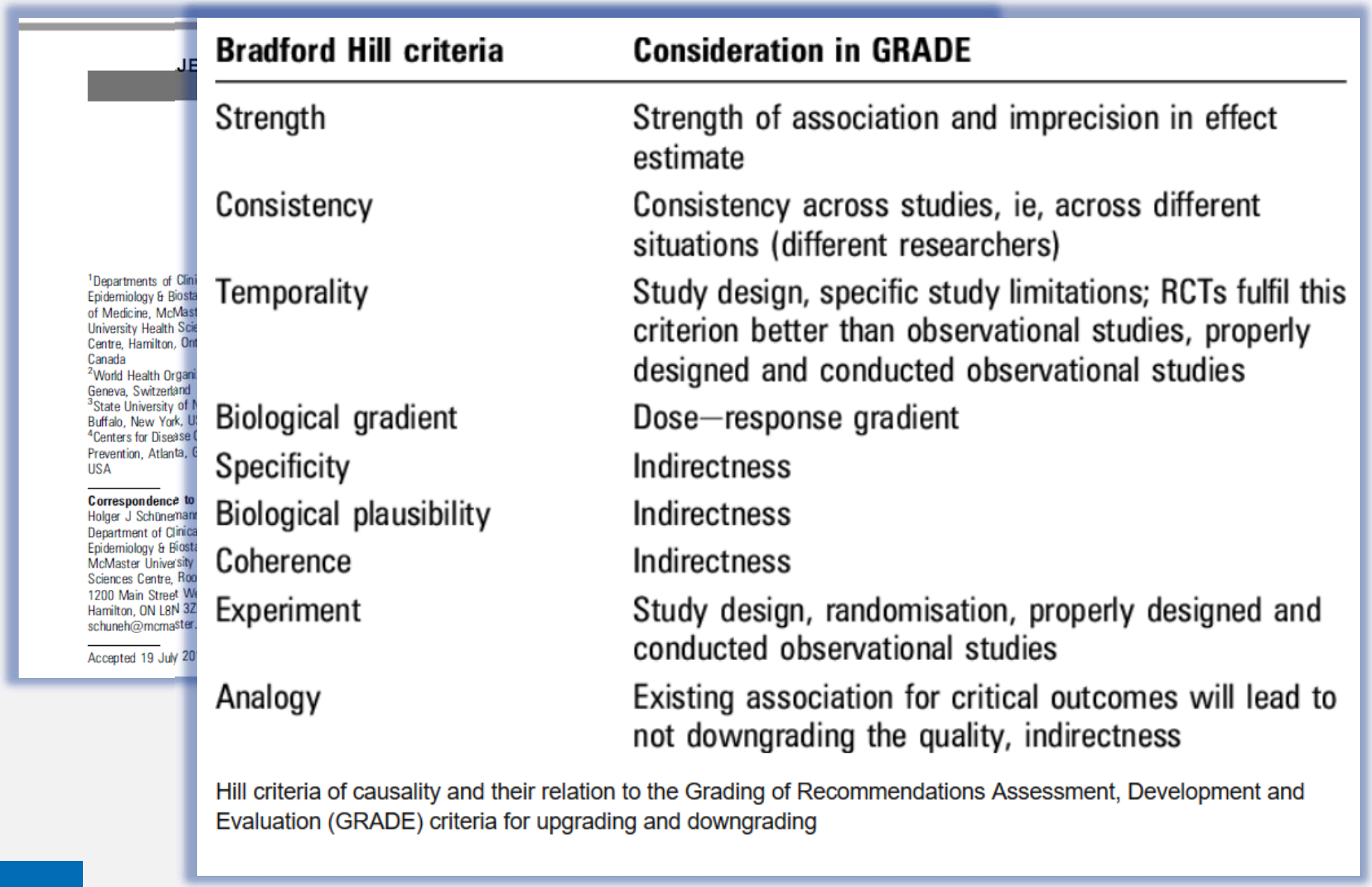
Accepted 19 July 2010

ABSTRACT
This article describes how the Grading of Recommendations Assessment, Development and Evaluation (GRADE) approach to grading the quality of evidence and strength of recommendations considers the Bradford Hill criteria for causation and how GRADE may relate to questions in public health. A primary concern in public health is that evidence from non-randomised studies may provide a more adequate or best available measure of a public health strategy's impact, but that such evidence might be graded as lower quality in the GRADE framework. GRADE, however, presents a framework that describes both criteria for assessing the quality of research evidence and the strength of recommendations that includes considerations arising from the Bradford Hill criteria. GRADE places emphasis on recommendations and in assessing quality of evidence; GRADE notes that randomisation is only one of many relevant factors. This article describes how causation may relate to developing recommendations and how the Bradford Hill criteria are considered in GRADE, using

First, concern has been expressed that herd immunity as a result of immunisation and indirect effects on the co-circulation of other pathogens are typically ascertained through the use of observational epidemiological methods. Although we do not disagree with this assessment, we would like to point out that, innovative randomised controlled trials (RCTs) using cluster-randomisation can be conducted to provide such information.[2] Second, concern is expressed that a quasi-RCT that found a 94% protective effect of a live, monovalent vaccine against measles was classified as 'moderate level of scientific evidence'. However, GRADE's strength of association criteria can be applied to quasi-RCTS and observational studies with no major threats to validity to upgrade the quality of evidence (see below). Such a judgement would be possible in this situation. Third, it is implied that GRADE ratings do not give credit to the 'gradient of effects with scale of population level impact compatible with degree of coverage'. However, we

Schünemann, H.S., et al. (2011). *The Grade Approach and Bradford Hill's Criteria for Causation. Vol. 65,2011. doi:10.1136/jech.2010.119933 .*

EPA United States Environmental Protection Agency



Bradford Hill criteria	Consideration in GRADE
Strength	Strength of association and imprecision in effect estimate
Consistency	Consistency across studies, ie, across different situations (different researchers)
Temporality	Study design, specific study limitations; RCTs fulfil this criterion better than observational studies, properly designed and conducted observational studies
Biological gradient	Dose—response gradient
Specificity	Indirectness
Biological plausibility	Indirectness
Coherence	Indirectness
Experiment	Study design, randomisation, properly designed and conducted observational studies
Analogy	Existing association for critical outcomes will lead to not downgrading the quality, indirectness

Hill criteria of causality and their relation to the Grading of Recommendations Assessment, Development and Evaluation (GRADE) criteria for upgrading and downgrading

Schünemann, H.S., et al. (2011). *The Grade Approach and Bradford Hill's Criteria for Causation. Vol. 65,2011. doi:10.1136/jech.2010.119933 .*

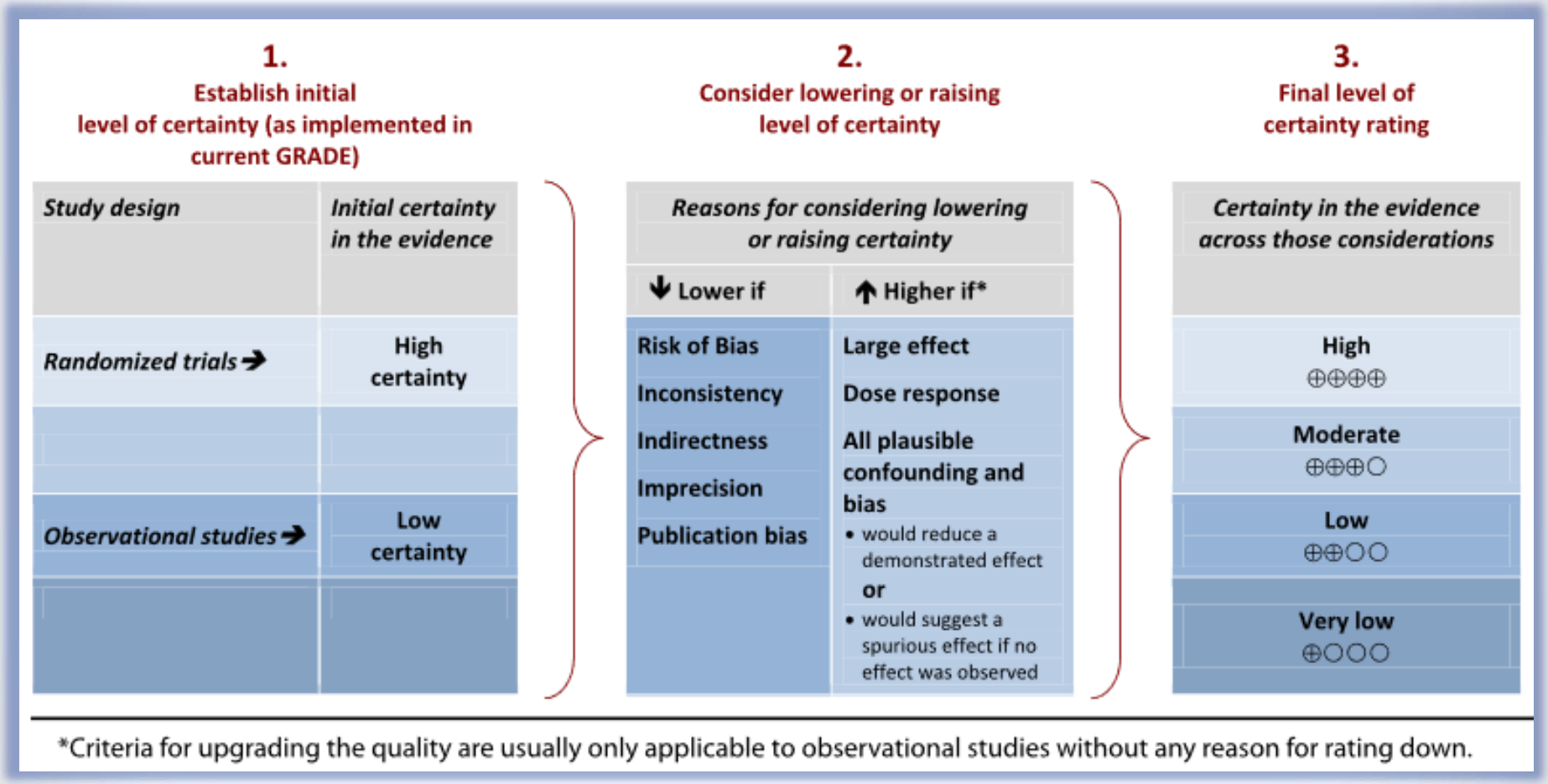
Interpreting the Certainty in the Evidence (CiE)

GRADE

Certainty rating	Definitions
⊕⊕⊕⊕ High	The panel is very confident that the true association lies close to that of the estimate of the association
⊕⊕⊕◯ Moderate	The panel is moderately confident in the association: The true association is likely to be close to the estimate of the association, but there is a possibility that it is substantially different
⊕⊕◯◯ Low	The panel's confidence in the association is limited: The true association may be substantially different from the estimate of the association
⊕◯◯◯ Very low	The panel has very little confidence in the association: The true association is likely to be substantially different from the estimate of association

HEI

Certainty in Evidence Framework



Schunemann, H. J., et al. (2018). "GRADE Guidelines: 18. How ROBINS-I and other tools to assess risk of bias in non-randomized studies should be used to rate the certainty of a body of evidence." J Clin Epidemiol.

EPA United States Environmental Protection Agency

Table 1
GRADE evidence profile: antibiotics for children with acute otitis media

Quality assessment						Summary of findings					
						Number of patients			Absolute risk		
No of studies (Design)	Limitations	Inconsistency	Indirectness	Imprecision	Publication bias	Placebo	Antibiotics	Relative risk (95% CI)	Control risk[a]	Risk difference (95% CI)	Quality
Pain at 24h											
5 (RCT)	No serious limitations	No serious inconsistency	No serious indirectness	No serious imprecision	Undetected	241/605	223/624	RR 0.9 (0.78−1.04)	367/1,000	Not Significant	⊕⊕⊕⊕ High
Pain at 2−7 d											
10 (RCT)	No serious limitations	No serious inconsistency	No serious indirectness	No serious imprecision	Undetected	303/1,366	228/1,425	RR 0.72 (0.62−0.83)	257/1,000	72 fewer per 1,000 (44−98)	⊕⊕⊕⊕ High
Hearing, inferred from the surrogate outcome abnormal tympanometry—1 mo											
4 (RCT)	No serious limitations	No serious inconsistency	Serious indirectness (because of indirectness of outcome)	No serious imprecision	Undetected	168/460	153/467	RR 0.89 (0.75−1.07)	350/1,000	Not Significant	⊕⊕⊕○ Moderate
Hearing, inferred from the surrogate outcome abnormal tympanometry—3 mo											
3 (RCT)	No serious limitations	No serious inconsistency	Serious indirectness (because of indirectness of outcome)	No serious imprecision	Undetected	96/398	96/410	RR 0.97 (0.76−1.24)	234/1,000	Not Significant	⊕⊕⊕○ Moderate
Vomiting, diarrhea, or rash											
5 (RCT)	No serious limitations	Serious inconsistency (because of inconsistency in absolute effects)	No serious indirectness	No serious imprecision	Undetected	83/711	110/690	RR 1.38 (1.09−1.76)	113/1,000	43 more per 1,000 (10−86)	⊕⊕⊕○ Moderate

Abbreviations: GRADE, Grading of Recommendations Assessment, Development, and Evaluation; RCT, randomized controlled trials; CI, confidence interval; RR, risk ratio.
[a] The control rate is based on the median control group risk across studies.

G. Guyatt et al. / Journal of Clinical Epidemiology 64 (2011) 383−394