

Navigating the Scientific Evidence to Improve Prevention:

***A Proposal to Develop A Transparent and Systematic Methodology
to Sort the Scientific Evidence Linking Environmental Exposures to
Reproductive Health Outcomes***

Discussion Document to Workshop Participants

July 29, 2009

*All scientific work is incomplete -
whether it be observational or experimental.
All scientific work is liable to be upset or modified by advancing knowledge.
That does not confer upon us a freedom to ignore the knowledge we already have,
or to postpone the action that it appears to demand at a given time.[1]*

*Sir Bradford Hill
1965 address to the Royal Society of Medicine*

Discussion Document to Workshop Participants

Prepared by the Navigating the Scientific Evidence to Improve Prevention Workshop Organizing Committee

Patrice Sutton, MPH, Co-Chair

University of California San Francisco, Program on Reproductive Health and the Environment

Heather Sarantis, MS

Commonweal/ Collaborative on Health & the Environment

Julia Quint, PhD

(Retired), Chief, Hazard Evaluation System and Information Service,
California Department of Public Health

Mark Miller, MD, MPH

University of California San Francisco, Pediatric Environmental Specialty Unit

Michele Ondeck, RN, MEd

Magee-Womens Hospital of University of Pittsburgh Medical Center

Rivka Gordon, PA-C, MHS

Association of Reproductive Health Professionals

Tracey Woodruff, PhD, MPH, Co-Chair

University of California San Francisco, Program on Reproductive Health and the Environment

INTRODUCTION.....	5
I. STATEMENT OF PROBLEM	5
<i>Reproductive Environmental Health Science Is Rapidly Evolving.....</i>	<i>5</i>
<i>Taking Action To Prevent Harm.....</i>	<i>7</i>
II. KEY COMMONALITIES AND DIFFERENCES IN THE USE OF SCIENTIFIC EVIDENCE IN CLINICAL AND ENVIRONMENTAL HEALTH SCIENCES	8
III. OVERVIEW: HOW SCIENTIFIC EVIDENCE IS INCORPORATED INTO CLINICAL PRACTICE	11
<i>Evidence-Based Medicine.....</i>	<i>11</i>
<i>Methods for Incorporating Evidence Into to Clinical Decision-Making</i>	<i>12</i>
<i>Identifying the Best of Best Practices.....</i>	<i>15</i>
IV. A PROPOSAL FOR THE NAVIGATION GUIDE	19
1. Study Question.....	22
2. Is the Substance on an Authoritative List?	22
3. Review Evidence	25
4. Rate Quality of Each Piece of Evidence.....	27
5. Determine Strength of Evidence.....	29
6. No Action At This Time.....	29
7. Take Action to Prevent Exposure.....	29
8. Grade Strength of Recommendations	31
9. Craft Effective Anticipatory Guidelines and Patient Messages.....	34
10. Clinician Provides Anticipatory Guidelines.....	34
V. FUTURE DIRECTIONS	35

List of Boxes, Figures and Appendices

Box A. Abstract of a Cochrane Review

Box B. American College of Obstetrics and Gynecologists (ACOG) Methodology for Evaluating the Scientific Evidence Grading Recommendations for Clinical Practice

Box C. Organizations that have endorsed or that are using GRADE

Figure 1: Streams of Evidence for Chemical Toxicity Assessment in Clinical and Environmental Health Sciences

Figures 2 A and 2 B: GRADE's Approach to Rating the Quality of Evidence and Grading the Strength of Recommendations

Figures 3 and 4: Proposed *Navigation Guide*

Figure 5: Criteria for "Authoritative Lists"

Figure 6: Proposed Lists to Evaluate According to "Authoritative Lists Criteria"

Figure 7: Proposed Inputs to be Used to Review Evidence

Figure 8: Criteria for Rating Quality of Individual Studies

Figure 9: Classifications for Levels of Reproductive Toxicity

Figure 10: Grading the Strength of a Recommendation

Appendix 1. Detailed Description of Classification of Levels of Toxicity

Appendix 2. The Wingspread Statement on the Precautionary Principle (January 1998)

Introduction

This paper is a pre-meeting discussion document for participants to help stimulate thinking and organize the discussion at the August 10-11, 2009 workshop, *Navigating the Scientific Evidence to Improve Prevention*. The document outlines a proposal for the *Navigation Guide*, a transparent and systematic methodology to sort the scientific evidence linking environmental exposures to reproductive health outcomes. The purpose of the *Navigation Guide* is to provide the evidence-based foundation for the timely development of anticipatory guidance in clinical care settings. Anticipatory guidance is prevention and other advice provided by a health care provider to her/his patients. For example, anticipatory guidance in pediatrics is a routine part of well-child care.[2]

The paper begins with a statement of the problem --- reproductive environmental health science is rapidly growing and the evidence has reached a threshold that demands timely action to prevent harm. Next, we describe the key commonalities and differences in the use of scientific evidence in clinical and environmental health sciences. We then provide an overview of how scientific evidence is currently incorporated into clinical practice through systems of evidence-based medicine. We conclude with a specific proposal for the *Navigation Guide* that builds on the strengths of similarities and bridges the differences in approaches between clinical and environmental health sciences.

I. Statement of Problem

Reproductive Environmental Health Science Is Rapidly Evolving

The fact that chemicals can harm human reproduction has been known since Roman times, when lead was first recognized to cause miscarriage and infertility in women and men.[3],[4] Over the past 60 years it has become clear that: (1) the placenta does not protect the fetus from damaging chemicals;[5] (2) the fetus can be uniquely sensitive to chemical exposures;[6, 7], and (3) intergenerational harm can result from *in utero* chemical exposures.[8] These discoveries stemmed from exposure to drugs and higher levels of environmental exposure than typically encountered by the general population. Hence it was generally assumed that environmental exposures experienced by an average person living in the U.S. would be below levels of reproductive harm. A rapidly expanding body of scientific evidence documents that this assumption has not held true.

Scientific indicators of declining reproductive function and increasing rates of reproductive illnesses since the mid-20th century suggest our reproductive health and, ultimately, our reproductive capacity are under strain.[9-12] The burgeoning evidence from studies of laboratory animals and human populations is further amplified by signals from wildlife showing altered reproductive performance in wild populations of annelids, mollusks, crustaceans, insects, fish, amphibians, and other species.[13, 14]

These trends in reproductive health have occurred in roughly the same time frame in which human exposure to both natural and synthetic chemicals has dramatically increased. Approximately 87,000 chemical substances are registered for use in U.S. commerce as of 2006, with about 3,000 chemicals manufactured or imported in excess of 1 million pounds each,[15] and 700 new industrial chemicals introduced into commerce each year.[16] These chemicals are distributed throughout homes,

workplaces and communities, and contaminate food, water, air and consumer products. Everyone in the U.S. has measurable levels of multiple environmental contaminants in their body.[17]

There is accumulating evidence that environmental exposure to low levels of chemicals can adversely impact reproductive health.[18, 19] Studies have demonstrated that the levels of chemicals that an average person is exposed to can perturb biological processes, such as preventing genes from functioning normally and interfering with the hormonal regulation critical to healthy reproduction.[20, 21] For example, certain chemicals in commonly used plastics [22] and persistent pesticides [23, 24] share the ability to alter the endocrine, neurological and/or other biological systems.

Exposure to plastics and pesticides is ubiquitous:

- Biological monitoring has shown that over 90% of people in the United States are exposed to bisphenol-A (BPA), a chemical found in many everyday places, including polycarbonate plastic containers and can linings.[25, 26] Studies in animals show that exposure to BPA during critical windows of development can result in permanent alterations to the reproductive system in a number of ways, thus increasing the risk of future health problems.[27, 28] Similarly, phthalates are a class of chemicals used as plasticizers and solvents in industrial, medical and consumer products. Findings in experimental animal studies document that exposure to some phthalates can cause reproductive damage. Because human exposure is so widespread, these data pose human health concerns.[22] The urgency of this concern is underscored by preliminary biological monitoring evidence that documented occupationally exposed populations have significantly higher urinary metabolite levels of some of these toxic phthalates than the general public.[29]
- Pesticides have been detected in human urine,[30] semen,[31] breast milk,[32, 33] ovarian follicular fluid,[34, 35] cord blood,[36, 37] and amniotic fluid,[38, 39] and are prevalent in food,[40] water,[41] and homes.[42, 43] Some pesticide exposures can interfere with all developmental stages of reproductive function in adult females,[44] and are associated with adverse outcomes that occur throughout the life course of males and females, including sterility in males, spontaneous abortion, diminished fetal growth and survival, and childhood and adult cancer.[45-49]

Plastics and pesticides are only two of the many chemical exposures encountered in daily life. Over 10,000 ingredients are used in personal care products; nearly 90% of these ingredients have not been evaluated for safety by any publicly accountable institution. People apply an average of 126 unique ingredients on their skin daily.[50]

In general, the human reproductive system is vulnerable to biological perturbations, particularly when these changes occur during critical windows of development. Even subtle perturbations caused by chemical exposures may lead to important functional deficits and increased risks of disease and disability in infants, children and across the entire span of human life.[10, 51]

The well-established linkage between environmental chemicals at higher levels and adverse reproductive and developmental health outcomes has been strengthened and

expanded over the past 20 years. The body of evidence has also invalidated previous assumptions about the benign nature of “low-level” environmental exposures,[18, 19] for example, in regards to neurological, reproductive and developmental health outcomes.[52, 53] While many scientific questions remain, a delay in the recognition and control of toxic environmental exposures translates into excess morbidity and mortality. The strength of the evidence is sufficiently high that leading scientists and reproductive health providers and other health care practitioners have called for timely action to prevent harm.[12, 51, 54]

Taking Action To Prevent Harm

Historically, clinicians and other health care practitioners have helped spur preventive public policy action on environmental and other public health issues. For example, physician involvement played an important role in shifting the public debate on smoking, and had they been active earlier in tobacco control campaigns, many people could have been spared immense suffering. Professional organizations of physicians have consistently called for regulatory and other efforts to address the environmental threats to human health. For example, the American Medical Association (AMA) has adopted policies promoting the incorporation of environmental health into medical education, supporting reforms in chemical policy, and addressing mercury exposure and other key environmental health issues. Most recently AMA has made a major commitment to participate in actions to address climate change and adopted a policy to promote the engagement of clinicians and policy makers in creating a healthy and sustainable food system. In 1968, two years before the 1970 Clean Air Act established the National Ambient Air Quality Standards for particulates, the California Medical Association (CMA) established policy to “vigorously support all rational efforts for the control of air pollution,” as well as to “urge the support of studies and the enactment of laws that will assure a healthful air supply in the future.” In the ensuing 40 years, the CMA has followed this pioneering step by enacting numerous additional policies calling for increasingly comprehensive steps to protect the public from the health effects of air pollution.^a

Clinical practice offers a key point of intervention to prevent harm from hazardous environmental exposures. The American Academy of Pediatrics has had an environmental health committee for over half a century, and since 1999, has published a clinicians’ handbook for the prevention of childhood diseases linked to environmental exposures.[55] This experience is directly applicable to reproductive health. Many individuals hoping to bear children are intensely interested in the impact of environmental exposures on their pregnancies and the health of their future children. Their health care provider can serve as a science-based source of guidance on how to avoid potentially adverse exposures. More importantly, many people who may eventually have or want to have children lack awareness of potential risks to their fertility and their future children’s health. In addition to the current queries about a patient’s alcohol and smoking history, clinicians need to be prepared to provide anticipatory guidance and respond to patient inquiries about hazardous environmental exposures encountered at

^a California Medical Association HOD Resolution 18-68 “Air Pollution”, HOD 123-87 “Air Pollution Crisis,” HOD 120-92 “Health Consequences of Ambient Air Pollution,” HOD 104-01 “Diesel Exhaust and Zero Emission Vehicles, HOD 105-02 “Air Pollution, Energy and Health,” and HOD 118-04 “Reducing Sources of Diesel Exhaust.” The American Thoracic Society, the American College of Cardiology, and the American Academy of Pediatrics all also urged the U.S. Environmental Protection Agency to adopt the most protective particulate matter air standards.

home, at work and in the community.[56] Thus, within and beyond the clinic, the active participation of well-informed health professionals is critical to translating scientific findings as they unfold into policies to improve health outcomes on a larger scale.

Through a process of engagement with hundreds of individuals,^b we identified the absence of a roadmap for evaluating the scientific evidence in a timely manner as one factor that impedes the translation of environmental health science into clinical health practice. While there are many steps and complexities involved in the use of current best evidence in health care settings, the process can be accelerated when knowledge-based information is readily available.[57] Therefore, we undertook a collaborative process to develop a timely, transparent and systematic methodology to sort the scientific evidence linking environmental exposures to reproductive health outcomes.

In the next section, we review the current approaches employed in the clinical and environmental sciences to evaluate the scientific evidence and make recommendations to preserve human health. Finally, we propose a methodology that builds on the strengths of similarities and bridges the differences in approaches between clinical and environmental health sciences and practices.

II. Key Commonalities and Differences in the Use of Scientific Evidence in Clinical and Environmental Health Sciences

Environmental health and clinical medicine share many common approaches to evaluating the scientific evidence linking exposures to exogenous compounds and health. For example, both disciplines:

- Rely on similar information streams to ascertain the relationship between exposure to a synthetic chemical and potential adverse health effects. *In vitro* and *in vivo* testing and observational human studies are part of the evidence basis in both clinical and environmental health.
- Lack direct experimental human data (randomized control trials (RCTs)) to assess exposure to chemicals with reproductive and developmental health toxicity. The evidence base for medication treatment of pregnant or lactating women does not include RCTs because it is unethical and undesirable to expose such women to medications of unknown risk in clinical studies.[58] The teratogenic risk in human pregnancy was still undetermined for 91.2 percent of drug treatments approved in the U.S. between 1980 and 2000.[59] For ethical reasons, experimental human evidence is also not part of the evidence stream in

^b The need to develop a science-based decision tool was identified through the interactive engagement of hundreds of individuals including scientists, advocates and representatives of communities disproportionately affected by environmental hazards to health. The 2007 Summit on Reproductive Health and the Environment convened by the University of California San Francisco and Commonwealth's Collaborative on Health and the Environment was attended by more than 400 community group members, clinicians, researchers, and policy makers from around the world, and was allied with 18 leading professional societies in the field of reproductive health. Summit recommendations that related to the translation of the emerging scientific findings in clinical and policy arenas were further honed through five planning meetings involving an interdisciplinary group of 16 scientists, and child, environmental and reproductive health advocates, scores of additional conversations with key informants and consultation with the scientific literature.

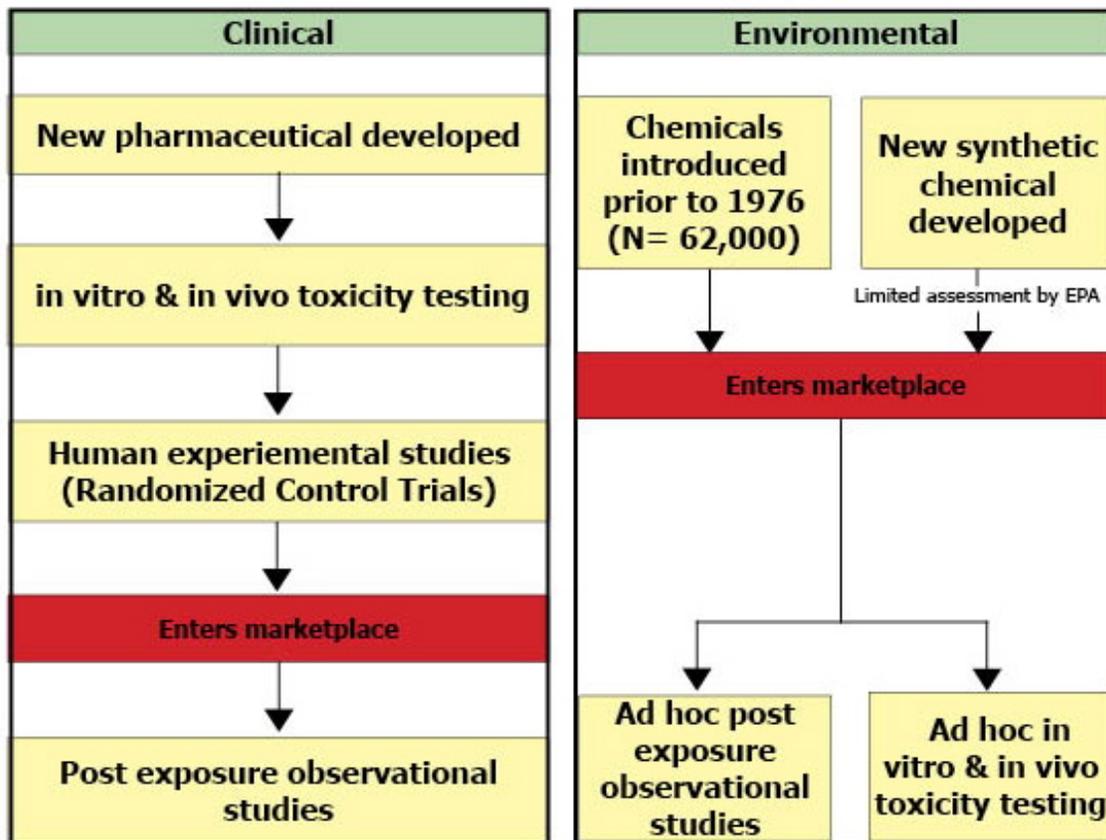
the realm of environmental health.

- Adhere to a hierarchy of evidence that places higher value on evidence that offers greater protection against bias and random error,[11, 60] and utilize explicit methods to identify, select, and critically appraise relevant research.[61, 62]
- Distinguish between the conduct of science and the practice of bringing the science to bear on a real world issue. In the clinical arena, there is a distinction between conducting clinical research and developing clinical practice guidelines, which incorporate research findings, patient values and preferences and other factors, depending on the system. In the environmental arena, there is also a distinction between conducting science and incorporating the scientific evidence, population characteristics and preferences and many other factors into practice. Thus, both disciplines share the need to integrate the overall context in which the evidence or intervention would apply, and each demand that the interpretation of evidence be effectively communicated.[63]^(p.24)

Environmental health and clinical medicine also have key differences regarding the evaluation of scientific evidence on the relationship between exposures to exogenous compounds and potential adverse health effects. For example:

- The type of experimental data available to each discipline differs. In the clinical arena RCTs are generally the “gold standard” of evidence for medical diagnostic and treatment decisions (except, as above, where RCTs are prohibited for testing pharmaceuticals on pregnant or lactating women). In contrast, because ethics precludes RCTs from the environmental health evidence stream, the primary experimental data in environmental health science stems from animal toxicity testing.
- The relative weight given to human observational studies differs. In clinical sciences, human observational studies are rated as less valuable to the evidence stream than a well-conducted RCT. In environmental health sciences, if human observational data are available and of sufficient quality, these data are always used and are afforded greater weight than results from animal studies.
- The timing of toxicity evaluation by regulatory agencies in relationship to a substance’s entry into the marketplace differs (Figure 1). In the clinical setting, before a drug can be tested in humans, the company or sponsor is required to perform *in vitro* and *in vivo* laboratory tests to discover how the drug works and whether it’s likely to be safe and work well in humans. In contrast, the vast majority of chemicals in commercial circulation have entered the marketplace without comprehensive and standardized information on their reproductive or other chronic toxicities.[64]

Figure 1: Streams of Evidence for Chemical Toxicity Assessment in Clinical and Environmental Health Sciences



- The weighing of the benefits and risks of human exposure differs. Before a drug is approved for sale, an independent and unbiased review must establish that a drug's health benefits outweigh its known risks.[65] Drugs are prescribed and administered in a manner that relates a specific exposure to this risk-benefit decision. This is effectively a regulatory requirement that human exposure to pharmaceuticals does not occur in the absence of some potential benefit greater than the known risks. Currently, there is no comprehensive comparable weighing of health benefits and risks in the environmental arena.[66] The benefits of environmental chemicals are mostly not health-related, and exposures vary and may or may not be significant depending on the toxicity of the agent. The current underlying regulatory decision in environmental health is to permit population exposure until such time the risks of exposure are deemed “unacceptably” high (i.e., first expose, then see if there’s harm).
- The ability to observe an adverse or beneficial health outcome differs. Very large population-wide environmental health impacts may not be observable on an individual level because individual risks for common exposures may be relatively small. For example, children’s blood lead levels are inversely related to IQ scores. A 6.2 IQ point decrement is estimated for an increase in blood lead levels

from < 1 to 10 micrograms per deciliter (95% CI, 3.8–8.6).[67] While for an individual child the effects of low-level exposure to lead are difficult to discern, on a population basis, the enormous societal consequences are apparent.[68] Assuming a large population with a mean IQ of 100 and a normal distribution, a five-point downward shift in IQ results in a 57% increase in the number of children with IQ scores in the extremely low-range (<70), and a 40% reduction in the number of children in the extremely high range (>130). Although not clinically apparent, these population-wide developmental health impacts also have demonstrated adverse impacts on individual health and wellbeing.[69]

In summary, clinical and environmental health sciences share the rules of scientific rigor and succeed when communication of the science is evidence-based, clear and concise. *In vivo* and *in vitro* testing and human observational studies are relied on in both disciplines to regulate substances with reproductive and developmental toxicity. For other types of chemical toxicities, the disciplines diverge in terms of the centrality of RCTs (to regulate diagnostic tools and treatments in clinical sciences) and human observational and animal data (to regulate chemical toxicities in environmental sciences). In the clinical sciences, exposure to an exogenous chemical is defined and permitted only with *a priori* knowledge about the substances toxicity and an evidence-based benefit to an exposed individual. In the environmental sciences, exposure levels are generally not known, and exposure is permitted most often absent *a priori* knowledge about the substances toxicity and with no assessment of the health or other benefits of exposure to either an individual or the overall population. Finally, even very large population health impacts may not be observable on an individual level. The *Navigation Guide* must incorporate an approach that will maximize these common strengths and bridge these key differences.

III. Overview: How Scientific Evidence is Incorporated into Clinical Practice

Evidence-Based Medicine

The purpose of the *Navigation Guide* is to bring scientific evidence in the realm of environmental health to bear on clinical practice. As such, the *Navigation Guide* seeks to integrate environmental health science into an evidence-based medicine framework.

Evidence-based medicine (EBM) is defined as, the “conscientious, explicit, and judicious use of the current best evidence in making decisions about the care of individual patients.[70] In general, systems of EBM combine: (1) existing scientific evidence; (2) clinical expertise; and (3) patient values and preferences to make diagnosis, prognosis and treatment decisions. An EBM clinical decision attempts to optimize the benefits to a patient by balancing each of these considerations.

Since the term EBM was first introduced in the scientific literature in 1991, there have been enormous advances in the incorporation of scientific evidence into healthcare decision-making on individual and policy levels.[61, 71] These trends are continuing at an accelerated pace such that a 2008 report by the Institute of Medicine (IOM) described EBM as the guiding framework for the development of healthcare delivery systems in the 21st century.[72] The IOM report describes how the incorporation of scientific evidence into health care decision-making continues to be transformed by the accelerating speed of scientific discovery, technological innovation and policy moves toward greater value and efficiency. These trends are equally applicable to environmental health sciences.

While the clinical care community has almost 20 years of research and practice in EBM systems to rate the quality of science and the strength of recommendations, there is no comparable transparent and comprehensive system for translating the science related to the health effects of environmental exposures into clinical practice.

Methods for Incorporating Evidence Into to Clinical Decision-Making

1. Systems to rate the strength of scientific evidence

A plethora of methodologies exist to evaluate health care research to guide clinical decision-making. In 2002, the Agency for Healthcare Research and Quality (AHRQ) conducted a study to describe existing systems used to rate the strength of the scientific evidence.[61] A goal was to provide guidance on “best practices” in rating the quality of individual studies that comprise the body of evidence on a specific question in health care. To this end, AHRQ identified three domains for grading the strength of a body of human evidence that are essential to robust review, i.e., quality, quantity and consistency, defined as:

Quality: Quality of an individual study is the extent to which a study’s design, conduct, and analysis have minimized selection, measurement, and confounding biases. Quality of the body of evidence is the aggregate of quality ratings for individual studies, predicated on the extent to which bias was minimized;

Quantity: Magnitude of effect, numbers of studies and sample size or power;

Consistency: For any given topic, the extent to which similar findings are reported using similar and different study designs.

AHRQ reviewed the available systems and found that of the 121 systems identified, 19 fully incorporated the elements deemed crucial to evaluating the quality of an individual study, and seven fully addressed all three domains for grading the strength of a body of evidence. The earliest system meeting all three necessary features was published in 1994, the next in 1999, and five were published in 2000.

An example of a methodology that met the AHRQ criteria for “best practices” for rating a body of evidence is the Cochrane Review.[73] The Cochrane Review is an example of a “systematic review.” Systematic reviews differ from traditional opinion-based narrative reviews in that they employ a rigorous methodology to evaluate a clearly formulated question using systematic and explicit methods to identify, select and critically appraise relevant research, and to collect and analyze data from the studies that are included in the review. Statistical methods (meta-analysis) may or may not be used to analyze and summarize the results of the included studies. The Cochrane Database of Systematic Reviews answers clinical questions about the effectiveness of treatment.[74] A Cochrane Review provides a summary of the state of weight of the evidence but does not make specific recommendations for treatment. An example of an abstract of a Cochrane Review is presented in Box A.

Box A. Abstract of a Cochrane Review*

[Intervention Review] Modification of the home environment for the reduction of injuries

Background

Injury in the home is extremely common, accounting for around a third of all injuries. The majority of injuries of children under five and people aged 75 and over, occur at home. Multifactorial injury prevention interventions have been shown to reduce injuries in the home. However, few studies have focused specifically on the impact of physical adaptations to the home environment and the effectiveness of such interventions needs to be ascertained.

Objectives

To review the evidence for the effect on injuries of modification of the home environment with a primary focus on interventions to reduce physical hazards.

Search strategy

We searched The Cochrane Library, MEDLINE, EMBASE, National Research Register and other specialised databases. We also scanned conference proceedings and reference lists. In addition, we contacted experts and trialists in the field. The searches were not restricted by language or publication status. The searches were last updated in December 2004.

Selection criteria

- Randomised controlled trials.
- Data collection and analysis
- All abstracts were screened by two authors for relevance, outcome and design. Two authors independently assessed methodological quality and extracted data from each eligible study.

Main results

We found 18 published and one unpublished trials. Trials were not sufficiently similar to allow pooling of data by statistical analyses, so this review takes a narrative form. Studies were divided into three groups based on the primary population sample; children (five studies), older people (14 studies) and the general population/mixed age group (no studies). None of the studies focusing on children demonstrated a reduction in injuries that might have been due to environmental adaptation in the home; one study reported a reduction in injuries and in hazards but the two could not be linked. Of the 14 included studies in older people, none demonstrated a reduction in injuries due to hazard reduction, although two demonstrated a reduction in falls that could be due to hazard reduction.

Authors' conclusions

There is insufficient evidence to determine the effects of interventions to modify environmental home hazards. Further interventions to reduce hazards in the home should be evaluated by adequately designed randomised controlled trials measuring injury outcomes. Recruitment of large study samples to measure effect must be a major consideration for future trials.

Plain language summary

More evidence is needed to show whether or not altering the physical home environment by removing potential hazards reduces injuries. Injuries in the home are very common. Most of the injuries of older people and children under five occur at home. Many people are encouraged to alter their home to try to reduce such injuries. Common alterations include the installation of locks on cupboards and covers on electrical sockets, improvement of lighting in halls and stairways, and the removal of rugs and other falls hazards. The review found that there is insufficient evidence from trials to show that such changes reduce the number of injuries in the home but does not conclude that these interventions are ineffective. Home alterations need to be evaluated by larger and better designed trials.

* Lyons RA, John A, Brophy S, Jones SJ, Johansen A, Kemp A, Lannon S, Patterson J, Rolfe B, Sander LV, Weightman A. Modification of the home environment for the reduction of injuries. *Cochrane Database of Systematic Reviews*, Issue 3, 2009.

2. Systems to grade the strength of recommendations

To communicate the evidence to clinicians, most professional practice guidelines for clinical care combine both a “rating” of the quality of the evidence and a “grading” of the strength of the recommendation that is derived from the evidence.[60] The quality of evidence is separated from the strength of recommendations in recognition of the role that patient values and preferences as well as clinical and social circumstances play in formulating practice recommendations. For example, the American College of Obstetricians and Gynecologists (ACOG) methodology for evaluating the scientific evidence and grading recommendations for clinical practice is presented in (Box B). This methodology results in recommendations such as,

“ ... based on good and consistent scientific evidence (Level A), abdominal myomectomy is a safe and effective alternative to hysterectomy for treatment of women with symptomatic leiomyomas. Based primarily on consensus and expert opinion (Level C), leiomyomas should not be considered the cause of infertility ... without completing a basic fertility evaluation to assess the woman and her partner.”

The ACOG grading system is one of a myriad of taxonomies in use in clinical practice.

Box B. American College of Obstetrics and Gynecologists (ACOG) Methodology for Evaluating the Scientific Evidence Grading Recommendations for Clinical Practice
<p>The MEDLINE database, the Cochrane Library, and ACOG's own internal resources are used to conduct a literature search to locate relevant articles published between January 1985 and [relevant date]. The search is restricted to articles published in the English language. Priority is given to articles reporting results of original research, although review articles and commentaries are also consulted. Abstracts of research presented at symposia and scientific conferences are not considered adequate for inclusion.</p>
<p>Guidelines published by organizations or institutions such as the National Institutes of Health and ACOG are reviewed, and additional studies are located by reviewing bibliographies of identified articles. When reliable research is not available, expert opinions from obstetrician–gynecologists are utilized.</p>
<p>Studies are reviewed and evaluated for quality according to the method outlined by the US Preventive Services Task Force.</p> <ul style="list-style-type: none">I Evidence obtained from at least one properly designed randomized controlled trial.II-1 Evidence obtained from well-designed controlled trials without randomization.II-2 Evidence obtained from well-designed cohort or case-control analytic studies, preferably from more than one center or research group.II-3 Evidence obtained from multiple time series with or without the intervention. Dramatic results in uncontrolled experiments could also be regarded as this type of evidence.III Opinions of respected authorities, based on clinical experience, descriptive studies, or reports of expert committees.
<p>Based on the highest level of evidence found in the data, the recommendations are graded according to the following categories</p> <ul style="list-style-type: none">A The recommendation is based on good and consistent scientific evidenceB The recommendation is based on limited or inconsistent scientific evidenceC The recommendation is based primarily on consensus and expert opinion

Identifying the Best of Best Practices

The abundance of methodologies for rating the strength of evidence and grading the strength of recommendations can lead to confusion rather than clarity, and attempts have been made to address this concern.[60, 75] A common, sensible approach that has built upon the strengths of existing systems and addressed shortcomings is The Grading of Recommendations, Assessment, Development and Evaluation (GRADE) system.[76] GRADE is based on contemporary principles of EBM.[71] GRADE systematically rates the quality of evidence and grades the strength of the recommendations to administer - or not administer- an intervention based on the tradeoffs between benefits on the one hand, and risks, burden and -potentially- costs on the other (Figures 2A and 2B). Grading of recommendations provides decision-makers with a qualitative estimate (strong or weak, with weak sometimes called “discretionary”) of the confidence in these estimates.

As of July 2009, 34 organizations throughout the world have endorsed or are using GRADE, including the Endocrine Society, World Health Organization, The Cochrane Collaboration – International, and the U.S. Agency for Healthcare Research Quality (AHRQ) (Box C). A recent series of articles in the *British Medical Journal* outlines the progress to date of the GRADE system.[77-82]

The developers of GRADE cite the following advantages of GRADE over other systems:[77]

- Developed by a widely representative group of international guideline developers;
- Clear separation between quality of evidence and strength of recommendations;
- Explicit evaluation of the importance of outcomes of alternative management strategies;
- Explicit, comprehensive criteria for downgrading and upgrading quality of evidence ratings;
- Transparent process of moving from evidence to recommendations;
- Explicit acknowledgment of values and preferences;
- Clear, pragmatic interpretation of strong versus weak recommendations for clinicians, patients, and policy makers; and
- Useful for systematic reviews and health technology assessments, as well as guidelines.

Figure 2A: GRADE's Approach to Rating the Quality of Evidence and Grading the Strength of Recommendations

**Step 1: Rate the Quality of Evidence
(High, moderate, low, very low)**

Factors in deciding on quality of evidence (1)

<p>Factors that might increase quality of evidence</p> <ul style="list-style-type: none"> • Large magnitude of effect • Situations in which all plausible biases would decrease the magnitude of effect • Dose-response gradient
<p>Factors that might decrease quality of evidence</p> <ul style="list-style-type: none"> • Study limitations • Inconsistency of results • Indirectness of evidence • Imprecision • Publication bias

**Step 2: Grade the Strength of the Recommendations
(Strong or Weak [also called Discretionary])**

Determinants of Strength of Recommendation (2)

Factor	Comment
Balance between desirable and undesirable effects	The larger the difference between desirable and undesirable effects, the higher the likelihood that a strong recommendation is warranted. The narrower the gradient, the higher the likelihood that a weak recommendation is warranted.
Quality of Evidence (See Figure 2B)	The higher the quality of evidence, the higher the likelihood that a strong recommendation is warranted.
Values and preferences	The more values and preferences vary, or the greater the uncertainty in values and preferences, the higher the likelihood that a weak recommendation is warranted.
Costs (resource allocation)	The higher the costs of an intervention — that is, the greater the resources consumed — the lower the likelihood that a strong recommendation is warranted.

(1) Guyatt GH, Oxman AD, Kunz R, Vist GE, Falck-Ytter Y, Schünemann HJ and for the GRADE Working Group. What is "quality of evidence" and why is it important to clinicians? *BMJ* 2008;336:995-998.

(2) Guyatt GH, Oxman AD, Kunz R, Falck-Ytter Y, Vist GE, Liberati A, Schünemann HJ and for the GRADE Working Group. GRADE: going from evidence to recommendations. *BMJ* 2008;336:1049-1051.

Figure 2B: How GRADE Links Quality of Evidence to Strength of Recommendations (60)

Rating of evidence quality	Clarity of risk/benefit	Description of supporting evidence	Implications
Strong recommendations High quality evidence	Benefits clearly outweigh harms and burdens, or vice versa	Consistent evidence from well-performed RCTs or exceptionally strong evidence from unbiased observational studies ^a	Recommendation can apply to most patients in most circumstances. Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality evidence	Benefits clearly outweigh harms and burdens, or vice versa	Evidence from RCTs with important limitations (inconsistent results, methodological flaws, indirect or imprecise evidence), or unusually strong evidence from unbiased observational studies	Recommendation can apply to most patients in most circumstances. Further research (if performed) is likely to have an impact on our confidence in the estimate of effect and may change the estimate.
Low quality evidence	Benefits clearly outweigh harms and burdens, or vice versa	Evidence for at least one critical outcome from observational studies, from RCTs with serious flaws, or indirect evidence	Recommendation may change when higher quality evidence becomes available. Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality evidence (very rarely applicable)	Benefits clearly outweigh harms and burdens, or vice versa	Evidence for at least one of the critical outcomes from unsystematic clinical observations or very indirect evidence	Recommendation may change when higher quality evidence becomes available; any estimate of effect, for at least one critical outcome, is very uncertain.
Weak recommendations High quality evidence	Benefits closely balanced with harms and burdens	Consistent evidence from well-performed RCTs or exceptionally strong evidence from unbiased observational studies	The best action may differ depending on circumstances or patient or societal values. Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality evidence	Benefits closely balanced with harms and burdens	Evidence from RCTs with important limitations (inconsistent results, methodological flaws, indirect or imprecise evidence), or unusually strong evidence from unbiased observational studies	Alternative approaches likely to be better for some patients under some circumstances. Further research (if performed) is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality evidence	Uncertainty in the estimates of benefits, harms, and burdens; benefits may be closely balanced with harms and burdens	Evidence for at least one critical outcome from observational studies, from RCTs with serious flaws, or indirect evidence	Other alternatives may be equally reasonable. Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality evidence	Major uncertainty in the estimates of benefits, harms, and burdens; benefits may or may not be balanced with harms and burdens	Evidence for at least one critical outcome from unsystematic clinical observations or very indirect evidence	Other alternatives may be equally reasonable. Any estimate of effect, for at least one critical outcome, is very uncertain.

Modified from Schunemann *et al.* (22).

^a Exceptionally strong evidence from unbiased observational studies includes: 1) evidence from studies that yield estimates of the treatment effect that are large and consistent; 2) evidence in which all potential biases may be working to underestimate an apparent treatment effect, and therefore, the actual treatment effect is likely to be larger than that suggested by the study data; and 3) evidence in which a dose-response gradient exists.

Box C. Organizations that have endorsed or that are using GRADE

- Guidelines for World Health Organization Guidelines
- Endocrine Society Clinical Guidelines - USA
- American College of Chest Physicians Guidelines - USA
- UpToDate - Putting Clinical Information Into Practice - USA
- Agenzia sanitaria regionale, Bologna - Italia
- Ministry of Health and Long-Term Care, Ontario - Canada
- Surviving Sepsis - International
- Ärztliches Zentrum für Qualität in der Medizin - Germany
- American Thoracic Society – USA
- American College of Physicians – USA
- The Cochrane Collaboration - International
- Kidney Disease: Improving Global Outcome – International
- European Society of Thoracic Surgeons - International
- British Medical Journal - UK
- Journal of Infection in Developing Countries - International
- Agency for Healthcare Research and Quality (AHRQ) – USA
- Society of Critical Care Medicine (SCCM) - USA
- National Institute for Clinical Excellence (NICE) - UK
- Norwegian Knowledge Centre for the Health Services - Norway
- The University of Pennsylvania Health System Center for Evidence-based Practice - USA
- German Center for Evidence-based Nursing "sapere aude" - Germany
- Evidence-based Nursing Südtirol, Alto Adige – Italy
- Society for Vascular Surgery - USA
- BMJ Clinical Evidence - UK
- EBM Guidelines - Finland/International
- Polish Institute for EBM - Poland
- European Respiratory Society (ERS) - Europe
- Japanese Society for Temporomandibular Joint - Japan
- National Board of Health and Welfare - Sweden
- COMPUS at The Canadian Agency for Drugs and Technologies in Health (CADTH) Canada
- Infectious Diseases Society of America - USA
- Spanish Society for Family and Community Medicine
- Emergency Medical Services for Children National Resource Center - USA
- The Swedish Council on Technology Assessment in Health Care - Sweden

Source: <http://www.gradeworkinggroup.org/society/index.htm>

IV. A Proposal for the *Navigation Guide*

The proposed *Navigation Guide* is an open-source methodology that proceeds from an EBM framework in general and GRADE specifically. The rationale for this decision is based on:

- EBM is expected to play an increasingly key role in the practice of medicine;
- A system widely used in clinical practice will be more recognizable and thus acceptable to a clinical audience than a novel system developed from an environmental health framework;
- The reported emerging and international consensus around the use of GRADE is expected to increase the *Navigation Guide*'s acceptance and uptake by its clinical target audience;
- GRADE incorporates scientific rigor and evidence-based recommendations for prevention that are integral to the environmental health framework;
- GRADE does not require quantitative assessments of risk, which, as described below, can lead to prolonged inaction in the field of environmental health; and
- It is efficient to build the *Navigation Guide* on a well-established knowledge base, and GRADE brings a depth and breadth of existing scholarship and practice.

It is important to note that the GRADE Working Group recommends against modifications to GRADE “because the elements of the GRADE process are interlinked, because modifications may confuse some evidence and guideline users, and because such changes compromise the goal of a single system with which clinicians, policy-makers and patients can become familiar.”[76] If we decide to use GRADE as a framework, we will need to determine how to best coordinate this effort with the GRADE Working Group.

The following section outlines a specific methodology for the *Navigation Guide* (Figures 3-10). Figure 3 presents an overview of architecture of the *Navigation Guide*, and Figure 4 provides more information about each of its components.

The structure of the *Navigation Guide* has two main components:

- (1) Decide if the exposure is toxic to human reproduction based on an authoritative list, or, if not on a list, a review of the quality of all available evidence;
- (2) Grade a recommendation for prevention of exposure by combining two factors: the strength of the body of evidence that the toxicity decision is based on, AND contextual factors, including public health best practices, exposure assessment, availability of safer alternatives, patient values and preferences, and other considerations.

The rationales for the framework overall and for each section of the methodology are described in detail, below.

Figure 3: Proposed Navigation Guide

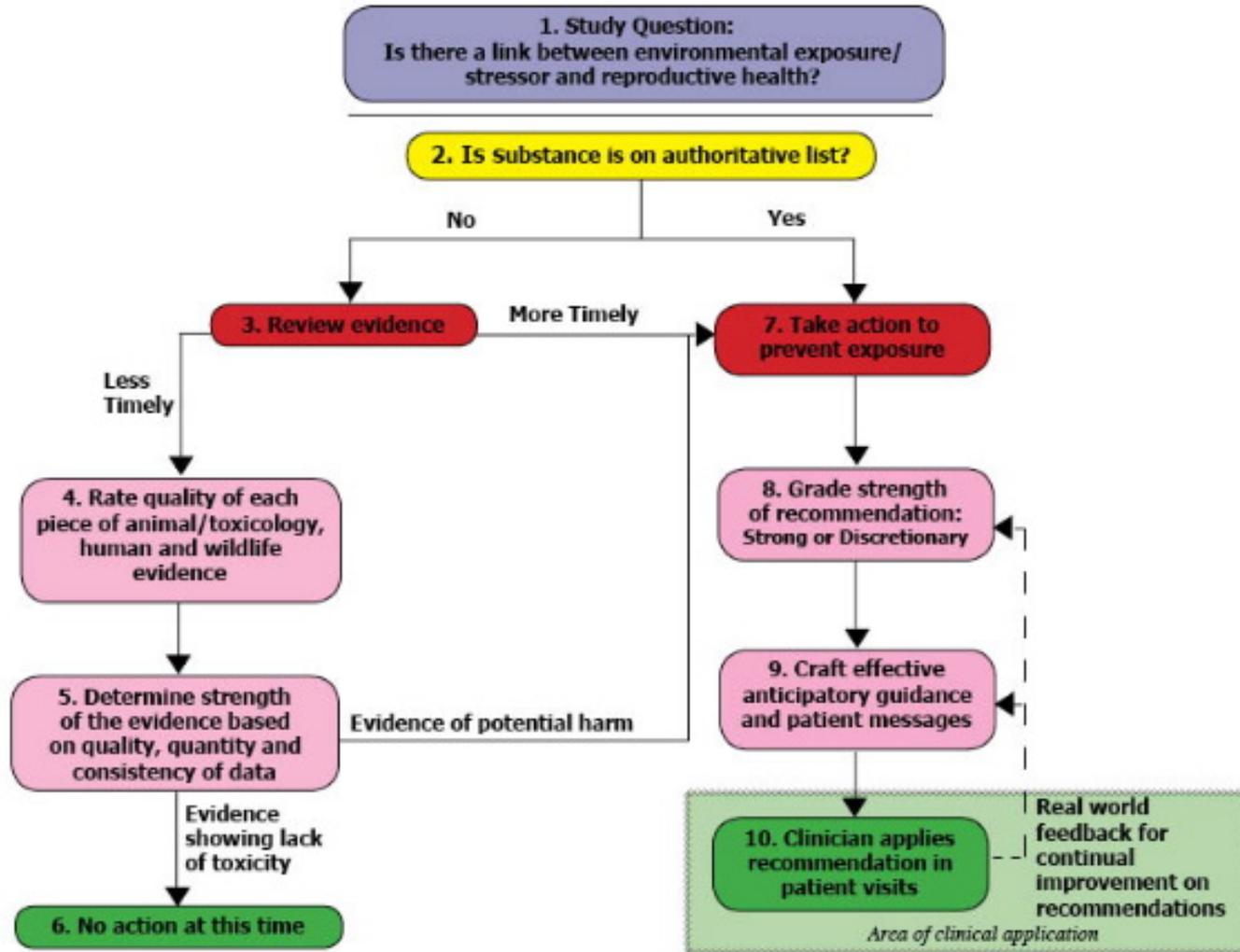
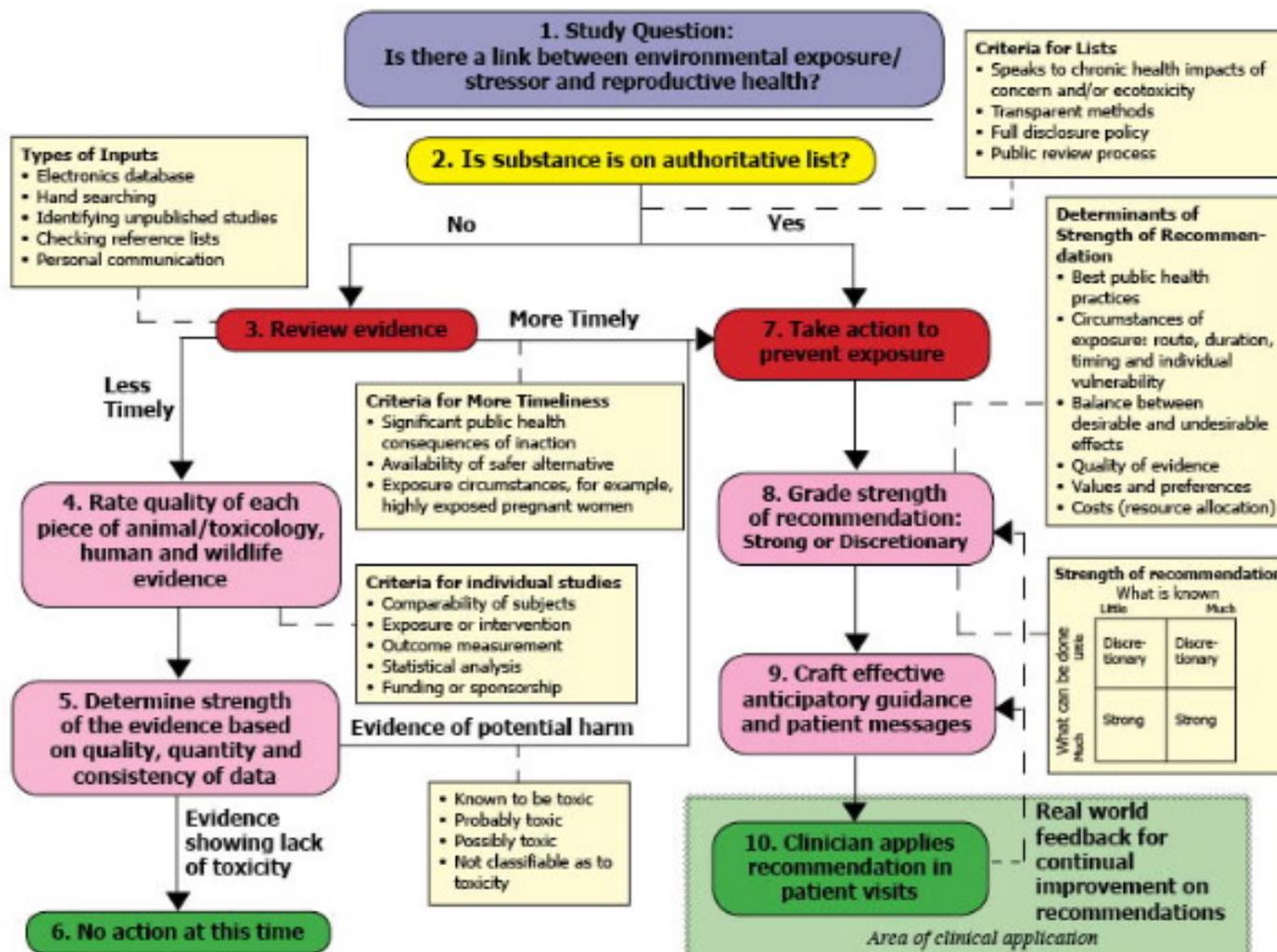


Figure 4: Proposed Navigation Guide



1. Study Question

The generic study question to be addressed by the *Navigation Guide* is:

Is there a link between an environmental exposure/stressor and a reproductive health outcome? In practice, this question will be refined to answer specific exposures or topics.

Definitions and Scope:

Reproductive Health Outcomes: All aspects of future reproductive health throughout the life course, including conception, fertility, pregnancy, child and adolescent development, and adult health. This includes chronic disease outcomes that may have a negative effect on reproductive health, such as cancer, asthma, diabetes, obesity, cardiovascular disease and metabolic syndrome. It includes eco-toxicity based on assumption that human reproductive health requires a healthy ecosystem to support life on earth. While this expansive framework is scientifically accurate, it poses the question: how to specify what is and what is not a reproductive health outcome? One way to focus the initial application of the *Navigation Guide* would be to prioritize the examination of those health outcomes linked to environmental exposures and stressors incurred by vulnerable populations. This is based on the consensus observation of leading scientists that the “accumulated research evidence suggests that prevention efforts against toxic exposures to environmental chemicals should focus on protecting the fetus and small child as highly vulnerable populations.”[51]

Environmental Exposures/Stressors: Any and all environmental pollutants (physical, chemical, biological (pests, molds) agents), built environment, and social determinants of health (poverty, nutrition, etc.). For practical purposes, we will initially focus on environmental pollutants, and as we gain experience with the methodology, expand it to encompass this broader range of exposures/stressors. The sources of exposure include: industrial emissions, transportation, consumer products, human activities (work, hobbies, entertainment, cleaning, redecorating, cooking, eating, drinking, smoking, sleeping, medical care delivery), etc.

2. Is the Substance on an Authoritative List?

First and foremost, healthcare providers, their patients, institutions, and policy-makers currently lack the capacity to readily identify substances recognized as reproductive and developmental toxicants. Thus, the *Navigation Guide* proceeds from identifying chemicals that have been deemed to have toxicity relevant to reproductive health (as defined above). First, the *Navigation Guide* list will include chemicals named as reproductive or developmental toxicants on an “authoritative list,” defined below. This takes advantage of the reviews and evaluations that have already been conducted through weight of the evidence criteria that are consistent with the *Navigation* approach. The *Navigation Guide* list would also include carcinogens, asthmagens, chemicals that persist or bioaccumulate in humans and the environment, and other agents with recognized toxicity related to reproductive health. The nature of toxicity data for each chemical will be designated, i.e., one could search the list for “reproductive/developmental toxicity” or “carcinogenicity” and generate a list of

chemicals with that specific toxicity. The rationale for a *Navigation Guide* list that encompasses more than chemicals with reproductive/developmental toxicity is two-fold. As above, exposure to carcinogens and chemicals with other toxicities can also impact human reproductive health. Second, evidence about a wide range of toxicities is needed to ensure recommendations for prevention do not replace a reproductive toxicant with an agent with another type of toxicity.

Figure 5 presents the minimal criteria for “authoritative lists” i.e., relevance to health outcomes of concern and transparency of methods. Workshop participants will determine additional criteria subsequent to a group discussion of what features make a list more or less “authoritative.” These criteria would initially be applied to over 40 lists identified by the developers of the “GoodGuide,”^c and additional lists identified in the scientific literature (Figure 6). We will cast a wide net for what is known about chemical toxicities because there is no one single authoritative source. For example, of 934 chemicals identified as reproductive toxicants from six databases, only four chemicals^d were common to all six sources.[83] A wide range of goals, objectives and rigorous scientific approaches are used by authoritative government bodies such as the International Agency for Research on Cancer (IARC), U.S. Environmental Protection Agency (USEPA), California Environmental Protection Agency (Cal-EPA), and others, and these differences in purpose and methods yield different resultant lists.

In a manner comparable to incorporating *in vivo* and *in vitro* data into decisions to permit drugs to proceed to human experimental trials, chemicals with toxicity recognized by an authoritative body would proceed to “take action to prevent exposure.” As described in detail below, being on a list is just one of many factors that determine the strength of the recommendation.

Figure 5: Criteria for “Authoritative Lists”

Criteria	Explanation
Speaks to chronic health impacts of concern and/or ecotoxicity.	Focuses on exposures to environmental contaminants, particularly during critical periods of development (such as before conception and during pregnancy) and their potential effects (outcomes) on all aspects of future reproductive health throughout the life course, including conception, fertility, pregnancy, child and adolescent development, and adult health. This includes chronic disease outcomes which may have a negative effect on reproductive health, such as cancer, asthma, diabetes, obesity, cardiovascular disease and metabolic syndrome, and/or may be a hazard to the environment. This also includes ecosystem toxicity, based on the assumption that human reproductive health requires a healthy ecosystem to support life on earth.
Transparent methods	Published and reproducible methods.
Other Criteria TBD	

^c The GoodGuide is a tool for rating the toxicity of consumer goods. Its developer, Dara O’Rorke generously provided the “list of lists” to PRHE. See <http://www.goodguide.com/>

^d Ethylene glycol methyl ether (CAS 109-86-4); ethylene glycol monopropyl ether (CAS 110-49-6); ethylene glycol ethyl ether (CAS 110-80-5); and toluene (CAS 108-83-3).

Figure 6: Proposed Lists to Evaluate According to "Authoritative Lists Criteria"

Reproduction/Development Focused Lists	Other Lists
American Conference of Governmental Industrial Hygienists (ACGIH)	CA Department of Public Health Hazard Evaluation System and Information Service (HESIS)
California Prop 65	EPA-IRIS (Known/Likely)
Catalog of Teratogenic Agents	EU DSD (Dangerous Substances Directive) Carc. Cat 1 (REACH)
Center for the Evaluation of Risks to Human Reproduction (CERHR)	EU DSD Carc. Cat. 2 (REACH)
Development and Reproductive Toxicology (DART) Database	EU ESIS (European chemical Substances Information System) PBTs
Developmental Neurotoxicants (Grandjean & Landrigan)	EU ESIS vPvBs
EU Endocrine Disruptors (ED) Human	EU R45 "May cause cancer"
EU Endocrine Disruptors (ED) Cat. 1 + 2 (DHI Report)	EU R46 "May cause heritable genetic damage"
EU Endocrine Disruptors (ED) Overall Assessment	EU R49 "May cause cancer by inhalation"
EU Endocrine Disruptors (ED) Wildlife	EU REACH (Registration, Evaluation, Authorisation and Restriction of Chemical substances) Mut. Cat. 2
EU R60 "May impair fertility"	EU vPvBs
EU R61 "May cause harm to unborn child"	International Agency for Research on Cancer (IARC)
EU REACH Repro. Tox. Cat. 1	National Institute for Occupational Safety And Health (NIOSH) Potential Carcinogens
EU REACH Repro. Tox. Cat. 2	National Toxicology Program (NTP)-Known Carc
Haz-Map	National Toxicology Program (NTP)-Rsnbly Carc
Janovic and Drake	NTP Program Report on Carcinogens
Material Data Safety Sheet (MSDS)	SIN (Substitute it Now) List
National Toxicology Program (NTP) Repro/Dev Toxicant	Stockholm Persistent Organic Pollutants (POPs)
Navy and Marine Corps Public Health Center Technical Manual	Stockholm POPRC Recommendations & Review
Quantitative Structure Activity Relationship (QSAR) Methodologies	USEPA Hazardous Air Pollutants
Registry of Toxic Effects of Chemical Substances (RTECS)	USEPA Priority PBTs
Reprotect	USEPA TRI (Toxic Release Inventory) Substance Persistent Bioaccumulative and Toxic Substances (PBTs) List
Scorecard	USEPA Waste Min Priority Chems
	WA State PBTs

3. Review Evidence

If a substance is not identified in one of the authoritative lists, then the next step is to evaluate the existing evidence. The inputs to methodically and comprehensively capture what is known about the toxicity of an environmental exposure are listed in Figure 7. These sources of data include electronic databases of peer-reviewed literature and government documents, identifying hard copies of papers that are not available electronically, identifying unpublished data, such as conference proceedings, and checking reference lists. Once the inputs are collected there are two review options which differ in terms of how quickly a decision is made to act or not to act.

Figure 7: Proposed Inputs to be Used to Review Evidence

Source	Examples
Electronic databases	MEDLINE, Web of Science, etc. (Peer-reviewed literature and government documents). Note the importance of findings in other languages. The neurotoxicity of arsenic was initially reported in Japanese journals and these results have often been overlooked even in the most thorough risk assessments.
Hand Searching	For identifying published studies which are not electronically available. (Request for hard copies from authors or UCSF library).
Identifying Unpublished Studies, also known as Gray Literature	<p>Gray literature includes "http://ssrc.tums.ac.ir/SystematicReview/CDT.asp" conference proceedings, "http://ssrc.tums.ac.ir/SystematicReview/CDT.asp" theses, "http://ssrc.tums.ac.ir/SystematicReview/CTRDB.asp" clinical trials registries and other reports. Using gray literature in preparing systematic reviews is mandatory to overcome or alleviate publication bias.</p> <p>Note: This is a potential opportunity for on-going stakeholder engagement. For example, CHE Fertility Group and other partners can be a major resource in helping to compile and keep current the gray literature. CHE Fertility has plans to compile all relevant conference proceedings; Above the Fold compiles news articles about environmental health "mailto:AboveTheFold@newsletters.environmentalhealthnews.org" AboveTheFold@newsletters.environmentalhealthnews.org; etc.</p>
Checking Reference Lists	This is to follow up references from one article to another including those from previously published systematic reviews, to identify relevant reports. The process of following up references from one article to another is generally an efficient means of identifying studies for possible inclusion in a review.
Personal Communication	To know of studies not found in the previous steps, we will send a list of the studies we have found to the authors of those studies asking if they are aware of any other relevant studies (published or unpublished).

More Timely

Sir Bradford Hill, the statistician who pioneered the randomized clinical trial and who together with Richard Doll, was the first to demonstrate the connection between cigarette smoking and lung cancer emphasized that “strong evidence” does not imply “crossing every ‘t’, and swords with every critic, before we act.”[1] He proposed differential standards of evidence for different actions. For example, “relatively slight” evidence would be needed to restrict the use of a drug for early-morning sickness in pregnant women; “fair evidence” to substitute a probably carcinogenic oil to a non-carcinogenic oil in an occupational setting; and “very strong evidence” to restrict public smoking and diets. The NAS echoed this approach over four decades later when they pointed to the need to cull those decisions that are not sensitive to the resolution of uncertainty from the risk assessment process, i.e., decisions for which additional information would have little or no value added to support the decision.[18]

Thus, a “more timely” decision pathway is proposed in the *Navigation Guide* to prevent what has been characterized as “paralysis by analysis.”[84] This “off-ramp” would lead to taking action on exposures that are not on an authoritative list via a review of the evidence that falls short of a time-intensive systematic review. The rationale for this is to identify and act on those circumstances where large individual and/or population-wide benefits of preventive action would be squandered by prolonged deliberations that are also not expected to advance decision-making beyond the available evidence. Examples of such circumstances would include situations where: (1) there are important public health implications of inaction; (2) high-risk exposure circumstances, such as those of a pregnant woman with a high exposure to a substance that is not on an authoritative list but for which there is evidence documenting reproductive toxicity; and/or (3) less toxic alternatives are already available.

Key criteria for the off-ramp to timely action would be direct evidence from animal and/or human observational studies of a large potential population impact, and/or, indirect evidence of toxicity from structure-activity relationships (SAR), and/or evidence that the compound influences early (or upstream) biological endpoints that have been linked to more downstream, overt effects. SAR’s are processes whereby a chemical’s structural similarity to other chemicals (for which data are available) is used to determine toxicity. For human health, this process can be used to assess absorption and metabolism, mutagenicity, carcinogenicity, developmental and reproductive effects, neurotoxicity, systemic effects, immunotoxicity, and sensitization and irritation. This is a qualitative assessment using terms such as good, not likely, poor, moderate, or high. The USEPA currently uses both toxicology data and a structure activity in assessment in their evaluation of toxicity.[85] The National Toxicology Program (NTP) *Report on Carcinogens* (RoC) uses “all available science” including if a chemical or substance has a structure or activity comparable to a chemical already listed in the RoC, in which case the NTP finds it is reasonable to assume that the chemical in question would also be a known or reasonably anticipated carcinogen, even if all the data needed to draw that conclusion are still not accessible.[63]^(p.31) The use of SAR is also consistent with clinical sciences in that weight of the evidence decision-making in the regulation of pharmaceuticals employs quantitative structure-activity relationship (QSAR) statistical models to predict the reproductive and developmental toxicity of exogenous agents.[86]

The *Navigation Guide* off-ramp would thus permit lessons learned about the adverse human health impacts of one chemical to be turned into action to prevent harm from chemicals with closely related structure-activity characteristics. For example, a chemical structurally like 1-bromopropane but which is not on an authoritative list would move directly to “take timely action.” This decision would be based on the fact that 1-bromopropane is on the Cal-EPA Prop 65 list, and there is evidence of long-lasting ovarian failure and the absence of sperm in workers exposed to the closely related chemical, 2-bromopropane, and evidence that many chemicals similar to 1-bromopropane, such as DBCP, are known to cause sterility in humans.[87]

Other examples of exposures potentially bypassing a systematic review would be exposures for which there are compelling data and readily available safer alternatives, for example, exposures to pesticides that can be reduced or eliminated by the use of integrated pest management practices or other exposure reduction alternatives.

Less Timely

Substances that are not on an authoritative list and do not meet the criteria for “more timely” action will undergo a systematic review of the available evidence. This assessment will proceed using the GRADE system as a model for evaluating the quality of each of the pieces of evidence.

4. Rate Quality of Each Piece of Evidence

GRADE uses four levels of evidence quality to evaluate each study: high, moderate, low, and very low.[60] Randomized trials begin as high and moderate quality evidence, and observational studies as low quality evidence, i.e., case-control studies and cohort studies, or “very low”, i.e., unsystematic clinical observations, case reports and series. GRADE relies solely on human RCTs and observational study data. GRADE does not incorporate animal data into the evidence stream. However, as noted above, evidence from experimental animal studies and *in vitro* testing are incorporated earlier in the evidence stream related to clinical decisions about exogenous chemicals.

The evidence hierarchy alone does not determine the GRADE system’s rating of the evidence. Study quality may be downgraded as a result of limitations in study design or implementation, imprecision of estimates (wide confidence intervals), variability in results, indirectness of evidence, or publication bias. Publication bias is accounted for because in addition to the many biases related to their design, RCTs are subject to bias due to the dramatic increase in for-profit funding of research;[71] industry funded researchers have been shown to systematically interpret their results differently and in favor of the industry product relative to not-for-profit funding.[88] Quality may be upgraded because of a very large magnitude of effect, a dose-response gradient, and if all plausible biases would reduce an apparent treatment effect. An unbiased observational study or other sources of scientific evidence can be of greater value than a poorly conducted, otherwise biased, unnecessary, or unachievable RCT.[60, 61, 89-92]

To rate the quality of the environmental health evidence stream, workshop participants will have to decide how to apply the GRADE categories --- “high”, “moderate”, “low” and “very low” to human observational studies and *in vivo* and *in vitro* studies. In the clinical sciences *in vivo* and *in vitro* evidence are used to

prevent harmful human exposure to exogenous chemicals (Figure 1). In environmental health sciences animal data are the main information source for prevention. This fact is recognized by every key national and international scientific institution. Animal data on the carcinogenicity of a variety of chemicals have preceded as well as predicted later epidemiological observations in humans and strong evidence exists that experimental results can be extrapolated qualitatively to human subjects.[93]The use of animal data is fundamental to timely prevention of adverse health outcomes. Whereas an experimental animal carcinogenic study typically lasts two years, it can take 20 years to get a result from a comparable human study.[93]

After determining the initial standing in the evidence hierarchy, the quality of individual human observational studies would be rated in the same manner as used by GRADE. The quality of individual experimental animal studies would be rated according to existing rigorous criteria employed by IARC and USEPA (Figure 8).

Figure 8: Criteria for Rating Quality of Individual Studies

Type of Study	Criteria for Rating Quality of Study	Source of Criteria
Animal/Toxicology Studies	<ul style="list-style-type: none"> IARC: Evaluation of the strength of the evidence for carcinogenicity from animal data is evaluated using conventional bioassays, bioassays with genetically modified animals and in-vivo bioassays. IARC uses the guidelines set out by the Organization for Economic Cooperation and Development (OECD) on the evaluation of chronic toxicity and carcinogenicity studies to evaluate the validity of animal studies (2002). US EPA: The US EPA considers a variety of experimental animal studies and relies heavily on long-term whole animal carcinogenicity studies when determining carcinogenic risk to humans. The US EPA has developed a health effects test guideline for carcinogenicity (OPPTS 870.4300) which is based on previous EPA test guidelines and the OECD 453 combined guideline for chronic toxicity and carcinogenicity studies. These guidelines provide the basis for evaluating experimental studies for carcinogenicity. In addition to its own guidelines (OPPTS 870.4300) and the OECDs, the EPA relies on other published guidelines (NTP 1984, OSTP 1985, Chhabra 1990). All available studies of tumor effects in whole animals are considered by the EPA. However the US EPA may discard studies they judge have inadequate protocols, conduct or results. 	<p>IARC: Studies of cancer in experimental animals http://monographs.iarc.fr/ENG/Preamble/currentb3studiesanimals0706.php</p> <p>IARC: Scientific Review and Evaluation http://monographs.iarc.fr/ENG/Preamble/currentb6evaluation0706.php</p> <p>OECD, 2002: Guidance Notes for Analysis and Evaluation of Chronic Toxicity and Carcinogenicity Studies http://www.oalis.oecd.org/olis/2002doc.nsf/LinkTo/NT00002BE2/\$FILE/JT00130828.PDF</p> <p>USEPA Cancer Guidelines, 2005, Section 2-15, page 42 http://oaspub.epa.gov/eims/eimscomm.getfile?p_download_id=439797</p>
Human Studies	<p>Agency for Healthcare Research and Quality concludes that systems that evaluate the following 5 domains represent acceptable approaches for assessing the quality of observational studies.</p> <ol style="list-style-type: none"> 1. Comparability of subjects 2. Exposure or intervention 3. Outcome measurement 4. Statistical analysis 5. Funding or sponsorship 	<p>AHRQ/ U.S. Department of Health and Human Services. Systems to Rate the Strength of Scientific Evidence. Evidence Report/Technology Assessment, Number 47.</p>
Wildlife Studies	TBD	

5. Determine Strength of Evidence

Next, the quality of each of the individually rated studies is aggregated into a summary statement about the overall quality of the evidence. The resulting “strength of the evidence” categories, i.e., known to be toxic, probably toxic, possibly toxic, not classifiable as toxic, and probably not toxic, are described in Figure 9 and in further detail, in Appendix 1. These criteria were developed based on the IARC/USEPA cancer criteria were made into a “generic” form for reproductive and developmental toxics.

Figure 9: Classifications for Levels of Reproductive Toxicity*

Category	Required Levels of Data	Description
Toxic	<ul style="list-style-type: none"> Sufficient (H), OR Strong (H) AND Sufficient (A) AND Relevant Mechanism (exceptionally) 	This category is used when there is sufficient evidence of toxicity in humans. Exceptionally, an agent (mixture) may be placed in this category when evidence of toxicity in humans is less than sufficient but there is sufficient evidence of toxicity in experimental animals and strong evidence in exposed humans that the agent (mixture) acts through a relevant mechanism of toxicity.
Probably Toxic	<ul style="list-style-type: none"> Limited (H) AND Sufficient (A), OR Inadequate (H) AND Sufficient (A) AND Relevant Mechanism (in some cases), OR Limited (H) (exceptionally) 	This category is used when there is limited evidence of toxicity in humans and sufficient evidence of toxicity in experimental animals. In some cases, an agent (mixture) may be classified in this category when there is inadequate evidence of toxicity in humans and sufficient evidence of toxicity in experimental animals and strong evidence that the toxicity is mediated by a mechanism that also operates in humans. Exceptionally, an agent, mixture or exposure circumstance may be classified in this category solely on the basis of limited evidence of toxicity in humans.
Possibly Toxic	<ul style="list-style-type: none"> Limited (H) AND <Sufficient (A), OR Inadequate (H) AND Sufficient (A), OR Inadequate (H) AND Limited (A) AND Other Supporting Evidence (in some cases) 	This category is used for agents, mixtures and exposure circumstances for which there is limited evidence of toxicity in humans and less than sufficient evidence of toxicity in experimental animals. It may also be used when there is inadequate evidence of toxicity in humans but there is sufficient evidence of toxicity in some experimental animals. In some instances, an agent, mixture or exposure circumstance for which there is inadequate evidence of toxicity in humans but limited evidence of toxicity in experimental animals together with supporting evidence from other relevant data may be placed in this group.
Not Classifiable as to Toxicity	<ul style="list-style-type: none"> Inadequate (H) AND Inadequate (A) OR Limited (A), OR Inadequate (H) AND Sufficient (A) AND Irrelevant Mechanism (exceptionally), OR Does not fit into criteria for groups 1, 2A/B, or 4 	This category is used most commonly for agents, mixtures and exposure circumstances for which the evidence of toxicity is inadequate in humans and inadequate or limited in experimental animals. Exceptionally, agents (mixtures) for which the evidence of toxicity is inadequate in humans but sufficient in experimental animals may be placed in this category when there is strong evidence that the mechanism of toxicity in experimental animals does not operate in humans. Agents, mixtures and exposure circumstances that do not fall into any other group are also placed in this category.
Probably Not Toxic	<ul style="list-style-type: none"> Evidence Suggesting Lack of Toxicity (H and A), OR Inadequate (H) AND Evidence Suggesting Lack of Toxicity (A) AND Other Supporting Evidence 	This category is used for agents or mixtures for which there is evidence suggesting lack of toxicity in humans and in experimental animals. In some instances, agents or mixtures for which there is inadequate evidence of toxicity in humans but evidence suggesting lack of toxicity in experimental animals, consistently and strongly supported by a broad range of other relevant data, may be classified in this group. (Revised IARC Group 4)

(H) Human observational data, (A) Animal data

* See Appendix 1 for further explanation.

6. No Action At This Time

The decision for substances for which there is evidence of a lack of toxicity (Figure 9 and Appendix 1) is to take no action at this time.

7. Take Action to Prevent Exposure

Current approaches to limiting population exposure to environmental hazards are generally undertaken long after the exposure has occurred. For example, evidence-based programs of exposure reduction related to lead have been highly successful but were initiated after many decades of delay.[94] Likewise, steps to prevent exposure to ionizing radiation and asbestos and to prevent further destruction of the ozone layer from chlorofluorocarbons (CFCs) were undertaken long after the first credible scientific evidence of harm emerged.[84, 95] It can take 10 to 20 years of efforts to complete

quantitative assessments of risk for well-recognized environmental hazards. The decades-long process reviewing the reproductive toxicity of dioxin is illustrative of such prolonged deliberations. A 2008 report by the National Academy of Sciences (NAS) emphasized the need to identify and protect against chemicals that can harm human health before entering into a long process to establish numerical levels of risk.[18]

Regulatory deficiencies also contribute to delayed response to preventing exposure to toxic substances. In the U.S., the Toxic Substances Control Act of 1976 (TSCA) provides the USEPA with authority to take regulatory action on hazardous chemical substances and/or mixtures both before and after they enter commerce, though with differing evidence requirements. However, at least five government studies conducted between 1984 and 2005 have all concluded that TSCA has not served as an effective vehicle for the public, industry, or government to assess the hazards of chemicals in commerce or control those of greatest concern.[64] For example, in the 33 years since it was granted this authority, USEPA has banned only five substances.^e

The *Navigation Guide* proposes timely action be taken to prevent or reduce exposure to substances with evidence of toxicity. The rationale for this is best public health practice that prioritizes prevention over treatment. The *Navigation Guide* defines evidence of toxicity in two ways (1) substances present on an authoritative list (as defined above); and (2) substances judged through a transparent review of the evidence as: known to be toxic, probably toxic, and possibly toxic (as described above, based on IARC/USEPA criteria for judging the carcinogenicity of a substance, Figure 9 and Appendix 1).

The *Navigation Guide* also proposes that action be taken to prevent or reduce exposure to some substances “not classifiable as toxic.” These are substances that lack sufficient data to classify as having known, probable or possible toxicity, but which also do not have evidence showing a lack of toxicity. This category is used by IARC/USEPA most commonly for agents, mixtures and exposure circumstances for which the evidence of toxicity is inadequate in humans and inadequate or limited in experimental animals (Figure 9).

The rationale for this is the “precautionary principle.” The idea at the core of the precautionary principle is that action should be taken to prevent harm to the environment and human health, even if scientific evidence is inconclusive. The concept of precaution was born in the environmental domain in the 1970s, and by the 1980s precaution began to be invoked in international environmental agreements in the context of mounting evidence of unprecedented environmental changes surrounded by vast uncertainties.[66] In 1998, the precautionary principle became incorporated as a framing concept in environmental health in the U.S. after participants at the Wingspread conference in Racine, Wisconsin, issued the statement, “When an activity raises threats of harm to human health or the environment, precautionary measures should be taken even if some cause and effect relationships are not fully established scientifically” (Appendix 2).[96] Currently, European law explicitly incorporates the principles of prevention, precaution and proportionality (which prevents the unreasonable use of

^e Polychlorinated biphenols (PCBs), halogenated chlorofluoroalkanes (CFCs), dioxins, hexavalent chromium, and asbestos. The 1989 ban on asbestos was overturned in court in 1991. Source: Shapiro M. Exposed. The Toxic Chemistry of Everyday Products and What's at Stake for American Power. Chelsea Green. White River Junction, VT: 2007

precaution) into environmental health decision-making.[97]

It is critical to keep in mind that the nature and extent of “taking action” will vary between these categories. These differences in the weight of evidence among these categories are incorporated later in the *Navigation Guide*, when the strength of the recommendation for action is graded.

8. Grade Strength of Recommendations

In GRADE and other related taxonomies, the quality of evidence is separated from strength of recommendations in recognition of the role that patient values and preferences as well as clinical and social circumstances play in formulating practice recommendations (Figure 2).[60] For example, evidence-based maternity care proceeds from a framework that informed decision-making should incorporate values and circumstances of individual women.[98] Notably, while it is widely accepted that patient values and preferences are a critical component of decision making in the clinical context,[99] there is currently no consensus on why and how to involve patients in clinical practice guideline development. A recent study from the U.K. documented a diversity of views on patient involvement, and highlighted a central division between perspectives that seek to adopt strategies to maximize benefits for a total population and those that seek to promote individual patient’ interests.[100]

The clinical framework that combines an objective measure of the quality of evidence with an understanding of the context of the decision has its counterpart in environmental health practice. Recent recommendations for improving risk assessment made by the NAS speak to formal and informal provisions for stakeholder involvement at all stages of the process, while the technical assessment of risk (rating of evidence) is carried out under its own standards and guidelines.[18] It has been observed that the strength of the National Toxicology Program is “a direct result of the amount of public input factored into the decision-making process.”[63]^(p.31) Grading the strength of the evidence is comparable to selecting the weight of the evidence needed to inform any risk management decision.

Key determinants of the strength of the recommendations used in the *Navigation Guide* would be: best public health practices, such as the hierarchy of exposure control measures in industrial hygiene that places highest value on controlling hazardous exposures by eliminating them and using safer alternatives to accomplish the task at hand; circumstances of exposure, such as the route, duration, timing and presence of individual vulnerabilities; quality of evidence; costs (resource allocation); and values and preferences. The balance between desirable and undesirable effects of a specific recommendation is also a key determinant, and this factor is anticipated to be especially relevant to occupational exposure scenarios, where the insufficiencies of current regulations are likely to produce decisions that pit the benefits of employment against reproductive health harms that may be incurred by workplace exposures.^f

^f Occupational health regulations generally do not reflect well-documented chronic health impacts. A 2007 study by the California Environmental Protection Agency (Cal-EPA) found that 44 of 106 workplace chemicals known to the state of California to cause cancer do not have a permissible exposure limit and 62 are not regulated as carcinogens. The report found that 5 of 19 workplace chemicals known to cause reproductive or developmental harm do not have a permissible exposure limit and 14 are not regulated as reproductive or developmental hazards. Risks that are considered acceptable for workers are much greater

All of these key factors would be incorporated into “what is known” and “what can be done” decision algorithm (Figure 10) by PRHE scientists and/or others who utilize the *Navigation Guide* to vet the scientific evidence. Under this algorithm, strong recommendations would apply to exposures about which much is known in terms of toxicity, exposure patterns, vulnerabilities, and for which much can be done to prevent or reduce exposure. Such an example would be a strong recommendation to prevent preconception, prenatal, childhood and occupational lead exposure.

Figure 10: Grading the Strength of a Recommendation

		What is known	
		Little	Much
What can be done	Little	<p>Discretionary: Little is known about the chemical toxicity and there are limited options for what can be done</p>	<p>Discretionary: Much is known about the chemical toxicity but there are limited options for what can be done</p>
	Much	<p>Strong: Limited understanding of a chemical's toxicity but safer alternatives are readily available</p>	<p>Strong: Clear understanding of a chemical's toxicity and safer alternatives are readily available</p>

In the clinical arena, GRADE permits strong recommendations to be derived from relatively weak evidence if the availability of an alternative can result in a clear decision. For example, the early case-control studies demonstrating the association between aspirin use and Reye's syndrome were relatively weak and left considerable doubt about the causal relationship. However the availability of a safe, inexpensive, and well-tolerated alternative, acetaminophen, justified use of this alternative agent in children at risk of Reye's syndrome.[101, 102] The *Navigation Guide* would also permit strong recommendations in cases where there is little data on chronic toxicity but when large populations of vulnerable individuals would be exposed and an alternative exists to accomplish the task and the efficacy of that alternative is known. Such an example

than risk levels established for the public, i.e., six workplace chemicals identified in the Cal-EPA report had risks of greater than 1 case of cancer for every 10 exposed workers. See: California Environmental Protection Agency. Office of Environmental Health Hazard Assessment, Reproductive and Cancer Hazard Branch. Occupational Health Hazard Risk Assessment Project for California: Identification of Chemicals of Concern, Possible Risk Assessment Methods, and Examples of Health Protective Occupational Air Concentrations. December 2007. <http://www.dhs.ca.gov/ohb/HESIS/riskreport.pdf>

would be a strong recommendation to prevent exposure to widespread aerial spraying of pesticides when a non-toxic method of preventing the spread of unwanted pests (i.e., sterilization of gypsy moths) is available.

A discretionary recommendation is warranted for exposures where much is known about toxicity and the efficacious and beneficial impacts of the alternative, but most individuals can do little to adopt a specific alternative due to high cost or lack of availability in the marketplace. An example here might be purchasing organic food as an alternative to incurring pesticide exposure via the food supply. This example also points to how the process of grading recommendations has the potential to identify policy gaps that need to be addressed in order for individuals to take action to prevent hazardous exposures.

A discretionary recommendation would also result from circumstances where little is known about a particular substance's toxicity, exposure is not large for any individual or widespread throughout the population, and an informed individual finds value (such as in accomplishing a necessary task or employment) in incurring the exposure.

Notably, the *Navigation Guide's* adoption of IARC/USEPA cancer criteria to evaluate the weight of the evidence for reproductive hazard assessment implies a "no threshold" dose-response relationship will inform the *Navigation Guide's* strength of the recommendations. A "no-threshold" assumption translates into a "no safe dose" recommendation for substances that are reproductive or developmental toxicants. The proposal to utilize a "no-threshold" assumption in the *Navigation Guide* is consistent with the direction of risk assessment outlined in a 2008 NAS report.[18] The NAS report described important structural problems with the current framework that assumes thresholds exist for non-carcinogens. Under the current framework, risk assessments of carcinogens have assumed that there is no threshold of effect, meaning there is some risk of cancer even at the lowest doses. Risk assessments for non-cancer health outcomes such as reproductive toxicity have assumed there is some threshold level of exposure below which effects do not occur or are extremely unlikely. When non-carcinogen dose-response relationships with no apparent thresholds have been observed, such as for subtle and common adverse endpoints, like IQ loss or neurobehavioral deficits associated with lead or methylmercury exposures, they have been treated as "exceptions" to the threshold rule, not the norm.

The NAS report raises many scientific and policy concerns about the current assumption of a threshold for non-carcinogenic health endpoints. A central concern of the NAS scientists is that other concurrent chemical exposures and biologic factors that influence the same adverse effect can modify the dose-response relationship at low doses and should therefore be considered. This is especially the case when an underlying disease can interact with the toxicant, such as the interaction of cardiopulmonary disease and exposure to particulate matter or ozone. Case studies of adverse upstream endpoints, (i.e., thyroid hormone disruption and related toxicities, antiandrogen-mediated male and reproductive effects, and immune function) illustrate the importance of considering preexisting exposures or continuous exposures to environmental chemicals as well as preexisting biological or disease susceptibilities that contribute independently to risk of overt disease.[103] The NAS report recommends that a new framework be devised that addresses these and other limitations.

The *Navigation Guide's* strength of the recommendations will also be congruent with the direction of risk assessment for early life exposure to carcinogens being pursued by the

Cal-EPA Office of Environmental Health Hazard Assessment (OEHHA).[104] Health risk varies depending on whether exposure to carcinogens occurs from conception through puberty or is incurred in adulthood. In recognition of this increased period of vulnerability, OEHHA is incorporating age-at-exposure adjustments into its quantitative cancer risk assessments. USEPA has indicated that it will be considering modifications to stay in line with OEHHA advances.

9. Craft Effective Anticipatory Guidelines and Patient Messages

As with GRADE, the *Navigation Guide* would result in evidence profiles that provide simple, transparent summaries, such as practice guidelines. Our long-term aspirations are to develop a web-based application of the *Navigation Guide* whereby the details would be nested electronically behind a clear and simple message. Depending on their interest level, clinicians and other end users can simply access the “bottom line” and/or can see every step in the decision-logic.

10. Clinician Provides Anticipatory Guidelines

This is the area of clinical application of the *Navigation Guide*. The *Navigation Guide* practice guidelines will be used to support a wide range of clinical, patient and public educational activities and policy recommendations. The commonality of these potentially wide-ranging efforts will be a timely, comprehensive, and transparent evidence-based foundation. The actual materials can be crafted to meet the language, literacy and cultural needs of diverse populations. Based on the experience of EBM, barriers to the uptake of the anticipatory guidance by clinicians will need to be addressed. For example, many factors shape views about maternity-care, and what is considered suitable care and patterns of practice often do not reflect the best current research.[98]

Real World Feedback (Evaluation)

Our goal is to incorporate multiple mechanisms of timely evaluation including a web-based portal to evaluate application of the *Navigation Guide* in “real-world” circumstances in on-going and timely manner.

It is anticipated that application of the *Navigation Guide* will point to the absence, strengths and weaknesses of data for comparing relative effectiveness of safer alternatives. In the clinical arena, for a variety of reasons, there is a lack of any evidence on the efficacy or effectiveness of many interventions. It is difficult to extrapolate from trials carried out on selected populations to those with multiple chronic conditions, and there is increased recognition that the benefits of interventions vary according to the underlying risk of the population.[105] A 2009 study of clinical guidelines of the American College of Cardiology and the American Heart Association documented that less than one in eight (314 of 2711) recommendations were based on evidence from multiple randomized trials or meta-analyses (level A).[106]

V. Future Directions

The *Navigation Guide* is not a panacea, but a missing tool in a much larger effort to address the public health impacts of widespread environmental exposure to toxic substances. The developers of GRADE have observed that using an approach that is systematic and transparent reduces (but does not eliminate) the likelihood of making judgments that cannot be substantiated. However it does allow others to inspect the basis for the judgments, and facilitates identifying the reasons for disagreements.

An EBM framework carries many implications for policy and law.[107] A recent commentary in *JAMA* noted that only when likely biases of industry and specialty societies have been either removed or overcome by countervailing interests can the promise of impartial recommendations be achieved.[108] The *Navigation Guide* will likewise have to account for the influence of non-impartial data in the evidence stream and an evidence-based approach to environmental health will require re-invigorated efforts to develop data streams free of commercial bias.

The evidence stream is rapidly changing in both clinical and environmental health sciences and the *Navigation Guide* and other evidence-based systems will need constant review to ensure the most current approaches to discerning the evidence are rapidly incorporated and evaluated. It is anticipated that EBM will increasingly rely on nonrandomized evidence. The speed and complexity with which new medical interventions and scientific knowledge are being created make it unlikely that the evidence base required for treatment and cost effective health care delivery across subpopulations can be built using only RCTs.[105] It is also expected that electronic medical records will revolutionize medical research by facilitating instant, comprehensive, longitudinal data that go back years into history and extend indefinitely into the future.[109] Harnessing these changes could greatly accelerate the creation of knowledge about the impact of the environment on human health. Advances in toxicity testing,[110, 111] risk assessment,[18, 19, 52, 112] and policy⁹[64] are likely to create a sea change in the how environmental chemicals are assessed and regulated in the future. The proposed *Navigation Guide* offers a framework to incorporate these and other innovations rapidly and transparently as they unfold.

⁹ REACH is a new European Community Regulation on chemicals and their safe use (EC 1907/2006). It deals with the Registration, Evaluation, Authorisation and Restriction of Chemical substances. The new law entered into force on 1 June 2007. http://ec.europa.eu/environment/chemicals/reach/reach_intro.htm

Appendix 1. Detailed Description of Classification of Levels of Toxicity

Category	Description
Toxic to Humans	The agent (mixture) is toxic to humans. The exposure circumstance entails exposures that are toxic to humans. This category is used when there is sufficient evidence of toxicity in humans[1]. Exceptionally, an agent (mixture) may be placed in this category when evidence of toxicity in humans is less than sufficient but there is sufficient evidence of toxicity in experimental animals[2] and strong evidence in exposed humans that the agent (mixture) acts through a relevant mechanism of toxicity.
Probably Toxic To Humans	The agent (mixture) is probably toxic to humans. The exposure circumstance entails exposures that are probably toxic to humans. This category is used when there is limited evidence of toxicity in humans[3] and sufficient evidence of toxicity in experimental animals[2]. In some cases, an agent (mixture) may be classified in this category when there is inadequate evidence of toxicity in humans[4] and sufficient evidence of toxicity in experimental animals and strong evidence that the toxicity is mediated by a mechanism that also operates in humans. Exceptionally, an agent, mixture or exposure circumstance may be classified in this category solely on the basis of limited evidence of toxicity in humans[3].
Possibly Toxic	The agent (mixture) is possibly toxic to humans. The exposure circumstance entails exposures that are possibly toxic to humans. This category is used for agents, mixtures and exposure circumstances for which there is limited evidence of toxicity in humans[3] and less than sufficient evidence of toxicity in experimental animals[2]. It may also be used when there is inadequate evidence of toxicity in humans[4] but there is sufficient evidence of toxicity in experimental animals. In some instances, an agent, mixture or exposure circumstance for which there is inadequate evidence of toxicity in humans but limited evidence of toxicity in experimental animals[5] together with supporting evidence from other relevant data[6] may be placed in this group.
Not Classifiable As To Toxicity	The agent (mixture or exposure circumstance) is not classifiable as to its toxicity to humans. This category is used most commonly for agents, mixtures and exposure circumstances for which the evidence of toxicity is inadequate in humans[7] and inadequate[8] or limited[5] in experimental animals. Exceptionally, agents (mixtures) for which the evidence of toxicity is inadequate in humans but sufficient in experimental animals[2] may be placed in this category when there is strong evidence that the mechanism of toxicity in experimental animals does not operate in humans. Agents, mixtures and exposure circumstances that do not fall into any other group are also placed in this category.
Probably Not Toxic	The agent (mixture) is probably not toxic to humans. This category is used for agents or mixtures for which there is evidence suggesting lack of toxicity in humans[9] and in experimental animals[10]. In some instances, agents or mixtures for which there is inadequate evidence of toxicity in humans[7] but evidence suggesting lack of toxicity in experimental animals, consistently and strongly supported by a broad range of other relevant data[6], may be classified in this group.

Notes

1. Sufficient evidence of toxicity in humans: A causal relationship has been established between exposure to the agent, mixture or exposure circumstance and human toxicity. That is, a positive relationship has been observed between the exposure and toxicity in studies in which chance, bias and confounding could be ruled out with reasonable confidence.
2. Sufficient evidence of toxicity in experimental animals: A causal relationship has been established between the agent or mixture and an increased incidence of the adverse health outcome in (a) two or more species of animals or (b) in two or more independent studies in one species carried out at different times or in different laboratories or under different protocols. Exceptionally, a single study in one species might be considered to provide sufficient evidence of toxicity when the adverse outcomes occur to an unusual degree with regard to incidence, site, type of outcome or age at onset.
3. Limited evidence of toxicity in humans: A positive association has been observed between exposure to the agent, mixture or exposure circumstance and toxicity for which a causal interpretation is considered by the Working Group to be credible, but chance, bias or confounding could not be ruled out with reasonable confidence.
4. Inadequate evidence of toxicity in humans: The available studies are of insufficient quality, consistency or statistical power to permit a conclusion regarding the presence or absence of a causal association between exposure and toxicity, or no data on toxicity in humans are available.
5. Limited evidence of toxicity in experimental animals: The data suggest a toxic effect but are limited for making a definitive evaluation because, e.g. (a) the evidence of toxicity is restricted to a single experiment; or (b) there are unresolved questions regarding the adequacy of the design, conduct or interpretation of the study; or (c) the agent or mixture increases the incidence of health outcomes which may occur spontaneously in high incidences in certain strains.
6. Other data relevant to the evaluation of toxicity and its mechanisms: Other evidence judged to be relevant to an evaluation of toxicity and of sufficient importance to affect the overall evaluation is then described. This may include data on preneoplastic lesions, tumour pathology, genetic and related effects, structure-activity relationships, metabolism and pharmacokinetics, physicochemical parameters and analogous biological agents. Data relevant to mechanisms of the toxic action are also evaluated. The strength of the evidence that any toxic effect observed is due to a particular mechanism is assessed, using terms such as weak, moderate or strong. Then, the Working Group assesses if that particular mechanism is likely to be operative in humans. The strongest indications that a particular mechanism operates in humans come from data on humans or biological specimens obtained from exposed humans. The data may be considered to be especially relevant if they show that the agent in question has caused changes in exposed humans that are on the causal pathway to toxicity. Such data may, however, never become available, because it is at least conceivable that certain compounds may be kept from human use solely on the basis of evidence of their toxicity in experimental systems. For complex exposures, including occupational and industrial exposures, the chemical composition and the potential contribution of toxicants known to be present are considered by the Working Group in its overall evaluation of human toxicity. The Working Group also determines the extent to which the materials tested in experimental systems are related to those to which humans are exposed.
7. Inadequate evidence of toxicity in humans: The available studies are of insufficient quality, consistency or statistical power to permit a conclusion regarding the presence or absence of a causal association between exposure and toxicity, or no data on toxicity in humans are available.
8. Inadequate evidence of toxicity in experimental animals: The studies cannot be interpreted as showing either the presence or absence of a toxic effect because of major qualitative or quantitative limitations, or no data on toxicity in experimental animals are available.
9. Evidence suggesting lack of toxicity in humans: There are several adequate studies covering the full range of levels of exposure that human beings are known to encounter, which are mutually consistent in not showing a positive association between exposure to the agent, mixture or exposure circumstance and any studied toxicity at any observed level of exposure. A conclusion of 'evidence suggesting lack of toxicity' is inevitably limited to the cancer sites, conditions and levels of exposure and length of observation covered by the available studies. In addition, the possibility of a very small risk at the levels of exposure studied can never be excluded. In some instances, the above categories may be used to classify the degree of evidence related to toxicity in specific organs or tissues.
10. Evidence suggesting lack of toxicity in experimental animals: Adequate studies involving at least two species are available which show that, within the limits of the tests used, the agent or mixture is not toxic. A conclusion of evidence suggesting lack of toxicity is inevitably limited to the species, tumour sites and levels of exposure studied.

Appendix 2. The Wingspread Statement on the Precautionary Principle (January 1998)

The release and use of toxic substances, the exploitation of resources, and physical alterations of the environment have had substantial unintended consequences affecting human health and the environment. Some of these concerns are high rates of learning deficiencies, asthma, cancer, birth defects and species extinctions; along with global climate change, stratospheric ozone depletion and worldwide contamination with toxic substances and nuclear materials.

We believe existing environmental regulations and other decisions, particularly those based on risk assessment, have failed to protect adequately human health and the environment - the larger system of which humans are but a part.

We believe there is compelling evidence that damage to humans and the worldwide environment is of such magnitude and seriousness that new principles for conducting human activities are necessary. While we realize that human activities may involve hazards, people must proceed more carefully than has been the case in recent history. Corporations, government entities, organizations, communities, scientists and other individuals must adopt a precautionary approach to all human endeavors.

Therefore, it is necessary to implement the Precautionary Principle: When an activity raises threats of harm to human health or the environment, precautionary measures should be taken even if some cause and effect relationships are not fully established scientifically.

In this context the proponent of an activity, rather than the public, should bear the burden of proof. The process of applying the Precautionary Principle must be open, informed and democratic and must include potentially affected parties. It must also involve an examination of the full range of alternatives, including no action.

Participants

- Dr. Nicholas Ashford, M.I.T.
- Katherine Barrett, Univ. of British Columbia
- Anita Bernstein, Chicago-Kent College of Law
- Dr. Robert Costanza, Univ. of Maryland
- Pat Costner, Greenpeace
- Dr. Carl Cranor, Univ. of California, Riverside
- Dr. Peter deFur, Virginia Commonwealth Univ.
- Gordon Durnil, Attorney
- Dr. Kenneth Geiser, Toxics Use Reduction Inst., Univ. of Mass., Lowell
- Dr. Andrew Jordan, Centre for Social and Economic Research on the Global Environment, Univ. of East Anglia
- Andrew King, United Steelworkers of America, Canadian Office
- Dr. Frederick Kirschenmann, Farmer
- Stephen Lester, Center for Health, Environment and Justice
- Sue Maret, Union Inst.
- Dr. Michael M'Gonigle, Univ. of Victoria, British Columbia
- Dr. Peter Montague, Environmental Research Foundation
- Dr. John Peterson Myers, W. Alton Jones Foundation
- Dr. Mary O'Brien, Environmental Consultant
- Dr. David Ozonoff, Boston Univ.
- Carolyn Raffensperger, Science and Environmental Health Network
- Dr. Philip Regal, Univ. of Minnesota
- Hon. Pamela Resor, Massachusetts House of Representatives
- Florence Robinson, Louisiana Environmental Network
- Dr. Ted Schettler, Physicians for Social Responsibility
- Ted Smith, Silicon Valley Toxics Coalition
- Dr. Klaus-Richard Sperling, Alfred-Wegener- Institut, Hamburg
- Dr. Sandra Steingraber, Author
- Diane Takvorian, Environmental Health Coalition
- Joel Tickner, Univ. of Mass., Lowell
- Dr. Konrad von Moltke, Dartmouth College
- Dr. Bo Wahlstrom, KEMI (National Chemical Inspectorate), Sweden
- Jackie Warledo, Indigenous Environmental Network

Literature Cited

1. Hill, A.B., *The Environment and Disease: Association or Causation?* Proc R Soc Med, 1965. 58: p. 295-300.
2. Schuster, M.A., N. Duan, M. Regalado, and D.J. Klein, *Anticipatory guidance: what information do parents receive? What information do they want?* Arch Pediatr Adolesc Med, 2000. 154(12): p. 1191-8.
3. Cunningham, M., *Chronic occupational lead exposure: the potential effect on sexual function and reproductive ability in male workers.* Aaohn J, 1986. 34(6): p. 277-9.
4. Hauser, R. and R. Sokol, *Science Linking Environmental Contaminant Exposures with Fertility and Reproductive Health Impacts in the Adult Male.* Fertil Steril, 2008. 89(2 Suppl): p. e59-65.
5. Rusyniak, D.E., R.B. Furbee, and R. Pascuzzi, *Historical neurotoxins: what we have learned from toxins of the past about diseases of the present.* Neurol Clin, 2005. 23(2): p. 337-52.
6. McBride, W.G., *Thalidomide and Congenital Abnormalities.* The Lancet, 1961. 278(7216): p. 1358-1358.
7. William, G.M., *Thalidomide embryopathy.* Teratology, 1977. 16(1): p. 79-82.
8. Newbold, R.R., *Lessons learned from perinatal exposure to diethylstilbestrol.* Toxicol Appl Pharmacol, 2004. 199(2): p. 142-50.
9. Colborn, T., D. Dumanoski, and J.P. Myers, *Our Stolen Future.* 1996: Penguin Books USA, Inc. 1-306.
10. Crain, D.A., S.J. Janssen, T.M. Edwards, J. Heindel, S.M. Ho, P. Hunt, T. Iguchi, A. Juul, J.A. McLachlan, J. Schwartz, N. Skakkebaek, A.M. Soto, S. Swan, C. Walker, T.K. Woodruff, T.J. Woodruff, L.C. Giudice, and L.J. Guillette, Jr., *Female reproductive disorders: the roles of endocrine-disrupting compounds and developmental timing.* Fertil Steril, 2008. 90(4): p. 911-40.
11. Schettler, T., G. Solomon, M. Valenti, and A. Huddle, *Generations at Risk. Reproductive Health and the Environment.* 1999, Cambridge, MA: MIT Press.
12. Woodruff, T.J., A. Carlson, J.M. Schwartz, and L.C. Giudice, *Proceedings of the Summit on Environmental Challenges to Reproductive Health and Fertility: Executive summary.* Fertil Steril, 2008. 89(2 Suppl): p. e1-e20.
13. Oehlmann, J., U. Schulte-Oehlmann, W. Kloas, O. Jagnytsch, I. Lutz, K.O. Kusk, L. Wollenberger, E.M. Santos, G.C. Paull, K.J. Van Look, and C.R. Tyler, *A critical analysis of the biological impacts of plasticizers on wildlife.* Philos Trans R Soc Lond B Biol Sci, 2009. 364(1526): p. 2047-62.
14. Guillette, L.J., Jr. and T.M. Edwards, *Environmental Influences on Fertility: Can We Learn Lessons from Studies of Wildlife?* Fertil Steril, 2008.
15. U.S. EPA. *What is the TSCA Chemical Substance Inventory.* 2006 [cited 2007 Apr 4]; Available from: <http://www.epa.gov/opptintr/newchems/pubs/invntory.htm>.
16. National Research Council, *Toxicity Testing in the 21st Century: A Vision and Strategy.* 2007, Washington, DC: The National Academies Press.
17. CDC, *Third national report on human exposure to environmental chemicals.* 2005, Centers for Disease Control and Prevention: Atlanta, GA.
18. National Academy of Science, *Science and Decisions: Advancing Risk Assessment.* 2008, National Research Council Committee on Improving Risk Analysis Approaches Used by the U.S. EPA: Washington, DC.

19. National Research Council, *Phthalates and Cumulative Risk Assessment: The Task Ahead*, ed. Committee on the Health Risks of Phthalates. 2008, Washington, D.C.: National Academies Press.
20. Welshons, W.V., K.A. Thayer, B.M. Judy, J.A. Taylor, E.M. Curran, and F.S. vom Saal, *Large effects from small exposures. I. Mechanisms for endocrine-disrupting chemicals with estrogenic activity*. Environ Health Perspect, 2003. 111(8): p. 994-1006.
21. Palanza, P., F. Morellini, S. Parmigiani, and F.S. vom Saal, *Prenatal exposure to endocrine disrupting chemicals: effects on behavioral development*. Neurosci Biobehav Rev, 1999. 23(7): p. 1011-27.
22. Talsness, C.E., A.J. Andrade, S.N. Kuriyama, J.A. Taylor, and F.S. vom Saal, *Components of plastic: experimental studies in animals and relevance for human health*. Philos Trans R Soc Lond B Biol Sci, 2009. 364(1526): p. 2079-96.
23. Chevrier, J., B. Eskenazi, N. Holland, A. Bradman, and D.B. Barr, *Effects of exposure to polychlorinated biphenyls and organochlorine pesticides on thyroid function during pregnancy*. Am J Epidemiol, 2008. 168(3): p. 298-310.
24. Maervoet, J., G. Vermeir, A. Covaci, N. Van Larebeke, G. Koppen, G. Schoeters, V. Nelen, W. Baeyens, P. Schepens, and M.K. Viaene, *Association of thyroid hormone concentrations with levels of organochlorine compounds in cord blood of neonates*. Environ Health Perspect, 2007. 115(12): p. 1780-6.
25. Vandenberg, L.N., R. Hauser, M. Marcus, N. Olea, and W.V. Welshons, *Human exposure to bisphenol A (BPA)*. Reprod Toxicol, 2007. 24(2): p. 139-77.
26. Stahlhut, R.W., W.V. Welshons, and S. Swan, *Bisphenol A Data in NHANES Suggest Longer Than Expected Half-Life, Substantial Non-Food Exposure, or Both*. Environ Health Perspect, 2009 In-Press.
27. vom Saal, F.S. and C. Hughes, *An extensive new literature concerning low-dose effects of bisphenol A shows the need for a new risk assessment*. Environ Health Perspect, 2005. 113(8): p. 926-33.
28. Welshons, W.V., S.C. Nagel, and F.S. vom Saal, *Large effects from small exposures. III. Endocrine mechanisms mediating effects of bisphenol A at levels of human exposure*. Endocrinology, 2006. 147(6 Suppl): p. S56-69.
29. Hines, C.J., N.B. Nilsen Hopf, J.A. Deddens, A.M. Calafat, M.J. Silva, A.A. Grote, and D.L. Sammons, *Urinary phthalate metabolite concentrations among workers in selected industries: a pilot biomonitoring study*. Ann Occup Hyg, 2009. 53(1): p. 1-17.
30. Riederer, A.M., S.M. Bartell, D.B. Barr, and P.B. Ryan, *Diet and nondiet predictors of urinary 3-phenoxybenzoic acid in NHANES 1999-2002*. Environ Health Perspect, 2008. 116(8): p. 1015-22.
31. Kumar, R., N. Pant, and S.P. Srivastava, *Chlorinated pesticides and heavy metals in human semen*. Int J Androl, 2000. 23(3): p. 145-9.
32. Solomon, G.M. and P.M. Weiss, *Chemical contaminants in breast milk: time trends and regional variability*. Environ Health Perspect, 2002. 110(6): p. A339-47.
33. Jaga, K. and C. Dharmani, *Global surveillance of DDT and DDE levels in human tissues*. Int J Occup Med Environ Health, 2003. 16(1): p. 7-20.
34. Younglai, E.V., W.G. Foster, E.G. Hughes, K. Trim, and J.F. Jarrell, *Levels of environmental contaminants in human follicular fluid, serum, and seminal plasma of couples undergoing in vitro fertilization*. Arch Environ Contam Toxicol, 2002. 43(1): p. 121-6.

35. Baukloh, V., H.G. Bohnet, M. Trapp, W. Heeschen, W. Feichtinger, and P. Kemeter, *Biocides in human follicular fluid*. Ann N Y Acad Sci, 1985. 442: p. 240-50.
36. Tan, B.L. and M. Ali Mohd, *Analysis of selected pesticides and alkylphenols in human cord blood by gas chromatograph-mass spectrometer*. Talanta, 2003. 61(3): p. 385-91.
37. Tan, J., A. Loganath, Y.S. Chong, and J.P. Obbard, *Exposure to persistent organic pollutants in utero and related maternal characteristics on birth outcomes: a multivariate data analysis approach*. Chemosphere, 2009. 74(3): p. 428-33.
38. Foster, W., S. Chan, L. Platt, and C. Hughes, *Detection of endocrine disrupting chemicals in samples of second trimester human amniotic fluid*. J Clin Endocrinol Metab, 2000. 85(8): p. 2954-7.
39. Bradman, A., D.B. Barr, B.G. Claus Henn, T. Drumheller, C. Curry, and B. Eskenazi, *Measurement of pesticides and other toxicants in amniotic fluid as a potential biomarker of prenatal exposure: a validation study*. Environ Health Perspect, 2003. 111(14): p. 1779-82.
40. U.S. Department of Agriculture, *Pesticide Data Program; Annual Summary, Calander Year 2006*, Washington, DC. [available: <http://www.ams.usda.gov/AMSV1.0/getfile?dDocName=STELPRDC5064786>].
41. U.S Department of the Interior- U.S. Geological Survey, *The Quality of Our Nation's Waters. Pesticides in the Nation's Streams and Ground Water, 1992-2001.*, N.W.-Q.A. Program, Editor. 2006.
42. Kiely, T., D. Donaldson, and A. Grube, *Pesticide Industry Sales and Usage. 2000 and 2001 Market Usage*. 2004, US EPA.
43. U.S. Environmental Protection Agency, *An Inroduction to Indoor Air Quality*. 2008.
44. Mendola, P., L.C. Messer, and K. Rappazzo, *Science linking environmental contaminant exposures with fertility and reproductive health impacts in the adult female*. Fertil Steril, 2008. 89(2 Suppl): p. e81-94.
45. Wigle, D.T., M.C. Turner, and D. Krewski, *A Systematic Review and Meta-analysis of Childhood Leukemia and Parental Occupational Pesticide Exposure*. Environmenal Health Perspectives, 2009 (epub).
46. Whorton, D. and D. Foliart, *DBCP: eleven years later*. Reprod Toxicol, 1988. 2(3-4): p. 155-61.
47. Whorton, D., R.M. Krauss, S. Marshall, and T.H. Milby, *Infertility in male pesticide workers*. Lancet, 1977. 2(8051): p. 1259-61.
48. Wigle, D.T., T.E. Arbuckle, M.C. Turner, A. Berube, Q. Yang, S. Liu, and D. Krewski, *Epidemiologic evidence of relationships between reproductive and child health outcomes and environmental chemical contaminants*. J Toxicol Environ Health B Crit Rev, 2008. 11(5-6): p. 373-517.
49. Infante-Rivard, C. and S. Weichenthal, *Pesticides and childhood cancer: an update of Zahm and Ward's 1998 review*. J Toxicol Environ Health B Crit Rev, 2007. 10(1-2): p. 81-99.
50. EWG. *Skin Deep Cosmetic Safety Database*. 2007 [cited 2008 March 13, 2008]; Available from: <http://www.cosmeticsdatabase.com/faq.php#2>.
51. Grandjean, P., D. Bellinger, A. Bergman, S. Cordier, G. Davey-Smith, B. Eskenazi, D. Gee, K. Gray, M. Hanson, P. van den Hazel, J.J. Heindel, B. Heinzow, I. Hertz-Picciotto, H. Hu, T.T. Huang, T.K. Jensen, P.J. Landrigan, I.C. McMillen, K. Murata, B. Ritz, G. Schoeters, N.E. Skakkebaek, S. Skerfving, and P. Weihe, *The faroes statement: human health effects of developmental*

- exposure to chemicals in our environment*. Basic Clin Pharmacol Toxicol, 2008. 102(2): p. 73-5.
52. Weiss, B., D. Cory-Slechta, S.G. Gilbert, D. Mergler, E. Miller, C. Miller, M.C. Newland, D. Rice, and T. Schettler, *The new tapestry of risk assessment*. Neurotoxicology, 2008. 29(5): p. 883-90.
 53. Cory-Slechta, D.A., B. Weiss, and J. Cranmer, *The environmental etiologies of neurobehavioral deficits and disorders: weaving complex outcomes and risk modifiers into the equation*. Neurotoxicology, 2008. 29(5): p. 759-60.
 54. Diamanti-Kandarakis, E., J.P. Bourguignon, L.C. Giudice, R. Hauser, G.S. Prins, A.M. Soto, R.T. Zoeller, and A.C. Gore, *Endocrine-disrupting chemicals: an Endocrine Society scientific statement*. Endocr Rev, 2009. 30(4): p. 293-342.
 55. American Academy of Pediatrics Committee on Environmental Health, *Pediatric Environmental Health*, R.A. Etzel, Editor. 2003, American Academy of Pediatrics: Elk Grove Village, IL.
 56. McDiarmid, M.A., P.M. Gardiner, and B.W. Jack, *The clinical content of preconception care: environmental exposures*. Am J Obstet Gynecol, 2008. 199(6 Suppl 2): p. S357-61.
 57. Titler, M.G., *Chapter 7: The Evidence for Evidence-based practice implementation*. , in *An Evidence-Based Handbook for Nurses*, R.G. Hughes, Editor. 2008, Agency for Healthcare Research and Quality.
 58. Lagoy, C.T., N. Joshi, J.D. Cragan, and S.A. Rasmussen, *Medication use during pregnancy and lactation: an urgent call for public health action*. J Womens Health (Larchmt), 2005. 14(2): p. 104-9.
 59. Lo, W.Y. and J.M. Friedman, *Teratogenicity of recently introduced medications in human pregnancy*. Obstet Gynecol, 2002. 100(3): p. 465-73.
 60. Swiglo, B.A., M.H. Murad, H.J. Schunemann, R. Kunz, R.A. Vigersky, G.H. Guyatt, and V.M. Montori, *A case for clarity, consistency, and helpfulness: state-of-the-art clinical practice guidelines in endocrinology using the grading of recommendations, assessment, development, and evaluation system*. J Clin Endocrinol Metab, 2008. 93(3): p. 666-73.
 61. West, S., V. King, T.S. Carey, K.N. Lohr, N. McKoy, S.F. Sutton, and L. Lux, *Systems to rate the strength of scientific evidence*. Evid Rep Technol Assess (Summ), 2002(47): p. 1-11.
 62. U.S. EPA, *Guidelines for Carcinogen Risk Assessment*, Risk Assessment Forum. Washington, DC. [available: http://oaspub.epa.gov/eims/eimscomm.getfile?p_download_id=439797].
 63. Institute of Medicine, *Environmental Health Sciences Decision-Making: Risk Management, Evidence, and Ehtics: Workshop Summary*. The National Academies Press, 2009.
 64. Wilson, M.P., D.A. Chia, and B.C. Ehlers, *Green chemistry in California: a framework for leadership in chemicals policy and innovation*. New Solut, 2006. 16(4): p. 365-72.
 65. U.S. Food and Drug Administration, *Development and Approval Process (Drugs)*, [available: <http://www.fda.gov/Drugs/DevelopmentApprovalProcess/default.htm>].
 66. Raffensperger, C. and J. Tickner, eds. *Protecting Public Health and the Environment: Implementing the Precautionary Principle*. 1999, Island Press: Washington, DC.
 67. Lanphear, B.P., R. Hornung, J. Khoury, K. Yolton, P. Baghurst, D.C. Bellinger, R.L. Canfield, K.N. Dietrich, R. Bornschein, T. Greene, S.J. Rothenberg, H.L. Needleman, L. Schnaas, G. Wasserman, J. Graziano, and R. Roberts, *Low-level*

- environmental lead exposure and children's intellectual function: an international pooled analysis.* Environ Health Perspect, 2005. 113(7): p. 894-9.
68. Gilbert, S.G. and B. Weiss, *A rationale for lowering the blood lead action level from 10 to 2 microg/dL.* Neurotoxicology, 2006. 27(5): p. 693-701.
 69. Jusko, T.A., C.R. Henderson, B.P. Lanphear, D.A. Cory-Slechta, P.J. Parsons, and R.L. Canfield, *Blood lead concentrations < 10 microg/dL and child intelligence at 6 years of age.* Environ Health Perspect, 2008. 116(2): p. 243-8.
 70. Sackett, D.L., W.M. Rosenberg, J.A. Gray, R.B. Haynes, and W.S. Richardson, *Evidence based medicine: what it is and what it isn't.* 1996. Clin Orthop Relat Res, 2007. 455: p. 3-5.
 71. Montori, V.M. and G.H. Guyatt, *Progress in evidence-based medicine.* JAMA, 2008. 300(15): p. 1814-6.
 72. Institute of Medicine, *Evidence-Based medicine and the changing nature of health care:2007 IOM annual meeting summary.* 2008, IOM: Washington, DC.
 73. Clarke, M. and A.D. Oxman, *Cochrane Reviewer's Handbook 4.0.* 1999: The Cochrane Collaboration.
 74. The Cochrane Collaboration. [cited; Available from: <http://www.cochrane.org/index.htm>.
 75. Ebell, M.H., J. Siwek, B.D. Weiss, S.H. Woolf, J. Susman, B. Ewigman, and M. Bowman, *Strength of recommendation taxonomy (SORT): a patient-centered approach to grading evidence in the medical literature.* J Am Board Fam Pract, 2004. 17(1): p. 59-67.
 76. GRADE Working Group. [cited; Available from: <http://www.gradeworkinggroup.org/society/index.htm>.
 77. Guyatt, G.H., A.D. Oxman, G.E. Vist, R. Kunz, Y. Falck-Ytter, P. Alonso-Coello, and H.J. Schunemann, *GRADE: an emerging consensus on rating quality of evidence and strength of recommendations.* BMJ, 2008. 336(7650): p. 924-6.
 78. Guyatt, G.H., A.D. Oxman, R. Kunz, Y. Falck-Ytter, G.E. Vist, A. Liberati, and H.J. Schunemann, *Going from evidence to recommendations.* BMJ, 2008. 336(7652): p. 1049-51.
 79. Guyatt, G.H., A.D. Oxman, R. Kunz, G.E. Vist, Y. Falck-Ytter, and H.J. Schunemann, *What is "quality of evidence" and why is it important to clinicians?* BMJ, 2008. 336(7651): p. 995-8.
 80. Schunemann, H.J., A.D. Oxman, J. Brozek, P. Glasziou, R. Jaeschke, G.E. Vist, J.W. Williams, Jr., R. Kunz, J. Craig, V.M. Montori, P. Bossuyt, and G.H. Guyatt, *Grading quality of evidence and strength of recommendations for diagnostic tests and strategies.* BMJ, 2008. 336(7653): p. 1106-10.
 81. Guyatt, G.H., A.D. Oxman, R. Kunz, R. Jaeschke, M. Helfand, A. Liberati, G.E. Vist, and H.J. Schunemann, *Incorporating considerations of resources use into grading recommendations.* BMJ, 2008. 336(7654): p. 1170-3.
 82. Jaeschke, R., G.H. Guyatt, P. Dellinger, H. Schunemann, M.M. Levy, R. Kunz, S. Norris, and J. Bion, *Use of GRADE grid to reach decisions on clinical practice guidelines when consensus is elusive.* BMJ, 2008. 337: p. a744.
 83. Quigley, D., F. Simmons, H. Whyte, J. Robertson, and D. Freshwater, *Variations in reproductive and developmental toxicant identification.* J. Chem. HealthSafety, 2000.
 84. European Environment Agency, *Late lessons from early warnings: The precautionary principle 1896-2000.* 2001: Luxembourg.
 85. US EPA. 2003 [cited; Available from: <http://www.epa.gov/EPA-PEST/2003/September/Day-05/p22546.htm>.

86. Matthews, E.J., N.L. Kruhlak, R. Daniel Benz, and J.F. Contrera, *A comprehensive model for reproductive and developmental toxicity hazard identification: I. Development of a weight of evidence QSAR database*. Regul Toxicol Pharmacol, 2007. 47(2): p. 115-35.
87. California Department of Public Health, *Health HazardAlert. 1 bromopropane*. 2003, Hazard Evaluation System and Information Service.
88. Als-Nielsen, B., W. Chen, C. Gluud, and L.L. Kjaergard, *Association of funding and conclusions in randomized drug trials: a reflection of treatment effect or adverse events?* JAMA, 2003. 290(7): p. 921-8.
89. Howick, J., P. Glasziou, and J.K. Aronson, *The evolution of evidence hierarchies: what can Bradford Hill's 'guidelines for causation' contribute?* J R Soc Med, 2009. 102(5): p. 186-94.
90. Glasziou, P., I. Chalmers, M. Rawlins, and P. McCulloch, *When are randomised trials unnecessary? Picking signal from noise*. BMJ, 2007. 334(7589): p. 349-51.
91. Concato, J., *Observational versus experimental studies: what's the evidence for a hierarchy?* NeuroRx, 2004. 1(3): p. 341-7.
92. Concato, J., N. Shah, and R.I. Horwitz, *Randomized, controlled trials, observational studies, and the hierarchy of research designs*. N Engl J Med, 2000. 342(25): p. 1887-92.
93. Zapponi, G.A., I. Marcello, and A. Carere, *Prevention, ethics and science: lessons from Lorenzo Tomatis*. Ann Ist Super Sanita, 2008. 44(1): p. 8-12.
94. Grandjean, P. and P.J. Landrigan, *Developmental neurotoxicity of industrial chemicals*. Lancet, 2006. 368(9553): p. 2167-78.
95. Megie, G., *From stratospheric ozone to climate change: historical perspective on precaution and scientific responsibility*. Sci Eng Ethics, 2006. 12(4): p. 596-606.
96. Science and Environmental Health Network. *Wingspread Statement on the Precautionary Principle*. 1998 [cited; Available from: <http://www.sehn.org/state.html#wUH>].
97. Gee, D., *Establishing evidence for early action: the prevention of reproductive and developmental harm*. Basic Clin Pharmacol Toxicol, 2008. 102(2): p. 257-66.
98. Sakala, C. and M.P. Corry, *Evidence-based maternity care: What is it and what can it achieve*. 2008, Millbrank Memorial Fund.
99. Murad, M.H., V.M. Montori, and G.H. Guyatt, *Incorporating patient preferences in evidence-based medicine*. JAMA, 2008. 300(21): p. 2483; author reply 2483-4.
100. Boivin, A., J. Green, J. van der Meulen, F. Legare, and E. Nolte, *Why Consider Patients' Preferences?: A Discourse Analysis of Clinical Practice Guideline Developers*. Med Care, 2009.
101. Soumerai, S.B., D. Ross-Degnan, and J.S. Kahn, *Effects of professional and media warnings about the association between aspirin use in children and Reye's syndrome*. Milbank Q, 1992. 70(1): p. 155-82.
102. Levine, M., S. Walter, H. Lee, T. Haines, A. Holbrook, and V. Moyer, *Users' guides to the medical literature. IV. How to use an article about harm. Evidence-Based Medicine Working Group*. JAMA, 1994. 271(20): p. 1615-9.
103. Woodruff, T.J., L. Zeise, D.A. Axelrad, K.Z. Guyton, S. Janssen, M. Miller, G.G. Miller, J.M. Schwartz, G. Alexeeff, H. Anderson, L. Birnbaum, F. Bois, V.J. Cogliano, K. Crofton, S.Y. Euling, P.M. Foster, D.R. Germolec, E. Gray, D.B. Hattis, A.D. Kyle, R.W. Luebke, M.I. Luster, C. Portier, D.C. Rice, G. Solomon, J. Vandenberg, and R.T. Zoeller, *Meeting report: moving upstream-evaluating adverse upstream end points for improved risk assessment and decision-making*. Environ Health Perspect, 2008. 116(11): p. 1568-75.

104. Miller, M., M. Marty, A. Arcus, J. Brown, D. Morry, and M. Sandy, *Differences between children and adults: Implications for risk assessment at California EPA*. . International Journal of Toxicology, 2002. 21: p. 402-418.
105. Peterson, E.D., *Research methods to speed the development of better evidence-the registries example.*, in *Evidence-based medicine and the changing nature of health care:2007 IOM annual meeting summary*. 2008, The National Academies Press: Washington, DC.
106. Tricoci, P., J.M. Allen, J.M. Kramer, R.M. Califf, and S.C. Smith, Jr., *Scientific evidence underlying the ACC/AHA clinical practice guidelines*. JAMA, 2009. 301(8): p. 831-41.
107. Rodwin, M.A., *The politics of evidence-based medicine*. J Health Polit Policy Law, 2001. 26(2): p. 439-46.
108. Shaneyfelt, T.M. and R.M. Centor, *Reassessment of clinical practice guidelines: go gently into that good night*. JAMA, 2009. 301(8): p. 868-9.
109. Halvorson, G.C., *Electronic medical records and the prospect of real time evidence development.*, in *Evidence-based medicine and the changing nature of health care: 2007 IOM annual meeting summary*. 2008, The National Academies Press: Washington, DC.
110. Lein, P., P. Locke, and A. Goldberg, *Meeting report: alternatives for developmental neurotoxicity testing*. Environ Health Perspect, 2007. 115(5): p. 764-8.
111. Cory-Slechta, D.A., *Studying toxicants as single chemicals: does this strategy adequately identify neurotoxic risk?* Neurotoxicology, 2005. 26(4): p. 491-510.
112. Callahan, M.A. and K. Sexton, *If cumulative risk assessment is the answer, what is the question?* Environ Health Perspect, 2007. 115(5): p. 799-806.