HHS Public Access
Author manuscript
Environ Int. Author manuscript; available in PMC 2017 July 01.

Published in final edited form as:
Environ Int. 2016 ; 92-93: 611–616. doi:10.1016/j.envint.2016.01.004.

GRADE: Assessing the quality of evidence in environmental and occupational health

Rebecca L Morgan, MPH[a], **Kristina A Thayer, PhD**[b], **Lisa Bero, PhD**[c], **Nigel Bruce, MBChB, PhD**[d], **Yngve Falck-Ytter, MD**[e], **Davina Ghersi, MPH, PhD**[f], **Gordon Guyatt, MD, MSc**[a], **Carlijn Hooijmans, PhD**[g], **Miranda Langendam, PhD**[h], **Daniele Mandrioli, MD**[i], **Reem A. Mustafa, MD, MPH, PhD**[a,j], **Eva A Rehfuess, MA (Oxon), PhD**[k], **Andrew A Rooney, PhD**[b], **Beverley Shea, PhD**[l], **Ellen K Silbergeld, PhD**[m], **Patrice Sutton, MPH**[n], **Mary Wolfe, PhD**[b], **Tracey J Woodruff, MPH, PhD**[n], **Jos H Verbeek, MD, PhD**[o], **Alison C. Holloway, PhD**[p], **Nancy Santesso, RD, PhD**[a], and **Holger J Schünemann, MD, PhD, M.Sc.**[a,q]

Rebecca L Morgan: morganrl@mcmaster.ca; Kristina A Thayer: thayer@niehs.nih.gov; Lisa Bero: lisa.bero@sydney.edu.au; Nigel Bruce: ngb@liv.ac.uk; Yngve Falck-Ytter: Yngve.Falck-Ytter@case.edu; Davina Ghersi: davina.ghersi@nhmrc.gov.au; Gordon Guyatt: guyatt@mcmaster.ca; Carlijn Hooijmans: Carlijn.Hooijmans@radboudumc.nl; Miranda Langendam: m.w.langendam@amc.uva.nl; Daniele Mandrioli: mandriolid@ramazzini.it; Reem A. Mustafa: ramustafa@gmail.com; Eva A Rehfuess: rehfuess@ibe.med.uni-muenchen.de; Andrew A Rooney: andrew.rooney@nih.gov; Beverley Shea: bevshea@uottawa.ca; Ellen K Silbergeld: esilber2@jhu.edu; Patrice Sutton: patrice.sutton@ucsf.edu; Mary Wolfe: wolfe@niehs.nih.gov; Tracey J Woodruff: tracey.woodrfuff@ucsf.edu; Jos H Verbeek: Jos.Verbeek@ttl.fi; Alison C. Holloway: hollow@mcmaster.ca; Nancy Santesso: santesna@mcmaster.ca; Holger J Schünemann: schuneha@mcmaster.ca, holger.schunemann@mcmaster.ca

[a]Department of Clinical Epidemiology & Biostatistics, McMaster University, Health Sciences Centre, Room 2C14, 1280 Main Street West, Hamilton, ON L8S 4K1 Canada [b]Division of the National Toxicology Program, National Institute of Environmental Health Sciences, National Institutes of Health, Department of Health and Human Services, P.O. Box 12233, Mail Drop K2-02, Research Triangle Park, NC USA 27709 [c]Charles Perkins Centre, The University of Sydney, D17, The Hub, 6th floor, Charles Perkins Centre, The University of Sydney, New South Wales, 2006 [d]Department of Public Health and Policy, University of Liverpool, L69 3GB, United Kingdom [e]Division of Gastroenterology, Case Western Reserve University and Louis Stokes VA Medical Center, Cleveland, 10701 East Blvd., Cleveland, OH, 44106, USA [f]Sydney Medical School, University of Sydney, NSW 2006; National Health and Medical Research Council, 16 Marcus Clarke Street, Canberra City, ACT 2601, Australia [g]Departments of SYRCLE and Anesthesiology, Radboud University Medical Centre, Geert Grooteplein-Noord 29, route 231, 6525 GA Nijmegen, The Netherlands [h]Department of Clinical Epidemiology, Biostatistics and Bioinformatics, Academic Medical Center, University of Amsterdam, Room J1B-211, P.O. Box 22660, 1100 DD Amsterdam, The Netherlands [i]Cesare Maltoni Cancer Research Center, Ramazzini Institute, Via Saliceto 3, Bentivoglio (Bologna), P.O. Box 40133, Italy [j]Departments of Medicine/Nephrology and Biomedical & Health Informatics, University of Missouri-Kansas City,

Corresponding author: Rebecca L. Morgan. Department of Clinical Epidemiology & Biostatistics, Health Sciences Centre, Room 2C14, 1280 Main Street West, Hamilton, ON L8S 4K1 Canada. morganrl@mcmaster.ca.

Conflict of interest: The authors declare they have no financial interests with respect to this manuscript, or its content, or subject matter.

School of Medicine, M4-303, 2411 Holmes St., Kansas City, Missouri 64108-2792 [k]Institute for Medical Informatics, Biometry and Epidemiology, University of Munich, Marchioninistr. 15, 81377 Munich, Germany [l]Bruyere Research Institute and Ottawa Hospital Research Institute, University of Ottawa, Ottawa, ON [m]Department of Environmental Health Sciences, Johns Hopkins Bloomberg School of Public Health, 615 N. Wolfe Street, E6644, Baltimore, Maryland 21205, USA [n]Program on Reproductive Health and the Environment, University of California-San Francisco, 550 16th Street, San Francisco, California USA, 94143 [o]Finnish Institute of Occupational Health, Cochrane Work, PO Box 310, 70101 Kuopio, Finland [p]Department of Obstetrics and Gynecology, McMaster University, Health Sciences Centre, Room 3N52A, 1280 Main Street West, Hamilton, ON L8S 4K1 Canada [q]Department of Medicine, McMaster University, Health Sciences Centre, Room 2C14, 1280 Main Street West, Hamilton, ON L8S 4K1 Canada

Abstract

There is high demand in environmental health for adoption of a structured process that evaluates and integrates evidence while making decisions and recommendations transparent. The Grading of Recommendations Assessment, Development and Evaluation (GRADE) framework holds promise to address this demand. For over a decade, GRADE has been applied successfully to areas of clinical medicine, public health, and health policy, but experience with GRADE in environmental and occupational health is just beginning. Environmental and occupational health questions focus on understanding whether an exposure is a potential health hazard or risk, assessing the exposure to understand the extent and magnitude of risk, and exploring interventions to mitigate exposure or risk. Although GRADE offers many advantages, including its flexibility and methodological rigor, there are features of the different sources of evidence used in environmental and occupational health that will require further consideration to assess the need for method refinement. An issue that requires particular attention is the evaluation and integration of evidence from human, animal, *in vitro,* and *in silico* (computer modelling) studies when determining whether an environmental factor represents a potential health hazard or risk. Assessment of the hazard of exposures can produce analyses for use in the GRADE evidence-to-decision (EtD) framework to inform risk-management decisions about removing harmful exposures or mitigating risks. The EtD framework allows for grading the strength of the recommendations based on judgments of the certainty in the evidence (also known as quality of the evidence), as well as other factors that inform recommendations such as social values and preferences, resource implications, and benefits. GRADE represents an untapped opportunity for environmental and occupational health to make evidence-based recommendations in a systematic and transparent manner. The objectives of this article are to provide an overview of GRADE, discuss GRADE's applicability to environmental health, and identify priority areas for method assessment and development.

Keywords

GRADE; Evidence-based; Risk of Bias; Environmental Health; Risk Assessment; Recommendations

1 Introduction

There is high demand in environmental and occupational health for using systematic review methodology and structured frameworks to evaluate and integrate evidence to support evidence-based and transparent decisions and recommendations (Agency for Toxic Substances and Disease Registry (ATSDR) 2012; Bruce and others 2014; EFSA 2010; Johnson and others 2014; Koustas and others 2014; Lam and others 2014; Mandrioli and Silbergeld 2015; Mandrioli and others 2014; Murray and Thayer 2014; NRC 2007; NRC 2014a; NRC 2014b; Silbergeld and Scherer 2013; Whaley and others 2015; Woodruff and Sutton 2011; Woodruff and Sutton 2014). Environmental health, which includes occupational health, is a broad field in which data address all the physical, chemical, and biological factors external to a person, and all the related factors impacting behaviors (WHO 2015). Environmental health questions focus on understanding whether an exposure is a potential health hazard or risk using exposure assessments to recognize the extent and magnitude of exposure, and interventions to prevent or mitigate exposure or risk.

The Grading of Recommendations Assessment, Development, and Evaluation (GRADE) approach has the potential to improve transparency in addressing these questions in environmental health assessments. GRADE represents a rigorous, structured, and transparent process to inform decision-making beginning with well-defined questions, followed by an assessment of the certainty in the evidence (also called confidence in the effect or other estimates, or quality of the evidence) (Guyatt and others 2011d; Schünemann and others 2003), and leading to development of recommendations and decisions.

GRADE is widely used internationally to address topics related to clinical medicine, public health, and health policy (Atkins and others 2004; Guyatt and others 2011d; Guyatt and others 2008; Schünemann and others 2008), including by programs within the U.S. Centers for Disease Control and Prevention (CDC), World Health Organization (WHO), the U.S. Agency for Healthcare Research and Quality (AHRQ), and National Institute for Health and Clinical Excellence (NICE) in the United Kingdom and the National Health and Medical Research Council in Australia (Ahmed and others 2011; National Health and Medical Research Council 2011; Thornton and others 2013; Viswanathan and others 2012; WHO 2014b). The Cochrane Collaboration, which prepares, maintains, and promotes the accessibility of systematic reviews, uses the GRADE system for reporting on the quality of evidence for outcomes in systematic reviews (Higgins and others 2011; Schünemann and others 2011b). Formed in 2000, the GRADE Working Group now includes over 500 active members from 40 countries and serves as a think tank for advancing evidence-based decision-making in multiple disciplines (Schünemann and others 2003)(see also http://www.gradeworkinggroup.org/).

Advantages of using the GRADE approach have already been recognized by some within the environmental health field. The Navigation Guide proposed adapting GRADE for an environmental health context (Woodruff and Sutton 2011) and followed-up with a series of case studies to demonstrate the feasibility of applying GRADE to epidemiological and animal studies (Johnson and others 2014; Koustas and others 2014; Lam and others 2014; Vesterinen and others 2014). In 2013, the National Toxicology Program's (NTP) Office of

Health Assessment and Translation (OHAT) at the National Institute of Environmental Health Sciences announced plans to use GRADE in its evaluations to assess the evidence for associations between environmental exposures and non-cancer health effects (NTP 2013; NTP 2015; Rooney and others 2014). The SYstematic Review Center for Laboratory animal Experimentation (SYRCLE), is currently applying the GRADE approach to assess the quality of evidence from preclinical animal intervention studies (Hooijmans and others 2014). GRADE has also been used in recent systematic reviews of epidemiological studies of shift work and breast cancer risk (Ijaz and others 2013), shift work and cardiovascular disease (Vyas and others 2012), and adverse effects related to reduced indoor air quality related to household fuel use (Bruce and others 2013; WHO 2014a). GRADE, including its adoption by NTP/OHAT and the Navigation Guide, was specifically identified in the National Academy of Sciences' National Research Council (NRC) review of the U.S. Environmental Protection Agency's (EPA) Integrated Risk Information System as an approach that would increase the transparency of evaluating evidence (NRC 2014a). Use of GRADE in environmental health is likely to grow as systematic reviews become more common in the field and the limitations of expert-based narrative review methods are increasingly recognized (Aiassa and others 2015; EFSA 2010; EPA 2013; Mandrioli and Silbergeld 2015; NRC 2014b; Woodruff and Sutton 2014).

An additional advantage of GRADE is the GRADE Working Group's commitment to ongoing methods development and assessment of applicability to different areas of research. This is critical because experience with GRADE in the environmental health context is limited. Work to-date from the Navigation Guide, NTP, and WHO show the GRADE framework is sufficiently flexible to support use now (Johnson and others 2013; Johnson and others 2014; Koustas and others 2014; Lam and others 2014; NTP 2015; WHO 2014a); however, areas for further method assessment have been identified. In this respect, the GRADE Working Group serves as a vehicle to leverage transdisciplinary skills, knowledge, and resources to bridge the fields of clinical and environmental health. The objectives of this article are to provide an overview of the GRADE framework, discuss applicability of GRADE to environmental and occupational health, and identify priority areas for method development.

2 GRADE Approach

2.1 Formulating the Research Question

GRADE requires that decision-makers specify key-elements to formulate a relevant and focused question for decision-making (e.g., to inform clinical and public health guidelines, formulate scientific consensus statements, etc.) (Aiassa and others 2015; Guyatt and others 2011b). The key elements are the components of the question that identify what information must be provided in a primary study to evaluate the intervention under assessment and hence answer the question (Aiassa and others 2015). For instance, for questions aimed at evaluating interventions, the key elements are the Population, Intervention, Comparator, and Outcome (PICO) (Guyatt and others 2011b; Richardson and others 1995). Both beneficial and harmful outcomes that the target population may experience as a result of the intervention should be considered. At present, GRADE focuses on answering decision-making (i.e., actionable) questions about interventions (including diagnostic tests and strategies), though the GRADE framework has been expanded to prognostic questions (Iorio and others 2015; Spencer and others 2012).

2.2 Quality of the Evidence

GRADE uses a structured framework to determine overall certainty in the evidence (CiE) for outcomes across a collection of research studies or body of evidence (Figure 1)(Schünemann and others). The GRADE approach does not remove judgment from decision-making; however, the approach provides a framework of critical components to assess, guidance on the consideration of empirical evidence, and emphasizes transparency throughout the process. An initial evaluation of the CiE is conducted based on whether or not the research studies used randomized allocation. In the current GRADE approach, the CiE from randomized controlled trials (RCT) receives an initial rating of “high”, whereas the CiE from observational (i.e., non-randomized) studies starts at “low”. After this initial evaluation of randomization, other aspects of risk of bias (RoB), i.e., internal validity, are assessed. GRADE does not recommend the use of a specific RoB tool, but suggests specific criteria that should be considered when assessing a body of randomized or non-randomized studies that address risk of bias (Guyatt and others 2011e). In addition to RoB, the certainty in a body of evidence can be rated down for inconsistency, indirectness, imprecision, or publication bias, or rated up for the magnitude of the effect, dose-response gradient, or direction and impact of residual plausible confounding. Different terminology may be used to describe these elements as long as the concepts are identical (GRADE Working Group 2010; Schünemann and others). Like RCTs, randomized experimental studies in animals would start as “high” and typically be downgraded for indirectness due to differences in the population (Guyatt and others 2011c). The evidence is assessed and presented in an evidence summary table separately for each critical or important outcome and expressed using four levels of certainty ratings (i.e., “high”, “moderate”, “low”, or “very low”) (Balshem and others 2011; Guyatt and others 2011a). This table, called a GRADE Evidence Profile or Summary of Findings table, requires transparent descriptions of the reasons for rating down and rating up (WHO 2014a).

2.3 Recommendations and the Evidence-to-Decision Framework

In addition to assessing the CiE across outcomes, the GRADE EtD framework explicitly considers the balance of benefits and harms, values and preferences, resource implications, feasibility, equity, and acceptability to determine the strength of the recommendation (strong or weak), and the direction (for or against) to make a final recommendation or decision (Andrews and others 2013; Schünemann and others 2012; Treweek and others 2013). The elements of the framework’s structure transparently display the important criteria for deliberation (including relevant research evidence, judgments from decision makers, and other considerations) to inform the balance about the desirable and undesirable consequences of the options or interventions considered. A judgment is needed for making decisions during all steps. However, the GRADE EtD framework provides a structure to maximize transparency and limit subjectivity throughout the process: in fact CiE is a key determinant for making a strong GRADE guidelines recommendation (Djulbegovic and others 2015).

3 Considerations for Environmental Health

3.1 Formulating the Research Question

The GRADE approach has been utilized predominantly to answer questions on interventions in health care, like "what is the impact of an intervention (including diagnostic tests and strategies) compared with an alternative on patient or population important outcomes?" or "should intervention A or B be used for X?" In the context of decision-making in environmental health, the term intervention has somewhat different connotations. First, an intervention can be thought of as a specific environmental factor (i.e., exposure) that is being evaluated in human, animal, *in vitro*, or *in silico* studies as a risk factor or causative agent for an undesirable health outcome. In this scenario, the PICO question can be rephrased as a PECO question, where the term "Intervention" is replaced with "Exposure" (Evidence. 2013; NTP 2015; Woodruff and Sutton 2014). The complexity of the exposure questions will vary, ranging from a single well-defined chemical to complex scenarios like wind farms, agricultural run-off, etc. To address the benefits and harms to humans from wind farms, PECO questions were developed to look at the exposure of physical emissions produced by wind farms or wind turbines (e.g., noise, infrasound, shadow flicker, and electromagnetic radiation), as compared with no exposure to the physical emissions produced by wind farms or turbines (Merlin and others 2015). Questions assessing exposures as risk factors or causative agents are used in risk assessments, which have several sub-questions (EPA 2012; Schünemann and others 2011a):

- Hazard identification: What health problems are caused by the environmental factor?
- Dose-response assessment: What are the health problems at different exposure levels?
- Exposure assessment: What is the extent and nature of the exposure in the target population?
- Risk characterization: What is the extra risk of health problems in the exposed population?

Second, an environmental intervention question could be formulated to evaluate the impact of interventions that prevent or mitigate an exposure or risk. Environmental exposure-related interventions typically address chemical or physical agents in the environment, such as air, soil, water, or food, in a public or occupational setting, with the goal of trying to prevent, remove, or reduce exposure levels (e.g., reduction at source, improved ventilation, ingredient reformulation) through regulatory, technical, or behavioral interventions. Questions assessing the effects of an intervention to prevent or reduce exposure should be based on an established relationship between the exposure and health outcome(s). For example, since the relationship between noise exposure and noise-induced hearing loss has been established, showing that an intervention reduces noise exposure is sufficient to also to conclude that the intervention decreases noise-induced hearing loss (Verbeek and others 2012). In studies of environmental health, such questions have the ability to compare the desirable consequences of reducing an exposure with potentially undesirable consequences of removing an exposure (e.g., costs, use of alternatives with unknown toxicity). While these types of questions are very similar to the clinical or public health intervention PICO questions GRADE was designed to assess, some challenges have been identified, such as how to assess complex interventions, use non-epidemiological evidence, and choosing outcomes and outcome measures (Rehfuess and Akl 2013). Methodological research has continued to address concerns with applying GRADE to studies of interventions (Guyatt and others 2011b; Schünemann 2013).

3.2 Quality of the Evidence

3.2.1 Human and Experimental Animal Data—In environmental health, observational human studies and experimental animal studies (where animals are randomly assigned to treatment groups), and observational animal studies (i.e., "wildlife studies" or natural population-based studies) are often the highest quality evidence available to understand *whether* there is an association (or, if possible, cause-effect relationship) between an exposure and health outcome, as in the case of carcinogens (Pearce and others 2015). The factors considered in GRADE when making and presenting judgments about the CiE (Figure 1) translate well to observational human and experimental animal studies, although harmonization of RoB tools and development of additional guidance on when rating down or rating up should be pursued. The WHO considered evidence from both non-randomized experimental and observational studies to inform their Recommendations for Indoor Air Quality (WHO 2014a). In the report, WHO assessed whether or not coal should be used as a household fuel. The decision to recommend against using unprocessed coal as a household fuel was informed by 1) the results from studies of cancer in humans and experimental animals; 2) systematic reviews of observational studies on particulate matter exposure and risk of lung cancer; and 3) population-level studies on the toxicity of coal and the impact of banning coal. While possible confounders of the different study types were recognized, they still provided the best available evidence to inform the recommendations. In addition, on-going methods development for rating the risk of bias (Bilotta and others 2014; Johnson and others 2014; Koustas and others 2014; Lam and others 2014; Morgan and others 2015; NTP 2015; WHO 2014a) includes searching for observational studies that might be considered equivalent to randomized trials for the initial assessment of the risk of bias (e.g., factors in study design and execution that mitigate the lack of randomization, such as steps taken to fully control or adjust for confounding). Examples, however are currently lacking.

3.2.2 Mechanistic Data—In environmental health, human and experimental animal data are often interpreted in conjunction with evidence from mechanistic data supporting the biological plausibility of an association and/or to prioritize chemicals for additional testing or evaluation. The GRADE framework does not explicitly address mechanistic data, but they may be used to inform judgments about indirectness. There are an estimated 85,000 chemicals in commerce, the vast majority of which have not been tested for toxicity, even though in many cases the evidence available for a chemical will be mechanistic in nature (EPA 2009; Judson and others 2009). The lack of toxicity data for most environmental chemicals has led to major initiatives to generate high throughput screening (HTS) data for chemicals. For example, the NTP's Tox21 HTS program has generated data for ~10,000 chemicals on ~75 biochemical- and cell-based assays that cover a range of activities including overall cellular health (cytotoxicity and apoptosis induction, mitochondrial toxicity, DNA damage), perturbation of cell signaling pathways, inflammatory response induction, agonists/antagonists for 15 nuclear receptors, and drug metabolism (Tice and others 2013). The US EPA's ToxCast HTS program currently has mechanistic data on 1860 chemicals tested in up to 821 assay endpoints (Kavlock and others 2012); however, many chemicals are still untested. Computer-modeling approaches are also being pursued to predict potential hazard and likelihood of significant exposure. For mechanistic data, tools to rate RoB for *in vitro* and *in silico* studies need to be developed and their contribution to the stream of evidence for different outcomes should be determined because these data are expected to be used more widely for prioritizing chemicals of concern as well as replacing traditional data in regulatory assessments (Mandrioli and Silbergeld 2015; NRC 2007). When assessing the effects of wind farms on human health, both direct and indirect evidence was considered to address the PECO question (Merlin and others 2015). When assessing the body of evidence across the outcome of shadow flicker, there was low quality direct evidence available; however, available indirect data suggested that shadow flicker can affect health by inducing seizures among persons prone to photosensitive epilepsy. The utility of the GRADE rating down and rating up factors also needs to be assessed, although the concepts should generally apply (e.g., magnitude of effect can be analogous to efficacy and potency in an *in vitro* system). Analyses to assess the predictive utility of mechanistic data are a high priority in toxicology, and results will inform indirectness ratings within the GRADE framework.

3.3 Evidence-to-Decision Frameworks

Very little work has been done to use structured and transparent decision-making frameworks to guide the development of recommendations in environmental health. The WHO Recommendations for Indoor Air Quality applied the GRADE EtD framework to guide their final recommendations (WHO 2014a). For their recommendation on household use of coal, in addition to the quality of evidence from studies on carcinogenicity of coal, risk of lung cancer, and population-level studies on toxicity, they also determined that the benefits of replacing unprocessed coal with cleaner alternatives clearly outweigh the harms of replacement, the values and preferences of replacing coal varied among stakeholders, and that there may be some limitations to the feasibility of implementing cleaner alternatives based on affordability and supply. The GRADE EtD framework, which has the capacity to integrate consideration of the CiE of a health hazard with evidence of benefit associated with mitigating exposure, values, preferences, resource implications and other criteria, has great potential for enhancing the transparency of decision-making in environmental and occupational health. The strength of the recommendation may be apparent and actionable, or application of GRADE may reveal gaps in our knowledge, and thus help efficiently and effectively target the allocation of scarce research funds.

The regulation of diesel is an example of an environmental topic that could be addressed with the GRADE EtD framework. Diesel engine exhaust is carcinogenic to humans and associated with increased hospital admissions, emergency room visits, asthma attacks, and premature death (IARC 2012; Office of Environmental Health Hazard Assessment 2007). At the same time, diesel engines have desirable consequences of higher fuel efficiency, lower carbon dioxide emissions, heavy duty hauling capacity, and durability. For example, EPA rule-making for diesel standards included consideration of the composition of diesel, technological feasibility, costs of retrofitting or replacing, cost-benefit analyses that include quantifying human health impacts, overall economic impact and alternatives assessment. Moreover, the rule-making applied to specific scenarios such as vehicles on highways, city streets, construction sites, and ports. These analyses have led to a number of emission standards for diesel fuel and diesel engines (NCDC 2014). By 2030, EPA estimates that particulate matter and nitrous oxides will be reduced by 380,000 tons/year and 7 million tons/year, respectively. This will result in annual benefits of over $290 billion, at a cost of approximately $15 billion. The GRADE EtD framework could also be applied to alternative assessments that look for safer chemicals by identifying and evaluating the safety of alternative chemicals (EPA 2011). Although such assessments are often not regulatory, they are used to inform consumer choice and encourage industry to move to safer alternatives and can complement regulatory actions.

The challenges of applying the GRADE EtD framework to environmental health topics are expected to be similar to clinical research, with most findings requiring a careful weighing of the health and other benefits or harms. A challenge specific to decision-making for environmental health is that many regulatory agencies require a determination of an allowable level or threshold of an exposure or risk, while in other cases there is no allowable exposure (for example asbestos ban). In studies where there is not a clear desirable effect of the exposure, the balance may focus on how frequently the undesirable effects occur. Research is also needed to increase understanding and acceptability of the format that desirable and undesirable consequences are presented in to end-users.

4 Future Directions

This paper provides an overview of important aspects of adapting GRADE to decision-making in environmental health. In 2014, several project groups were formed within the GRADE Working Group to focus on methods assessment needs that are directly applicable to environmental and occupational health, including project groups for environmental health, observational studies, public health, application of GRADE to laboratory animal research, and non-randomized study risk of bias integration. Priority areas for the environmental and occupational health project group include (1) developing approaches to evaluate and integrate evidence from observational human, animal, *in vitro*, and *in silico* (computer modeling) studies to determine whether an association exist between exposure and health outcome(s); (2) applying GRADE to evaluations of interventions to mitigate exposure or reduce risk when an association has been identified; and (3) gaining experience in applying the GRADE frameworks for evidence-to-decision (EtD) and determining the direction and strength of recommendations for environmental and occupational health topics. Critically adapting GRADE to environmental health also requires consideration of how to rate the overall strength of the evidence and to integrate evidence across multiple evidence streams.

5 Conclusions

This paper examines several key components of GRADE as they can be assessed and expanded as a standardized methodology for research and decision-making in environmental and occupational health. Over 90 organizations from 18 countries worldwide have adopted the GRADE framework to assess evidence and inform decision-making. With a focus on rigorous and transparent methods, the GRADE approach has been applied successfully to clinical medicine, public health, diagnostic decision-making, questions about prognosis, and has great potential for the field of environmental and occupational health. In parallel to the methods development that has occurred over the past decades in the clinical and public health field, environmental health scientists have developed topic specific expertise about the evidence that informs how the environment shapes our health and sets the stage for knowledge transfer across disciplines to strengthen the scientific basis of decision-making for public policy. Leveraging this synergy will increase the transparency of, and scientific basis for, decision-making in environmental health, and thus help secure improved health outcomes for individuals and populations.

Acknowledgments

This research was supported by the intramural research program of the National Institute of Environmental Health Sciences and the MacGRADE center at McMaster University. The contribution of UCSF Program on Reproductive Health and the Environment co-authors (TW and PS) to this research was supported by the Clarence Heller Foundation, the National Institute of Environmental Health Sciences (grants ES018135 and ESO22841), and U.S. EPA STAR grants (RD83467801 and RD83543301). Authors would like to acknowledge the contributions of Elisa Aiassa and Annette Martine Pruss-Ustun as members of the GRADE Environmental Health Project Group.

Abbreviations

AHRQ	Agency for Healthcare Research and Quality
ASTDR	Agency for Toxic Substances and Disease Registry
CDC	Centers for Disease Control and Prevention
CiE	Certainty in the Evidence
EFSA	European Food Safety Authority
EPA	Environmental Protection Agency
EtD	Evidence-to-decision
GRADE	Grading of Recommendations Assessment, Development, and Evaluation working group
OHAT	Office of Health Assessment and Translation
PECO	Population, Exposure, Comparator, Outcome
PICO	Population, Intervention, Comparator, Outcome
NRC	National Research Council
NTP	National Toxicology Program
RoB	Risk of Bias

SYRCLE	SYstematic Review Center for Laboratory animal Experimentation
WHO	World Health Organization

References

Ahmed F, Temte JL, Campos-Outcalt D, Schünemann HJ. AEBRW Group. Methods for developing evidence-based recommendations by the Advisory Committee on Immunization Practices (ACIP) of the US Centers for Disease Control and Prevention (CDC). Vaccine. 2011; 29:9171–9176. [PubMed: 21839794]

Aiassa E, Higgins J, Frampton G, Greiner M, Afonso A, Amzal B, Deeks J, Dorne J-L, Glanville J, Lövei G. Applicability and feasibility of systematic review for performing evidence-based risk assessment in food and feed safety. Critical reviews in food science and nutrition. 2015; 55:1026–1034. [PubMed: 25191830]

Andrews J, Guyatt G, Oxman AD, Alderson P, Dahm P, Falck-Ytter Y, Nasser M, Meerpohl J, Post PN. Kunz RGRADE guidelines: 14, Going from evidence to recommendations: the significance and presentation of recommendations. Journal of clinical epidemiology. 2013; 66:719–725. [PubMed: 23312392]

Atkins D, Eccles M, Flottorp S, Guyatt GH, Henry D, Hill S, Liberati A, O'Connell D, Oxman AD, Phillips B. Systems for grading the quality of evidence and the strength of recommendations I: critical appraisal of existing approaches The GRADE Working Group. BMC health services research. 2004; 4:38. [PubMed: 15615589]

ATSDR. The Future of Science at ATSDR: A Symposium. Atlanta, GA: US Department of Health and Human Services (DHHS) Agency for Toxic Substances and Disease Registry (ATSDR); 2012.

Balshem H, Helfand M, Schunemann HJ, Oxman AD, Kunz R, Brozek J, Vist GE, Falck-Ytter Y, Meerpohl J, Norris S. Guyatt GHGRADE guidelines: 3, Rating the quality of evidence. Journal of clinical epidemiology. 2011; 64:401–406. [PubMed: 21208779]

Bilotta GS, Milner AM, Boyd IL. Quality assessment tools for evidence from environmental science. Environmental Evidence. 2014; 3:1–14.

Bruce N, Dora C, Krzyzanowski M, Adair-Rohani H, Morawska L, Wangchuk T. Tackling the health burden from household air pollution: Development and implementation of new WHO Guidelines. 2013

Bruce N, Pope D, Rehfuess E, Balakrishnan K, Adair-Rohani H, Dora C. WHO indoor air quality guidelines on household fuel combustion: Strategy implications of new evidence on interventions and exposure–risk functions. Atmospheric Environment. 2014 Available online 27 August 2014.

Djulbegovic B, Kumar A, Kaufman RM, Tobian A, Guyatt GH. Quality of evidence is a key determinant for making a strong GRADE guidelines recommendation. Journal of clinical epidemiology. 2015

EFSA. Application of systematic review methodology to food and feed safety assessments to support decision making. EFSA Journal. 2010; 8:1637.

EPA. EPA Announces Actions to Address Chemicals of Concern, Including Phthalates. 2009.

EPA. Design for the Environment Alternatives Assessments. 2011.

EPA. Hazard Identification. 2012.

EPA. Applying systematic review to assessments of health effects of chemical exposures; EPA Workshop; Washington, DC. 2013;

Evidence., C.f.E. Environmental Evidence. Collaboration for Environmental Evidence; 2013. Guidelines for Systematic Review and Evidence Synthesis in Environmental Management. www.environmentalevidence.org/Documents/Guidelines/Guidelines4.2.pdf

GRADE Working Group. [accessed May 13, 2015] Criteria for applying or using GRADE. 2010. http://www.gradeworkinggroup.org/intro.htm#criteria

Guyatt GH, Oxman AD, Akl EA, Kunz R, Vist G, Brozek J, Norris S, Falck-Ytter Y, Glasziou P, Debeer H, Jaeschke R, Rind D, Meerpohl J, Dahm P. Schunemann HJGRADE guidelines: 1.

Introduction-GRADE evidence profiles and summary of findings tables. Journal of Clinical Epidemiology. 2011a; 64:383–394. [PubMed: 21195583]

Guyatt GH, Oxman AD, Kunz R, Atkins D, Brozek J, Vist G, Alderson P, Glasziou P, Falck-Ytter Y. Schunemann HJGRADE guidelines: 2. Framing the question and deciding on important outcomes. J Clin Epidemiol. 2011b; 64:395–400. [PubMed: 21194891]

Guyatt GH, Oxman AD, Kunz R, Woodcock J, Brozek J, Helfand M, Alonso-Coello P, Falck-Ytter Y, Jaeschke R, Vist G, Akl EA, Post PN, Norris S, Meerpohl J, Shukla VK, Nasser M. Schunemann HJGRADE guidelines: 8. Rating the quality of evidence--indirectness. Journal of clinical epidemiology. 2011c; 64:1303–1310. [PubMed: 21802903]

Guyatt GH, Oxman AD, Schunemann HJ, Tugwell P, Knottnerus A. GRADE guidelines: A new series of articles in the Journal of Clinical Epidemiology. Journal of clinical epidemiology. 2011d; 64:380–382. [PubMed: 21185693]

Guyatt GH, Oxman AD, Vist G, Kunz R, Brozek J, Alonso-Coello P, Montori V, Akl EA, Djulbegovic B. Falck-Ytter YGRADE guidelines: 4. Rating the quality of evidence—study limitations (risk of bias). Journal of clinical epidemiology. 2011e; 64:407–415. [PubMed: 21247734]

Guyatt GH, Oxman AD, Vist GE, Kunz R, Falck-Ytter Y, Alonso-Coello P, Schünemann HJ. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. Bmj. 2008; 336:924–926. [PubMed: 18436948]

Higgins JP, Altman DG, Gøtzsche PC, Jüni P, Moher D, Oxman AD, Savovid J, Schulz KF, Weeks L, Sterne JA. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. Bmj. 2011; 343:d5928. [PubMed: 22008217]

Hooijmans CR, Rovers MM, de Vries RB, Leenaars M, Ritskes-Hoitinga M, Langendam MW. SYRCLE's risk of bias tool for animal studies. BMC medical research methodology. 2014; 14:43. [PubMed: 24667063]

IARC. IARC: Diesel engine exhaust carcinogenic. 2012. Press release

Ijaz S, Verbeek J, Seidler A, Lindbohm M-L, Ojajarvi A, Orsini N, Costa G, Neuvonen K. Night-shift work and breast cancer—a systematic review and meta-analysis. Scand J Work Environ Health. 2013; 39:431–447. [PubMed: 23804277]

Iorio A, Spencer FA, Falavigna M, Alba C, Lang E, Burnand B, McGinn T, Hayden J, Williams K, Shea B. Use of GRADE for assessment of evidence about prognosis: rating confidence in estimates of event rates in broad categories of patients. bmj. 2015; 350:h870. [PubMed: 25775931]

Johnson, PI.; Sutton, P.; Atchley, D.; Koustas, E.; Lam, J.; Robinson, K.; Sen, S.; Axelrad, D.; Woodruff, TJ. [accessed 29 November, 2014] Applying the Navigation Guide: Case Study #1: The impact of developmental exposure to perfluorooctanoic acid (PFOA) on fetal growth (Final protocol). 2013. http://prhe.ucsf.edu/prhe/navigationguide.html

Johnson PI, Sutton P, Atchley DS, Koustas E, Lam J, Sen S, Robinson KA, Axelrad DA, Woodruff TJ. The Navigation Guide—evidence-based medicine meets environmental health: systematic review of human evidence for PFOA effects on fetal growth. Environ Health Perspect. 2014; 122:1028–1039. [PubMed: 24968388]

Judson R, Richard A, Dix DJ, Houck K, Martin M, Kavlock R, Dellarco V, Henry T, Holderman T, Sayre P. The toxicity data landscape for environmental chemicals. Environ Health Perspect. 2009; 117:685–695. [PubMed: 19479008]

Kavlock R, Chandler K, Houck K, Hunter S, Judson R, Kleinstreuer N, Knudsen T, Martin M, Padilla S, Reif D. Update on EPA's ToxCast program: providing high throughput decision support tools for chemical risk management. Chemical research in toxicology. 2012; 25:1287–1302. [PubMed: 22519603]

Koustas E, Lam J, Sutton P, Johnson PI, Atchley DS, Sen S, Robinson KA, Axelrad DA, Woodruff TJ. The Navigation Guide—evidence-based medicine meets environmental health: systematic review of nonhuman evidence for PFOA effects on fetal growth. Environ Health Perspect. 2014; 122:1015–1027. [PubMed: 24968374]

Lam J, Koustas E, Sutton P, Johnson PI, Atchley DS, Sen S, Robinson KA, Axelrad DA, Woodruff TJ. The Navigation Guide—evidence-based medicine meets environmental health: integration of animal and human evidence for PFOA effects on fetal growth. Environ Health Perspect. 2014; 122:1040–1051. [PubMed: 24968389]

Mandrioli, D.; Silbergeld, EK. Environmental health perspectives. 2015. Evidence from Toxicology: The Most Essential Science for Prevention.

Mandrioli, D.; Sillbergeld, E.; Bero, L. Preperation of Evidence Based Toxicology Handbook; Cochrane Colloquium expert meeting; Hyderabad, India. September 26, 2014; 2014. https://colloquium.cochrane.org/meetings/evidence-based-toxicology-handbook

Merlin T, Newton S, Ellery B, Milverton J, Farah C. Systematic review of the human health effects of wind farms. 2015

Morgan, RL.; Thayer, KA.; Guyatt, G.; Blain, R.; Eftim, S.; Ross, P.; Santesso, N.; Holloway, AC.; Schunemann, HJ. Cochrane Colloquium. Vienna, Austria: 2015. Assessing the Usability of ACROBAT-NRSI for Studies of Exposure and Intervention in Environmental Health Research.

Murray HE, Thayer KA. Implementing systematic review in toxicological profiles: ATSDR and NIEHS/NTP collaboration. Journal of environmental health. 2014; 76:34–35. [PubMed: 24749224]

National Health and Medical Research Council. Procedures and requirements for meeting the 2011 NHMRC standard for clinical practice guidelines. 2011.

NCDC. Tools & Resources Regulatory Standards. 2014. http://www.epa.gov/cleandiesel/reg-prog.htm

NRC. Toxicity testing in the 21st century: A vision and a strategy. National Academies Press; 2007.

NRC. [accessed 1 January 2015] Review of EPA's Integrated Risk Information System (IRIS) Process. 2014a. (http://www.nap.edu/catalog.php?record_id=18764)

NRC. [accessed 1 January 2015] Review of the Environmental Protection Agency's State-of-the-Science Evaluation of Nonmonotonic Dose-Response Relationships as they Apply to Endocrine Disrupters. 2014b. (http://www.nap.edu/catalog.php?record_id=18608)

NTP. [accessed 9 August 2014] Board of Scientific Counselors June 25, 2013 meeting. Meeting materials. 2013. available at http://ntp.niehs.nih.gov/go/40246

NTP. Handbook for Conducting a Literature-Based Health Assessment Using Office of Health Assessment and Translation (OHAT) Approach for Systematic Review and Evidence Integration. 2015. January 9, 2015 release. Available at http://ntp.niehs.nih.gov/go/38673

Office of Environmental Health Hazard Assessment. HEALTH EFFECTS OF DIESEL EXHAUST: A fact sheet. Cal/EPA's Office of Environmental Health Hazard Assessment and the American Lung Association; 2007. http://oehha.ca.gov/public_info/facts/dieselfacts.html

Pearce NE, Zahm SH, Andersen A, Antó i Boqué JM, Cardis E, Grimsrud TK, Kjaerheim K, Kogevinas M, Porta Serra M. IARC monographs: 40 years of evaluating carcinogenic hazards to humans. 2015

Rehfuess EA, Akl EA. Current experience with applying the GRADE approach to public health interventions: an empirical study. BMC public health. 2013; 13:9. [PubMed: 23294803]

Richardson WS, Wilson MC, Nishikawa J, Hayward RS. The well-built clinical question: a key to evidence-based decisions. Acp j club. 1995; 123:A12–13. [PubMed: 7582737]

Rooney AA, Boyles AL, Wolfe MS, Bucher JR, Thayer KA. Systematic review and evidence integration for literature-based environmental health science assessments. Environ Health Perspect. 2014

Schünemann H, Brożek J, Oxman G. Handbook for grading the quality of evidence and the strength of recommendations using the GRADE approach. 2013

Schünemann H, Hill S, Guyatt G, Akl EA, Ahmed F. The GRADE approach and Bradford Hill's criteria for causation. Journal of epidemiology and community health. 2011a; 65:392–395. [PubMed: 20947872]

Schünemann H, Oxman A, Higgins J, Vist G, Glasziou P, Guyatt G. Cochrane handbook for systematic reviews of interventions version 5.1. 0 (updated March 2011). 2011b

Schünemann HJ. Methodological idiosyncracies, frameworks and challenges of non-pharmaceutical and non-technical treatment interventions. Zeitschrift für Evidenz, Fortbildung und Qualität im Gesundheitswesen. 2013; 107:214–220.

Schünemann HJ, Best D, Vist G, Oxman AD. Letters, numbers, symbols and words: how to communicate grades of evidence and recommendations. Canadian Medical Association Journal. 2003; 169:677–680. [PubMed: 14517128]

Schünemann HJ, Oxman AD, Brozek J, Glasziou P, Jaeschke R, Vist GE, Williams JW Jr, Kunz R, Craig J, Montori VM. Grading quality of evidence and strength of recommendations for diagnostic tests and strategies. Bmj. 2008; 336:1106–1110. [PubMed: 18483053]

Schünemann, HJ.; Oxman, AD.; Vist, GE.; Higgins, JPT.; Deeks, JJ.; PG; Guyatt, GH. on behalf of the Cochrane Applicability and Recommendations Methods Group. Chapter 12: Interpreting results and drawing conclusions. In: Higgins, JPT.; Green, S., editors. Cochrane Handbook for Systematic Reviews of Interventions Version 510. The Cochrane Collaboration, 2011; 2012. updated March 2011Available at www.cochrane-handbook.org [accessed 13 July 2012]

Silbergeld E, Scherer RW. Evidence-based toxicology: Strait is the gate, but the road is worth taking. Altex. 2013; 30:67–73. [PubMed: 23338807]

Spencer FA, Iorio A, You J, Murad MH, Schünemann HJ, Vandvik PO, Crowther MA, Pottie K, Lang ES, Meerpohl JJ. Uncertainties in baseline risk estimates and confidence in treatment effects. Bmj. 2012; 345:e7401. [PubMed: 23152569]

Thornton J, Alderson P, Tan T, Turner C, Latchem S, Shaw E, Ruiz F, Reken S, Mugglestone MA, Hill J. Introducing GRADE across the NICE clinical guideline program. Journal of clinical epidemiology. 2013; 66:124–131. [PubMed: 22406196]

Tice RR, Austin CP, Kavlock RJ, Bucher JR. Improving the human hazard characterization of chemicals: a Tox21 update. Environ Health Perspect. 2013; 121:756–765. [PubMed: 23603828]

Treweek S, Oxman AD, Alderson P, Bossuyt PM, Brandt L, Brozek J, Davoli M, Flottorp S, Harbour R, Hill S, Liberati A, Liira H, Schunemann HJ, Rosenbaum S, Thornton J, Vandvik PO, Alonso-Coello P. Consortium D. Developing and Evaluating Communication Strategies to Support Informed Decisions and Practice Based on Evidence (DECIDE): protocol and preliminary results. Implement Sci. 2013; 8:6. [PubMed: 23302501]

Verbeek JH, Kateman E, Morata TC, Dreschler WA, Mischke C. Interventions to prevent occupational noise-induced hearing loss. The Cochrane Library. 2012

Vesterinen HM, Johnson PI, Atchley DS, Sutton P, Lam J, Zlatnik MG, Sen S, Woodruff TJ. Fetal growth and maternal glomerular filtration rate: a systematic review. The Journal of Maternal-Fetal & Neonatal Medicine. 2014:1–6.

Viswanathan M, Ansari MT, Berkman ND, Chang S, Hartling L, McPheeters M, Santaguida PL, Shamliyan T, Singh K, Tsertsvadze A. Assessing the risk of bias of individual studies in systematic reviews of health care interventions. 2012

Vyas MV, Garg AX, Iansavichus AV, Costella J, Donner A, Laugsand LE, Janszky I, Mrkobrada M, Parraga G, Hackam DG. Shift work and vascular events: systematic review and meta-analysis. Bmj. 2012; 345:e4800. [PubMed: 22835925]

Whaley P, Halsall C, Agerstrand M, Benford D, Aiassa E, Bilotta GS, Coggon D, Dempsey C, Duarte-Davidson R, Lipworth S, Mackenzie Ross S, Martin O, Meads C, Meyer-Baron M, Miller JW, Pease C, Rooney A, Sapiets A, Stewart G, Taylor D. Implementing systematic review techniques in chemical risk assessment: Challenges, opportunities and recommendations. Environment International. 2015

WHO. Indoor air quality guidelines: household fuel combustion. 2014a.

WHO. WHO Handbook for guideline development. World Health Organization; 2014b.

WHO. Environmental Health. 2015. http://www.who.int/topics/environmental_health/en/

Woodruff TJ, Sutton P. An evidence-based medicine methodology to bridge the gap between clinical and environmental health sciences. Health Affairs. 2011; 30:931–937. [PubMed: 21555477]

Woodruff TJ, Sutton P. The Navigation Guide Systematic Review Methodology: A Rigorous and Transparent Method for Translating Environmental Health Science into Better Health Outcomes. Environmental health perspectives. 2014

Highlights

- A structured framework is needed for decision-making in environmental health.
- GRADE has been applied in many disciplines and holds great promise for the field.
- Methods development and assessment is needed to address environmental health data.
- Methods assessment priorities are evaluation and integration of diverse evidence streams.
- GRADE evidence-to-decision framework informs risk and other management decisions.

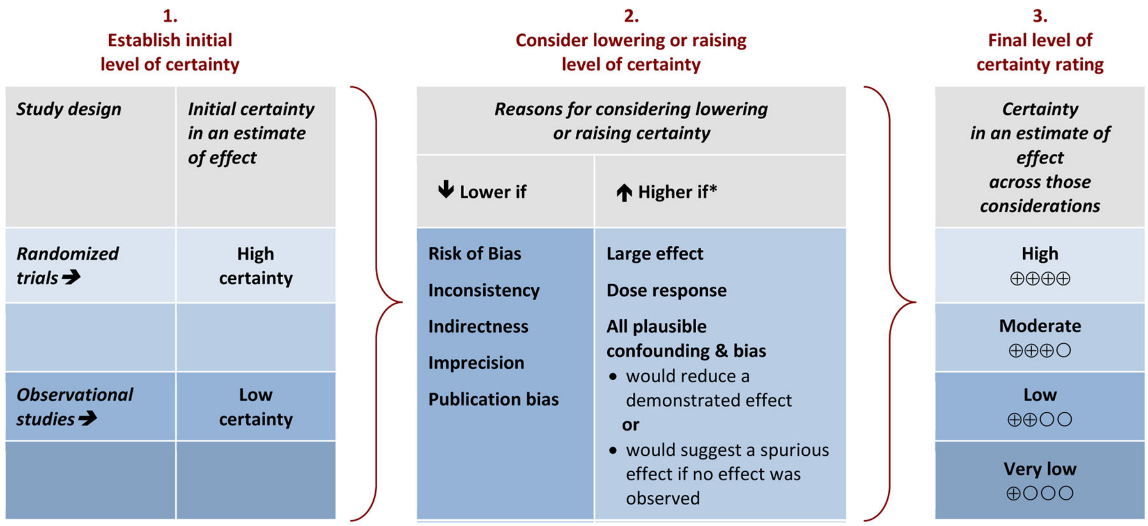

Figure 1.
GRADE's approach to developing certainty ratings across a body of evidence for each outcome based on a systematic review and across outcomes (lowest quality across the outcomes critical for decision-making).
*upgrading criteria are usually applicable to observational studies only.
Adapted from "Methodological idiosyncracies, frameworks and challenges of non-pharmaceutical and nontechnical treatment interventions" (Schunemann 2013)