

26 Science for precautionary decision-making

Philippe Grandjean ⁽¹⁾

The goals of academic researchers may differ from those of regulatory agencies responsible for protecting the environment. Thus, research must take into account issues such as feasibility, merit and institutional agendas, which may lead to inflexibility and inertia.

A large proportion of academic research on environmental hazards therefore seems to focus on a small number of well studied environmental chemicals, such as metals. Research on environmental hazards should therefore to a greater extent consider poorly known problems, especially the potential hazards about which new information is in particular need.

Misinterpretation may occur when results published in scientific journals are expressed in hedged language. For example, a study that fails to document with statistical significance the presence of a hazard is often said to be negative, and the results may be misinterpreted as evidence that a hazard is absent. Such erroneous conclusions are inspired by science traditions, which demand meticulous and repeated examination before a hypothesis can be said to be substantiated.

For prioritising needs for action, research should instead focus on identifying the possible magnitude of potential hazards. Research is always affected by uncertainties and many of them can blur a real association between an environmental hazard and its adverse effects, thereby resulting in an underestimated risk. Environmental health research therefore needs to address the following question: are we sufficiently confident that this exposure to a potential hazard leads to adverse effects serious enough to initiate transparent and democratic procedures to decide on appropriate intervention?

The choice of research topics must consider societal needs for information on poorly known and potentially dangerous risks. The research should be complementary and extend current knowledge, rather than being repetitive for verification purposes, as required by the traditional science paradigm. Research findings should be openly available and reported so that they inform judgements concerning the possible magnitude of suspected environmental hazards, thereby facilitating precautionary and timely decision-making.

⁽¹⁾ The author would like to thank Mette Eriksen for conducting the searches on SciFinder. Helpful contributions on earlier versions were made by John Bailar, Carl Cranor, David Gee and David Kriebel.

26.1 Science and the Precautionary Principle

The case studies in this volume illustrate that science can provide powerful evidence for targeted prevention to protect against hazards to the environment and public health. However, the chapters also shows how science can be insufficient, and it can be misinterpreted or ignored, so that appropriate intervention is deferred or abandoned. This chapter explores the strengths and weaknesses of environmental health research seen from the perspective of the wider needs of society and the use of 'Precautionary Principle' (PP).

At the outset, societal investment in environmental research would seem unwise if it is irrelevant, poor, or difficult to access. Research support should favour studies that stimulate timely decision-making and prudent action to prevent hazards. While not disregarding the need for basic research, I shall focus primarily on the weaknesses of current applied research in the environmental field and the possible avenues for science to become more useful for future environmental health decision-making.

Some researchers have raised the concern that the PP may potentially make further research redundant, given that an intervention has already been decided upon (Goldstein and Carruth, 2004). But any decision on environmental health hazards should be considered tentative and amenable to change, as justified by further research (including intervention studies to determine if the action had the intended effect). The basic problem is that prevention has too often been deferred due in part to the alleged absence of convincing scientific evidence, as illustrated by the case studies in this volume⁽²⁾. The error is recognised only when decisive evidence has finally been gathered, and it is realised that action should have been initiated much earlier on. With time, nearly all exposure limits for hazardous agents have decreased as new evidence documented that harm occurred at lower exposure levels than previously believed. Thus, when scientific evidence is incomplete, environmental standards are more lenient. But can science provide better support for prudent decision-making, so that adequate protection may be decided upon from the beginning?

For research to provide sufficient documentation for potential intervention, it has to be both reliable and pertinent. Thus, the quality of research has two sides — the methodology and the utility. One could also refer to these two aspects as the validity and the relevance. The two are of course related, but even research considered 'poor' from a narrow methodological perspective could nonetheless be highly relevant. Still, a study of limited validity is most likely also to have little impact, especially if the conclusions cannot be trusted. While the researchers should focus on securing a high methodological level, that should not turn them into sceptical ivory-tower nit-pickers preoccupied with methodological precision and technical detail. On the other hand, focusing mainly on environmental implications of the research can lead to inappropriate (or apparent) advocacy for particular policies or precautionary action that may be inspired, though perhaps not justified, by the research.

Environmental health is often considered a field of applied research, usually multidisciplinary. Researchers and their employers are engaged in science not just for purely altruistic reasons. Universities and other research institutions are enterprises that need to fulfil the institution's mandates, satisfy requirements stipulated by funding sources, and avoid going into debt. Within the EU, more than half of the research and development activities carried out are funded by industry, while slightly more than one-third is paid for from public sources (Eurostat, 2011). The EU's new Horizon 2020 research programme is intended to increase the public financing of 'smart investment' in research and innovation while dealing with pressing societal challenges, including climate change and environmental health problems (EC, 2011). Given the substantial public investment in research (van den Hove et al., 2011), one would anticipate that environmental research, especially the part of it that is reported in academic journals, would somehow reflect priorities expressed by regulatory agencies and other public bodies. The next section of this chapter will therefore examine the research coverage of environmental chemicals and whether poorly documented and potential hazards receive appropriate attention. But there is more to it than the coverage of priority topics.

(²) As the preface to the first volume of *Late lessons from early warnings* (EEA, 2001) pointed out, 'the absence of political will to take action to reduce hazards in the face of conflicting costs and benefits seems to be an even more important factor in these histories than is the availability of trusted information'.

Under PP-based decision-making, scientific proof or a very high degree of certainty are not required. Incomplete, but reliable evidence can be sufficient to justify a precautionary intervention. On the other hand, if extensive evidence is available, then a conventional risk assessment and subsequent prevention are indicated, and there would be little need to invoke the PP. However, traditional risk assessment is sometimes anti-precautionary when it demands convincing evidence and thus ignores emerging insight and incomplete documentation. Due to its focus on scientific justification, risk assessment may inspire continued elaboration of fairly well documented hazards, so that remaining uncertainties can be resolved to allow firm decisions. When decisions are PP-based, less extensive evidence is required, and some uncertainties are accepted as being inevitable or impossible to remove in the time available for preventing plausible harm. The less extensive requirement regarding scientific evidence can have significant implications for the ways that research is planned, performed, analysed, interpreted and reported (Grandjean, 2008b).

We rely on science as evidence to help justify decisions on environmental hazards. But, as the case studies in the present and the first volume of *Late lessons from early warnings* (EEA, 2001) clearly demonstrate, science does not automatically lead to appropriate prevention or precautionary action. Thus, neither the quality nor the relevance of the science as such will necessarily translate into responsible and prudent decisions. Still, the interpretation of incomplete research data, the evaluation of uncertainties and misunderstandings of the findings can obfuscate the discussion on the urgency of possible environmental protection. So the question must therefore be asked: Can science somehow better serve to support better public policy decisions?

I think that the answer is yes, although better quality and relevance in terms of PP-based decision-making may not be easy to achieve. This chapter will focus on four main issues listed below.

Concerns regarding science as evidence for decisions on environmental hazards:

- 1) Does the research cover the societal needs for supporting information on suspected, poorly documented or potential hazards?
- 2) Does the research explore new and emerging hazards so that it could serve as an early warning system?

- 3) Is the reporting of research findings appropriate to serve as evidence for reducing environmental hazards?
- 4) If the research is available, is it reliable and independent of vested interests?

26.2 Current research focus is on well-known hazards

The most appropriate and feasible way to assess the topics covered by environmental research is to carry out bibliometric analyses using internet-based databases on scientific publications. Environmental journals are usually categorised in the fields of toxicology, environmental sciences and public health (a total of 78 major journals in both Web of Science and the PubMed database). The Web of Science covers scientific literature back to 1899, but searches are limited to chemical names in the titles of journal articles. However, for recent publications, it is possible to use the SciFinder database, where individual environmental chemicals can be identified from their Chemical Abstract Service (CAS) registry numbers. Using these internet resources, information can be retrieved on how often scientific publications have dealt with chemicals of interest from an environmental viewpoint (Grandjean et al., 2011).

As a starting point, we first used Web of Science to examine the coverage of the seven chemical substances from the 14 case studies reviewed in *Late lessons from early warnings* Volume 1 (EEA, 2001). Table 26.1 shows the number of articles published in the relevant journals during the years 2000–2009, i.e. the 10 years right after the completion of the report. One could have expected that these early warning substances would have faded somewhat from the science radar, given that their environmental impact had already been recognised during the 20th century and that some had been banned several decades ago. However, the number of scientific publications on these substances during 2000–2009 corresponded to about 40 % of all articles available since 1899. The relative coverage before and after year 2000 differed somewhat between the substances. Both sulfur dioxide and DES clearly faded during recent years, with only about one quarter of all titles available in environmental and toxicological publications since 1899 being published during 2000–2009. On the other hand, MTBE became more popular, with three-quarters of all papers available since 1899 having been published during the first decade of this century.

Table 26.1 Total numbers of articles published 1899–2009 and during the most recent 10 years (and percent of total) in environmental and toxicology journals listing 'early warnings' substances in the title

Name	Total number of articles 1899–2009	Articles published 2000–2009	
		Number	% of total
Polychlorinated biphenyls (PCBs)	5 809	2 738	47
Asbestos	2 735	809	30
Sulfur dioxide	2 380	548	23
Benzene	2 075	879	42
Tributyltin (TBT)	672	344	51
Diethylstilboestrol (DES)	603	147	24
Methyl-tert-butyl ether (MBTE)*	542	411	76
Total	14 816	5 876	40

Note: *MBTE uses were expanded during the late 20th century, not restricted like the other substances in the table.

Source: Data from the Web of Science (Grandjean et al., 2011).

These numbers suggest that substantial research continued to be published on these substances, long after the recognition of their importance as environmental contaminants. However, the numbers extracted from the Web of Science are incomplete, as a research article might well address a chemical without the substance name appearing in the title. Thus, when extracting data from SciFinder, we obtained a greater number of articles (total of 8 267 during 2000–2009, asbestos not included). With an average of over 10 scientific articles per substance per month, these early warnings chemicals remained a significant focus of research reports published since 2000. PCBs, in particular, remained very much in focus, as I shall discuss shortly.

Given the continued attention paid to these chemicals characterised by 'early warnings', what about environmental chemicals in general? Thousands of potentially toxic chemicals are being released into the environment, and there is a need to determine their persistence, dissemination, biomagnification and toxic effects, especially when only minimal information is available. So how does published research reflect the societal needs to cover a wide range of potential hazards?

Based on CAS number links from the science journals during 2000–2009, the substances can be ranked in accordance with their numbers of publication links (as SciFinder is not limited to environmental chemicals, we had to manually exclude radioactive isotopes, enzymes, metabolites, etc.). All told, 119 636 articles were published by the 78 scientific journals during the first ten years of this millennium. SciFinder listed a total of 760 056 CAS links from these articles

(Grandjean et al., 2011). Thus, on average, each of the many scientific articles had six CAS links, thus not only describing a single substance at a time. The total numbers of publications and links are large and reflect an intense publication activity. However, the coverage turned out to be extremely uneven.

We focused on the 100 most frequent environmental chemicals. Each of them was covered in a minimum of 600 articles — and up to 10 000 — during the 10-year period. Thus, each of the top-100 substances would be addressed in about five to 80 articles every month. The total number of links to the top-100 environmental chemicals was 180 822. Thus, the vast majority of the many thousand chemicals listed were far less popular than the top-100. This finding suggests that research on environmental chemicals is and has been for some time fairly narrowly focused on a limited number of substances.

This conclusion becomes very clear when we examine the 20 most commonly studied environmental chemicals. Each had between 2 000 and 10 000 CAS links during the first ten years of the millennium. The sum of article links corresponds to 12 % of all CAS number links. Assuming they also represent 12 % of all published articles, one or more of these substances would be featured in 14 264 publications during the 10 years, or 119 articles per month, on average. To keep up with the literature in the top-20 substances only, one would have to read five or six papers every work day, without holiday breaks.

All of the top-10 substances are metals (including arsenic, which is regarded as a semimetal). Also well covered are several tar chemicals (polyaromatic

hydrocarbons), solvents and the PCBs — already known from *Late lessons from early warnings* Volume 1 (EEA, 2001) and Table 26.1. For the top-20 substances, an average of 51 % of all articles available in 2009 had been published within the most recent 10 years. Some variation was present: arsenic increased in popularity (74 % during the most recent 10 years), while aluminium decreased after the year 2000 (31 %). Also, the tar chemicals often found in air pollution (e.g. benzo[a]pyrene and phenanthrene) tended to appear more often in recent article titles. Overall, these results show that the chemicals most commonly studied in recent years had already been extensively studied during the previous century. Thus, the chemicals that were popular during the previous century remained a focus (Grandjean et al., 2011).

Two of the top-20 chemicals — lead and mercury — are included in the case studies, and Table 26.2 shows the results for the main substances reviewed in this volume. Thus, whether or not the chemicals are persistent in the environment or the human body, some of them have clearly become persistent and highly prominent in the scientific literature. The tens of thousands of articles on lead, mercury, and other well recognised environmental hazards testify to the enormous investments in studying, reporting and publishing on these prominent substances. It would therefore seem that the choice of research topic in the field of environmental health greatly benefits the well-known chemicals.

The next question to consider is whether research addresses societal needs for more poorly known and potentially dangerous risks. Does academic research in environmental chemicals ignore less-well known compounds that need documentation? We conducted additional studies to examine this question.

26.3 Ignoring new potential environmental hazards

We now focus on the other end of the spectrum, as many environmental chemicals have not been adequately tested. When the US National Research Council conducted a study in the 1980s on toxicity testing, 78 % of the industrial chemicals most commonly produced was found without even minimal test data for toxicity (NRC, 1984). Later follow-up showed little improvement (US EPA, 1998). Even today, the European Chemical Agency complains that gaps in safety data remain and that little has been done to mend the problem so far (Gilbert, 2011). Thus, as metals and tar chemicals attract much research attention, are substances of importance to society being neglected by environmental researchers?

To examine this question, we looked at the high-production chemicals considered in particular need of scientific documentation (US EPA, 2009). This high-priority list was first published in 2006 and included thirteen important substances lacking both a robust hazard data set and exposure information. For the time period of 2000–2009, we found that these chemicals had a total of only 352 links to scientific articles, i.e. an average of only three per month for the entire group (Grandjean et al., 2011). Five of the thirteen high-priority substances were not encountered at all in the 78 journals during the ten years. One could excuse the lack of coverage up to 2006, when the EPA published its list, and perhaps 2007. However, when extending the search to 2010 and 2011, the result was pretty much the same — the priority listing had not inspired any increased number of publications in scientific journals. When compared with the staggering numbers for top-20 substances, the

Table 26.2 Numbers of articles published in environmental science journals during 2000–2009 on chemicals covered in the present volume, as determined from SciFinder links to the CAS numbers

Name	CAS no.	Number of links	Rank
Lead	7439-92-1	8 926	2
Mercury	7439-97-6	4 399	9
<i>p,p'</i> -DDT	50-29-3	1 968	21
Bisphenol A	80-05-7	952	62
Perchloroethylene	127-18-4	898	68
Beryllium	7440-41-7	400	235
Vinylchloride	75-01-4	319	276
Dibromochloropropane (DBCP)	96-12-8	41	> 1 000

publication rates for these high-priority substances appear tiny.

Other substances may be considered likely emerging hazards, about which further information would be highly useful. Triclosan is a biocide often used in cosmetics, but releases into the environment have raised concern (Dann and Hontela, 2011). There were 259 articles on this chemical during 2000–2009, much better than the high-priority substances listed by the US EPA, but way below the popularity of toxic metals. Likewise, the perfluorinated compounds have been in use for decades, and concerns about their environmental fate and toxicity have grown (Lindstrom et al., 2011). The most prevalent member of this group, perfluorinated octanoic sulfate was covered in 271 articles, about the same as triclosan. Thus, each of them was addressed only in about two dozen scientific articles in the scientific journals every year. Accordingly, about 35 articles would focus on lead (and close to 20 on mercury) each time a single article would present evidence on one of these emerging hazards. But can we trust these numbers?

Although the bibliometric data do not distinguish between short, descriptive reports and thorough reviews, the overwhelming emphasis on a small minority of environmental chemicals cannot be explained away. Also, the scientific journals may not reflect research activities outside academic institutions, but one would have to imagine huge numbers of reports outside the mainstream journals to make up for the differences. The conclusion therefore seems inevitable that the long-term prominence of substances commonly covered in articles in environmental journals does not match the societal needs or those of regulatory agencies. Substances that were highly popular in research during the previous century remained so during the first ten years of the present millennium, despite the changing needs for evidence on environmental impacts.

26.4 Inertia and its reasons

An important reason for such inertia and continued focus on well-known substances may relate to the traditional science paradigm, where solid conclusions depend on replication and verification. While a single study should not be relied upon as

firm evidence, the extent to which replication is needed seems to have been stretched to the extreme, when well-known environmental chemicals inspire almost 1 000 publications per year.

It may well be that academic researchers do not know or contemplate the needs for environmental health documentation. We may question environmental researchers, who keep studying lead toxicity to obtain even more detailed or perfect results⁽³⁾. However, individual researchers and their institutions may have insufficient access to public and private funding that would allow an unrestricted choice of research topics. This limitation would especially refer to young researchers of low academic rank. Further, if students are taught to replicate and extend their mentor's own research, they will later become the seniors with the same type of expertise and narrow focus on well-known environmental hazards. Existing expertise as well as facilities may favour a continued focus on the same hazards, thereby propagating long-term traditions and ignoring society's changing needs for early warning investigations. In more general terms, a tendency to maintain a narrow focus is likely to be counter-productive in regard to scientific discovery and innovation, as there would potentially be much more to learn from studying new hazards than from replicating studies on old ones.

Several factors may contribute to the estrangement of academic research from societal needs for documentation on environmental hazards. Research institutions have an interest in maintaining highly qualified personnel and efficient use of costly infrastructure. All of the most popular chemicals can be inexpensively measured by instruments that became widely available already in the 1970s and 1980s. Analytical methodologies are already established and well documented. These instruments (atomic absorption spectrometers and gas chromatographs) make it possible within a week or so to generate results sufficient to justify a scientific paper on one or more of the top-20 substances. Under these conditions, why would ambitious researchers and their students take on new substances that might require the purchase of expensive equipment and arduous development of new methods?

The loyalty to established methods and research topics is not just a matter of convenience. In

⁽³⁾ The author has published numerous articles on lead, mercury, and other top-20 chemicals in scientific journals, thus being part of the inertia. However, as the chapters on lead and mercury show, real or alleged uncertainties were often used to argue against hazard abatement, thereby requiring more research.

academic research, competition is fierce and each researcher must demonstrate his or her qualifications by frequently publishing articles in scholarly journals. By endeavouring to research the unknown, these researchers would face longer time periods between publications, if any. The mere number of publications is a crucial metric for academic prestige and for obtaining a tenured position. By using existing instruments and methods, a researcher can more effectively expand the CV, especially if the reports can be framed into small incremental manuscripts, each of them contributing an entry on the publications list. So-called vanity publications may contaminate the scientific literature, as they contain little new information, but primarily serve to augment the author's credentials. Whether they contribute new insight then becomes a secondary concern. Similarly, the budget in many research departments is tied to the number of scientific publications, thus also favouring quantity over quality. Such a focus on publication numbers may deny the higher societal goals of environmental research while promoting earthly aspects of personal desires and academic reputation.

The pressure to complete a project on time (or even before the deadline) and to publish the findings with minimal delay also invites the use of short-cuts. Convenience and lack of funds may determine that some parameters in a study are not measured in appropriate detail, e.g. by relying on questionnaire responses rather than actual measurements, which may be too expensive. When a study claims to address an environmental hazard using study parameters that are unreliable or perhaps not representative, the results will often be non-informative. Worse, the results may be interpreted as evidence against the hazard causing any risk at all. Such misleading conclusions are sometimes referred to as Type III errors (Schwartz and Carpenter, 1999). I shall return to this problem shortly (see Section 26.5 below).

The inertia and reliance on convenience are not restricted to researchers themselves, or public research institutions, for that matter. It also affects the funding agencies. If a proposed project deals with a known environmental problem, the principal investigator probably has an impressive track record, the protocol is feasible and easy to comprehend and capable reviewers are readily recruited. That may not be the case with poorly studied substances and emerging hazards. The funding agency can feel comfortable about the proposed time schedule and the anticipated outcome of the project, as the exposures and effects rely on established methods. Uncomfortable surprises are unlikely. Hence, it may

be safer and more convenient for grant managers to concentrate on the known hazards.

Scientific journals probably also play a role in maintaining a focus on well-known substances. Peer review of submitted manuscripts is rarely a problem with a manuscript on lead exposure. Bias toward publication of the report may occur when the reviewer finds that his or her own research has been cited, thus demonstrating the sound judgment of the authors of the manuscript. Some of the journals that we explored in the bibliometric databases are regarded as prestigious, with high citation rates. The possibility exists that some environmental chemicals may be held in higher esteem than others, thereby adding to their continued prominence, or publication persistence, no matter what the societal needs may be. This means that there may be an element of circular reasoning involved, where a substance is a popular research item simply because it has been widely studied in the past — a self-prophetic bias that maintains a continued prominence of a small number of scientists and their publications.

The science sociologist Robert K. Merton (Merton, 1968) dubbed this phenomenon a 'Matthew' effect, referring to the New Testament (*'For unto every one that hath shall be given, and he shall have abundance: but from him that hath not shall be taken away even that which he hath'*). Popularity among scientists in the past seems to provide justification for the importance or relevance of continued research in a particular field. The opposite strategy would appear more attractive from the point of view of innovation.

However, it must be said that some conventional research into well-known substances have identified novel scientific breakthroughs that are not only relevant to our understanding of these well-characterised substances, such as mercury and lead, but they have also been scientifically valuable, via analogy, to many other substances.

Clearly, academic research has multiple purposes, a number of constraints and some limitations, when viewed from an environmental health angle. Societal needs for evidence on priority substances or emerging risks are apparently not seen as a high priority for academic research in general. But the choice of research topic is not the only problem.

26.5 Research methodologies and assumptions

Jointly with the inertia in the choice of research topics, traditional scientific thinking may also represent

an obstacle. According to the standard paradigm, we need to justify our conclusions by replicating our findings, securing the highest possible data quality and documenting each component of the anticipated causal link. Such high standards will protect science from making mistakes by claiming, e.g. that lead is toxic to the brain, unless extensive documentation is truly available to back this assertion. The links to scientific traditions extend back to the Leonardo da Vinci's and Galileo's writings. In studying environmental health hazards, the prevalent paradigm determines how the problem is usually framed as below:

'The traditional scientist will address an environmental research question as follows: Have we reliably documented through meticulous study and replication that this substance is mechanistically and causally linked to an adverse biological change?'

Along with the demand for replication, tradition calls for a narrow focus. Uncertainty is commonly restricted through rigorous control of the study setting. The advantage of a well-defined study is that it addresses only a single factor under specific

circumstances and therefore more likely will lead to firm or indisputable conclusions. However, due to its limited scope, the study will at best result only in an incremental increase in knowledge about the overall issue at hand, including multiple or complex exposure scenarios and the significance of individual vulnerability. Thus, the disadvantage is that this approach leads to reductionism and explores only limited or individual aspects of each hazard. Such proximate and simplistic risks poorly represent the true complexity of environmental hazards.

Examination of the chapters on human health hazards in this volume and *Late lessons from early warnings* Volume 1 (EEA, 2001) allows identification of several assumptions that were, at first, considered valid and important, but were later found to be misleading. Table 26.3 shows some of the most crucial – and erroneous – assumptions that were initially made in regard to one or more of the environmental hazards included in this volume. The case studies in this volume show that relying on these assumptions, while seemingly meaningful in terms of the prevailing research paradigm, led to proliferation of environmental hazards due to the substantial delay in their recognition.

Table 26.3 Erroneous assumptions made in initial evaluations of environmental hazards and the subsequent scientific recognition of the true complexity

Initial assumption	Late scientific lesson
1. Presence of environmental chemicals in the body can be tolerated at 'safe' or natural doses	Delayed effects, cumulated or re-mobilised doses, or toxic metabolites may occur at exposures previously thought to be safe
2. Absence of harm in adult male workers (from routine medical data or mortality) means absence of risk to the general public	Sub-populations, such as children and the elderly, may be more vulnerable to the exposure
3. Acute or short-term effects also reflect chronic or long-term effects	Dose-response relationships for acute effects may substantially differ from those for chronic effects
4. Biological effects may not necessarily be adverse and can be considered harmless	Early changes can predict more serious adverse effects which can develop later on
5. Dose-response relationships are consistent (and 'monotonic'), and no risk occurs at doses below apparent thresholds	Some substances show 'low dose' effects that are not readily predictable from responses to high doses
6. Short-term assessment of exposures from a single pathway can generally be considered sensitive and valid	Most methods for exposure assessment are imprecise, and imprecision usually results in underestimation of the toxicity
7. The placenta and the blood-brain barrier amply protect sensitive life-stages and organs from toxic chemicals	The barriers may be bypassed, as they offer limited protection against industrial chemicals
8. Average findings in exposed subjects indicate the potential for harm to the exposed population	Sensitive sub-groups may show effects that are not apparent from the average data
9. Toxicity evidence from animals and wildlife is not relevant to human toxicity	Animal data have reliably predicted most known carcinogens and many other hazards, and humans may be more vulnerable than other species

A crucial assumption was that a biological change may not necessarily reflect an adverse health effect and could therefore be ignored. Within normal variability, it may indicate adaptation or 'hormesis', therefore being innocuous. With proper justification, the assumption may be true, but biological changes should not be disregarded just because they are prevalent (or unwelcome, for some reason). For many years, researchers believed that the inhibition of an enzyme called ALAD due to lead exposure at blood-lead concentrations thought to be low, was a biological change that had no health implications (as described in the Chapter 3). That may be true in a strict sense, as the enzyme in red blood cells has no important function. However, recent research has shown that serious adverse effects do occur at lead exposures that were previously regarded as too low to be harmful.

Also, habitual levels of lead exposure were called 'natural' simply because they were normal or habitual. But prevalent lead exposures were the result of centuries of increasing lead use. Analyses of lead isotopes and of mummified tissues documented that normal lead exposure were far above what could be considered natural.

For efficiency reasons, toxicology studies have aimed at avoiding considerations of the sex, age and strain of the animals used. If focusing on inbred, adult male rats only, important sources of variability were ignored while making the study more efficient and precise. This problem became centre stage when reproductive toxicology and endocrine disruption began to attract attention (see Chapter 13 on ethinyl oestradiol in the aquatic environment and Chapter 10 on BPA, as well as PCBs, DES and TBT in *Late lessons from early warnings* Volume 1 (EEA, 2001, Ch. 6, 8, 13)). Also, it is only a recent discovery that exposures to environmental chemicals may cause much more toxicity if they happen during vulnerable developmental windows (Grandjean et al., 2008). However, prospective studies of birth cohorts take a long time and are extremely costly, and even multi-generation animal assays are often resisted due to economic burdens on industry.

The assumptions in Table 26.3 prevailed for a long time due to the failure of available, though incomplete data to show clear evidence of a risk. If adverse effects were not proven to exist, the

erroneous conclusion was drawn that adverse effects must be absent. Perhaps this is the underlying assumption, which represents the greatest error. It survived, as uncertainties were ignored, whether in regard to exposure assessment, sensitivity of outcome measures, individual vulnerability, statistical analysis methodology or statistical power of the study. Overlooking imprecisions and incompleteness will most often result in underestimation and may lead to rejection of the presence of a (true) risk. Also, these uncertainties are not likely to create spurious associations, unless confounding factors are present.

26.6 Vulnerability of research to criticism

The downside of the traditional strategy to provide ample verification is that science becomes vulnerable to a critique that raises concerns about various possible sources of error or bias, particularly in regard to emerging insights and early warnings. The desire to document the truth, preferably the 'full' truth, makes science vulnerable to purported weaknesses. Thus, while careful scientists must pay meticulous attention to the methodological standards and quality assurance, some colleagues primarily exert these skills when judging the work of colleagues. Such critique may be unjust, but the halo earned from emphasis on the quality of scientific methods thrives from the collusion of admiring colleagues and students (e.g. at scientific conferences).

However, harsh critique and exaggerated scepticism may be particularly inappropriate in regard to emerging insights and early warnings which are often innovative and necessarily tentative⁽⁴⁾. Thus, the case studies illustrate that astute observations by clinicians, factory inspectors, workers, anglers, bee keepers and community members can sometimes provide valid hypotheses on new hazards that are only confirmed by in-depth research much later.

A common strategy is to disregard studies that do not satisfy certain methodological criteria, sometimes abusing 'criteria' for causality. Although such criteria are useful, UK statistician Austin Bradford Hill noted:

'All scientific work is incomplete... All scientific work is liable to be upset or modified by

⁽⁴⁾ When Joe Forman first observed the hole in the ozone layer using low technology instruments he could not believe his results as they conflicted with the satellite data. He returned to the Antarctica to observe the hole three times before he — under pressure from his funding sources — felt confident enough to report his findings. See the chapter on Halocarbons in the first volume of *Late lessons from early warnings* (EEA, 2001).

advancing knowledge. That does not confer upon us the freedom to ignore the knowledge we already have, or to postpone the action that it appears to demand at the given time' (Hill, 1965).

Despite Hill's prudent advice, some researchers may mistake the validity of their own conclusions for meticulousness in identifying presumed violations of the causal criteria or other validity requirements committed by their colleagues.

The overly sceptical focus on scientific methodology may lead to bias and narrow-mindedness. Thus, special interest groups have praised what they call 'sound science', which supports conclusions that are considered attractive (but, of course, is no more reliable than other research, and is sometimes actually less). In particular, so-called black-box epidemiology studies of health hazards have received harsh critique (Taubes, 1995). Some of that exaggerated critique is echoed in the chapters of the present volume.

Expert committees that advise national and international bodies are often tempted to express unreasonable critiques of research results and stress the preponderance of uncertainties. Such critiques may be considered appropriate for highly respected experts and is in accordance with their high methodological standards and unrelenting scepticism. However, a narrow focus on scientific methodology may be coupled with blindness to environmental degradation and social injustice. Not surprisingly, the strategy of criticising research methodologies has been vigorously explored by vested interests, often with the purpose of manufacturing doubt (Michaels, 2005; Michaels, 2008; Oreskes and Conway, 2010).

When a call for guidelines on 'Good Epidemiological Practice' was first promoted, it was at first embraced by researchers as a useful tool to stimulate high quality (and sound) science. However, strict interpretation of epidemiological rules could also be applied in order to disregard epidemiological findings that for other reasons were regarded as unwelcome. It turned out that the initiative originated with industry groups in order to disqualify unwelcome 'junk science' (as described in Chapter 7 on tobacco) (Ong and Glantz, 2001). The scientific

rigour that had been considered a prerequisite in the traditional science paradigm was now turned around and became an unrealistic requirement for repetitive, controlled studies that could furnish virtual statistical certainty⁽⁵⁾. Using strict criteria, unwanted results could then be criticised as junk and the uncertainties were then erroneously interpreted as an indication that no hazard was present.

26.7 Statistics and confidence limits

A key issue is the statistical data analysis. When analysing their results, researchers use statistical methods to determine whether the observed data were 'statistically significant', or whether they can be attributed to chance. The probability that their results are significant is usually expressed as p values, or probability values. The p was originally proposed by the UK statistician Ronald Fisher along with a limit of 5 % thought to be appropriate. This method allowed the researcher to identify findings that deviated significantly — unlikely due to random variation — so that the hypothesis that no difference was present would be rejected.

From its early application to agricultural plant breeding test designs, the 5 % limit has since been applied much more widely and has become almost sacrosanct amongst scientists from many disciplines. Using Fisher's p value limit allowed researchers to classify research findings that — when the p value was above 5 % — did not reliably support the 'null' hypothesis of no difference or no association, as the results could be due to random variation. Accordingly, the 'null' hypothesis could be rejected only when the p value was lower. A few studies and many anecdotes suggest that scientists place greater emphasis on results that have a p value of, say, 4.9 % than on results with a p value of 5.1 % (Holman et al., 2001). Statistically, there is no meaningful difference between outcomes with such similar p values. But if Fisher's proposed limit is applied in a strict sense beyond Fisher's own recommendations, then one set of results with a p value of 4.9 % would be interpreted as rejecting the hypothesis (hereby providing evidence of possible causality), whilst the other with a p value of 5.1 % would not refute the null hypothesis and would be considered non-informative.

⁽⁵⁾ An additional criterion often used was that only a 2-fold increased risk above background would be believable, e.g. from childhood leukaemia in residences close to power lines or from heart disease from environmental tobacco smoke. Apart from the much greater impact of a 2-fold increase in heart disease, there is no meaningful statistical difference between increases by a factor of 1.9 and 2.1, one of which would satisfy the criterion for a hazard, the other one not.

As commonly applied and interpreted, the p value is used mainly to determine the viability of a hypothesis. Although science in principle aims at falsifying hypotheses — since a definite proof cannot be obtained — it seems to make too little use of the data if we are just determining whether or not the p value is below 5 %. If the p value is high (well above 5 %), the results are rendered useless, as they failed the only criterion for success, namely to refute the hypothesis (although a hypothesis may theoretically be correct, even though the data deviate substantially from prediction). Sometimes, repeated attempts at falsifying a hypothesis fail, but a joint calculation (so-called meta-analysis) could result in an overall p value that perhaps finally reaches statistical significance with p less than 5 %, or some other specified level.

In most cases, the null hypothesis is that an exposure has no effect. Thus, in environmental research, the p value is used to test a null hypothesis that may be unrealistic or obviously wrong. This would seem to be a serious limitation. Would we ever be tempted to conclude that lead is not toxic, just because a small study has resulted in a p value that is greater than 5 %? Of course not. But the traditional use of the significance limit means that scientists are very reluctant to draw conclusions if the p value is 5.1 %.

The so-called frequentist tradition in statistics considers the data in isolation and evaluates them in regard to a theoretical null hypothesis, which may or may not be appropriate. Combined with a sacrosanct 5 % limit, the research results may not be as useful as they could be, and the conclusions could even be confusing and counterproductive (Goodman, 2008). The point is that we may be testing the wrong hypothesis and not making ample use of all of the available data. Thus, several case studies have shown that early warnings are often initially not statistically significant, such as the first IARC study of passive smoking, but nevertheless turned out to be robust.

Even if a study has reached statistical significance, this could still be due to chance. If we are conducting a large number of comparisons, then in all likelihood a small proportion of them could happen to be unusual and perhaps deviate from expectation at a statistically significant level. But such deviation is accidental and would be associated with a large number of comparisons. A common method is to adjust the p values using a procedure named after the Italian mathematician Bonferroni, thereby requiring p values to be significant only at lower values, the larger the number of comparisons. However, this technique, too, can also be used

erroneously to disregard an unwelcome study (Perneger, 1998).

The use of an alternative approach to frequentist statistics started back in the 18th century, when UK Reverend Thomas Bayes designed a formula that let the study results modify the prior probability of a hypothesis, thereby generating a posterior probability of the hypothesis based on the new evidence obtained (Greenland, 2008). Bayes allowed inclusion of any results, whether few or large-scale, and no matter the p value, to help modify our reliance on a hypothesis and to determine its updated plausibility. One could still focus on the null hypothesis, or perhaps rather the overall outcome of all previous studies. This way, each study would still be useful and would be utilised to modify and fine-tune the hypothesis under consideration. Although attractive, Bayesian statistics sometimes results in serious mathematical complications that limit their usefulness. Also, we may not have a good idea about the exact hypothesis under study, and a prior probability of that hypothesis may be impossible to obtain. Bayesian statistics has therefore been criticised for being subjective and overly laborious. Still, empirical use of Bayesian statistics is gaining support (McGrayne, 2011).

Some scientists and some scientific journals now reject the use of p values (Lang et al., 1998). But if we are to limit our reliance on p values, how can we best extract a robust statistical summary of a complex study? A key parameter will always be the point estimate of the average effect. But instead of calculating whether this estimate is 'significantly' different (p less than 5 %) from no effect, many researchers recommend using the confidence interval (Thompson, 1987). It represents the range of values within which 95 % of averages would fall if a large number of similar studies were conducted. In other words, given the point estimate and the calculated variability, the study would be in accordance with any hypothesis that postulated an effect within the confidence interval. If zero is included in the interval, then the results do not deviate significantly from the null hypothesis. However, they also do not deviate from many other hypotheses, some perhaps suggesting a serious effect. The upper confidence limit indicates how large an effect that would be in agreement with the data. In a precautionary setting, the upper limit would often represent a plausible worst case scenario that would serve as a useful basis when considering intervention.

The two studies illustrated in Figure 26.1 show the same average effect, though with different degrees

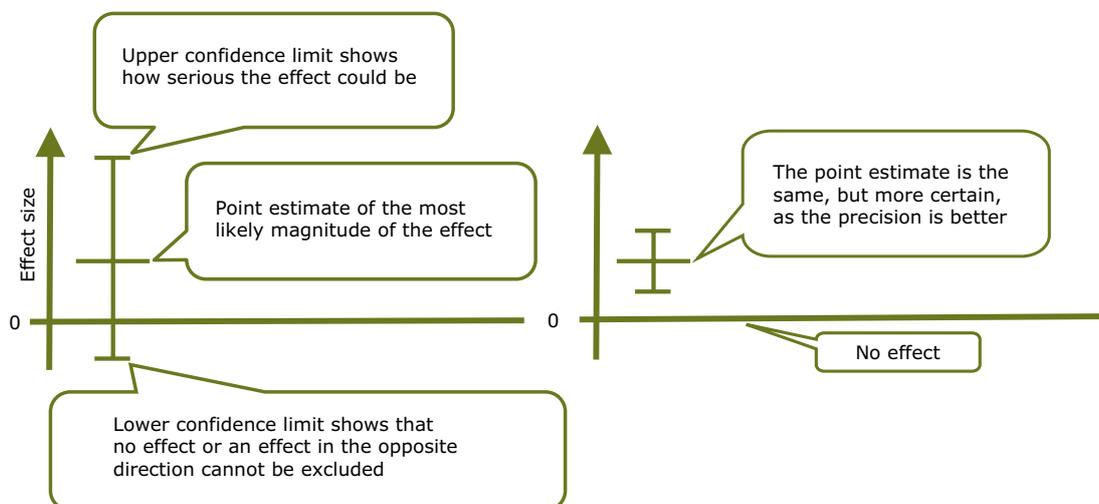
of certainty. The study on the right shows an effect that is statistically significant, as the no-effect hypothesis (zero effect) can be excluded. The study on the left has less precision, perhaps because it is smaller, and the point estimate does not deviate significantly from no effect (the null hypothesis). However, the upper confidence limit suggests that the study cannot exclude a large effect. In contrast, the significant study on the right would speak against the hazard being very large. Both of these perspectives are relevant, for both studies. A focus on the upper confidence limit would have the additional advantage that it would inspire larger studies with greater precision.

From a precautionary viewpoint, the use of confidence intervals is highly attractive. Instead of concluding that we are not sure that there is an effect at all, we can now also say that the results do not contradict an effect, and that it could possibly be up to a certain magnitude. If a study is large, and when results from two or more studies are combined, the confidence interval will narrow due to the decrease in statistical uncertainty. If a small study (like the one on the left) is in accordance with a potentially large effect, it would call for extended studies to explore whether such a serious hazard is indeed realistic. However, from the 'frequentist' viewpoint, the small study cannot reject the null hypothesis and would therefore not call for any further attention. Hence, the two perspectives differ substantially as to the interpretation of research, the conclusions, and the priorities for further information. Both are

useful, and a narrow focus on p values should be avoided (Stang et al., 2010).

The choice of statistical analysis is even more important in situations, where we do not have the option of calling for more studies. If a disease is serious but very rare, that number of subjects included will be small, and it may take a long time before enough information has been gathered in order to obtain a p value below 5%. Perhaps most dramatically, in regard to endangered species, it is simply not possible to sample sufficiently large materials to reach 'significance'. Thus, wildlife biologists some years ago concluded: 'At least part of the blame for the spectacular overexploitation of the great whales can be placed on scientists being unable to agree... In certain circumstances, a population might go extinct before a significant decline could be detected' (Taylor and Gerrodette, 1993). When the researchers examined the frequency and precision of recent monitoring efforts, they concluded that the percentage of precipitous declines that would not be detected as statistically significant would be between 72% and 90% for various whale species and 55% for polar bears. Thus, more than half of the world's polar bears and the great majority of the whales would have to disappear before current studies would be able to conclude that the decrease is 'significant' (based on a one-sided p value limit of 5%) (Taylor et al., 2007). Similarly, to the extent that monitoring and effect studies of environmental hazards are patchy, we are probably overlooking adverse effects, even those

Figure 26.1 The importance of confidence limits



Note: Two studies show the same average effect (horizontal line), but the vertical line suggests that the study on the left has a larger confidence interval and more uncertainty, so that it is both in accordance with no effect (it includes zero and is therefore not statistically significant), but it also cannot exclude a large effect. The study on the right shows the same effect, now statistically significant, but due to the greater precision, this study can exclude the presence of a large effect.

that are serious, simply because the information is uncertain (like the study on the left in Figure 26.1).

To avoid inconclusive results, researchers often carry out power analyses to determine the sensitivity of a proposed study, that is, the likelihood that the proposed study will lead to conclusions on the existence of a hazard of a certain relevant magnitude. If the protocol is not able to ascertain with any confidence the presence of an important risk, then the statistical power is insufficient (as in the monitoring of polar bears). Either the study would be a waste of time and should be disbanded, or the protocol should be expanded to allow sufficient power.

26.8 Bias in research

For the reasons listed above, the research results are often non-informative. Such inconclusive studies are sometimes called 'negative', although this term could suggest that an effect was in the direction opposite to expectation. Worse, such studies have sometimes been thought to represent 'no risk', rather than 'no information'. Such aspects of the traditional science paradigm involve inherent biases toward the null hypothesis. Based on the case studies in the present and the previous volume, Table 26.4 has been revised from previous compilations (Gee, 2009; Grandjean et al., 2004).

Most of the aspects listed in Table 26.4 have to do with the design of the research study and therefore

refer to the methodology, rather than the relevance of the research. So, in that respect, greater attention to methodology would be beneficial. However, the main problem is that even though the research results may be less informative than desired, the research may well contain information that is more relevant than the simple claim that the null hypothesis of no effect cannot be excluded. As illustrated in Figure 26.1, we need to ask: How large an effect can the study have overlooked? This question should also take into account the possible existence of vulnerable subgroups, long-term effects, and other issues that may have been ignored.

Two entries in the table refer to the possible existence of publication bias. It is quite likely that some science journals, and more often the mass media, prefer to publish alleged scares rather than to report that there is nothing to worry about (Ioannidis, 2008). But the bias may also be in the opposite direction (Oreskes, 2004). More importantly, our data on publication frequencies (Grandjean et al., 2011) suggest the opposite. The journals publish extensively on well-known chemicals, where new scares are rare, and only occasionally publish on the unknown and emerging environmental hazards which could possibly represent much scarier risks, given that so little attention is paid to them. So the few scares that catch occasional headlines should be interpreted in light of the overwhelming background of environmental hazards that are and have been ignored, some of which could well represent

Table 26.4 Key aspects of research likely to affect the outcome of a study, whether underestimating (false negative) or exaggerating (false positive) the possible existence of an environmental hazard

Methodological features and their main direction of error

Inadequate statistical power	
Lost cases and inadequate follow-up for long-term effects	
Exposure misclassification	
Insensitive or imprecise outcome measures	
Adjustment for confounders with better precision than the exposure	
Failure to adjust for confounder with effects in the opposite direction	False negative
Disregarding vulnerable subgroups	
5 % probability level to minimize risk of false positives (Type I error)	
20 % probability level to minimize risk of false negatives (Type II error)	
Pressure to avoid false alarm	
Incomplete adjustment for confounders with similar effects	
Post hoc hypothesis	False positive
Publication bias towards positive findings	

serious hazards. Chapter 2 on false positives shows that erroneous alarms are fairly rare.

In summary, the context of justification needs to be balanced with the context of application or, in other words, the quality of the research must be linked with its relevance. So the focus on methodology issues and the preoccupation with verification studies should not happen at the expense of providing evidence on issues of major environmental and social relevance. While polishing the same stone over and over again, we should not ignore all the other shingles and rocks where some scientific gems may yet be hiding.

26.9 The changing research paradigm

Science sociologist Robert K. Merton characterised traditional science by the acronym CUDOS, which stands for Communalism, Universalism, Disinterestedness, Originality and Scepticism. These traits are still valued and are prevalent in many scientific disciplines, but differences occur and the research paradigm within the environmental sciences is changing. All of these attributes for science, whether basic or applied, are needed to secure a meaningful and trustworthy research activity in society. However, the preoccupation with publication, credentials and funding that is common in academia today can lead to social apathy, thereby providing fertile ground for dependence on narrow interests that may include corporate money (Grandjean, 2008b).

Of particular note, Merton characterised ideal science as 'disinterested', but vested interests of whatever origin may make research less neutral and less reliable. Several case chapters in this volume describe how industries have withheld evidence, lambasted whistle-blowers and promoted research that supported the conclusions desired (Kurland, 2003). Especially when cover-up is included in such diversions, the result is that research loses credibility. Transparency in regard to conflicts of interest has been recommended, but complete elimination of financial ties may be the best way to secure trust-worthy research (Krimsky, 2003). Although conflicts of interest undoubtedly occur within academia at large, perhaps an additional problem is that the academic agenda is likely to differ from the priorities of regulatory agencies in environmental health.

Another science culture has developed, as research contracts or privately funded research have grown. They differ in several respects from the CUDOS ideal. The results may not necessarily be published

in scholarly journals (and would therefore be missed by our SciFinder searches). When the research is kept secret, it will not inspire further studies at public institutions. A particularly important chemical, bisphenol A (952 publications during 2000–2009), has enjoyed vast industrial popularity and became widely used in food packaging materials and beverage containers (see Chapter 10 on BPA). It was said to be safe at the very low exposures that consumers were likely to receive. However, after several decades of expanding use, independent research eventually uncovered evidence of health risks (Myers et al., 2009). The same pattern was seen with the perfluorinated compounds, where a major US producer for decades claimed that little would escape into the environment, and that essentially no toxicity occurred (Lindstrom et al., 2011). Only recently was it discovered that current exposures may be far from safe (Grandjean et al., 2012), but these chemicals have been disseminated into the global environment and cannot be recalled.

Physics professor John Ziman characterised the 'industrial' (or contracted) research as Proprietary, Local, Authoritarian, Commissioned, and Expert, thereby stressing that this activity builds on local expertise to reach specific goals. The same characteristics may apply to contract research carried out with public funding, but the initiator may not always be apparent. Thus, the Center for Indoor Air Research, the Electric Power Research Institute or the Chlorine Council may sound like charitable donors, rather than industry front groups. But they are in fact organisations funded by corporations with vested interests in the research outcome. However, the reader may be led to erroneously believe that the sponsored research reflects CUDOS values.

The source of funding will also affect the choice of study topics. Accordingly, comparatively little research is devoted to the risks associated with pesticide exposures and the advantages of alternative crop protection methods (Krimsky, 2003). Booster biocides (see Chapter 12), such as Diuron (389 links to articles in 2000–2009 in SciFinder) and Dichlofluanid (39 links), received only a little attention in independent research, some of them much less than the organotin compounds (see tributyltin in Table 26.1) that have been phased out. SciFinder also located only 133 links to Gaucho[®], the pesticide that endangered bee populations (see Chapter 16). As there are clear commercial interests in these compounds, the paucity of complementary academic research publications is unfortunate, although perhaps not surprising. Similarly, much less attention is paid to adverse

effects of new technology than to its advantages, although this has recently changed in regard to mobile telephony. Perhaps there is a parallel to physicians collaborating with the pharmaceutical industry in clinical trials of new drugs, which, with patent protection, will be capable of yielding great monetary returns. In contrast, older drugs no longer protected by patent are the subject of far less research, but may be as effective as modern drugs costing far more (Washburn, 2005).

Because evidence is the basis upon which the evaluation of risks must rely, researchers publishing results at odds with certain vested interests have become targets of criticism and intimidation with the aim of suppressing or throwing suspicion on unwelcome information about health risks. Perhaps the best known case involves Herbert Needleman, who supplied the first, weighty documentation of prevalent lead pollution damaging brain development (see Chapter 3 on lead). He was angrily persecuted and harassed with unfounded accusations of dishonesty (Needleman, 2000).

Disagreement usually focuses on the uncertainties and the scientific inference, not the choice of study topic. Harsh critique has sometimes been voiced, as have angry accusations of bias in differing interpretations of evidence (Gori, 1996). Research that has direct implications in regard to considerations of pollution abatement usually receives more wrath than reports on already recognised hazards. Perhaps this is another key as to why researchers favour well-known hazards.

In order to introduce dissent into the literature, possible strategies involve publication in trade magazines disguised as scientific journals. The best examples are *Indoor and Built Environment* (Tong et al., 2005) and *Regulatory Toxicology and Pharmacology* (Axelson et al., 2003). These journals tend to publish articles that contain conclusions favourable to the industrial sponsors, no matter their scientific weaknesses. This strategy is counter to the Precautionary Principle, as they argue for 'no risk' when the evidence is uncertain or non-informative. In addition to the tobacco industry, other examples include studies supported by the pharmaceutical industry, which are much more likely to conclude that a drug is safe and efficacious than studies conducted without such support (Jorgensen et al., 2006), but the same seems to happen in toxicology and environmental research (Myers et al., 2009).

As a consequence, public trust is abused by deceit. The purpose of research seeking truth is betrayed, when undisclosed ties taint the research and its conclusions.

Under such contentious conditions, researchers may choose to hedge their conclusions by incessant use of words, such as 'maybe', 'perhaps', 'in theory' and similar terms (Hyland, 1998). By softening the conclusions and avoiding attribution of specific causality, the researchers protect themselves against critique by appearing well-balanced, unassuming and even sceptical toward the implications of one's own findings⁽⁶⁾. However, this strategy has a downside. To the lay reader, who is not familiar with the traditions of scientific writing, the caveats and reservations may sound like the new results really do not prove anything, and that we are still left with the same uncertainty. To readers with a vested interest, the soft wording can be exploited through selective quotation and by emphasising real or alleged weaknesses (Grandjean, 2008a).

Because of the involvement of research funders, the industrialised (or contracted) science can be better characterized by the PLACE acronym (Ziman, 2000), although often posing like independent, basic research in accordance with CUDOS. If all research today earned CUDOS, no matter its funding, there would be little to worry about. But the weaknesses and biases outlined above suggest that PLACE needs to be supplemented by an additional research, one that better fits with the use of the PP in decision-making.

In this complementary paradigm, environmental research in support of PP-based decision-making would involve stake-holders and therefore become Participatory, rather than Communal or Proprietary as in the other paradigms. It would be Accessible, Transparent, Inventive and Open-minded. Although the various attributes may perhaps not be compared horizontally in Table 26.5, the PATIO characteristics would seem to fit better the research that is needed in a precautionary setting.

A key aspect is that, given the absence of final proof, an integrated evaluation must include uncertainty as a normal condition that needs to be explored and addressed, rather than minimised for the purpose of making research more efficient. An additional feature is the inclusion of the public in exploring how the uncertainty should affect

⁽⁶⁾ Please note how often I use the words 'may' and 'perhaps'. I do so, too, because I do not want to jump to conclusions and therefore present my case with understatement rather than the opposite.

Table 26.5 Main properties of research in three different settings

Academic (normal) CUDOS (a)	Industrial* PLACE (b)	Precautionary PATIO
Communalism	Proprietary	Participatory
Universalism	Local	Accessible
Disinterestedness	Authoritarian	Transparent
Originality	Commissioned	Inventive
Skepticism	Expert	Open-minded

Note: * 'Industrial' science is driven by private or other special interest and may violate some of the CUDOS norms in its pursuit of knowledge within fields of commercial or other defined interest, where public interest may be ignored.

Source: (a): Merton, 1973; (b): Ziman, 2000.

the decision-making. As discussed above, and in agreement with the PATIO paradigm, both study designs and the reporting of research results need to change.

In contrast to the traditional science paradigm, where replication is held as a key to supporting conclusions on causation, the PP does not inspire repetitive verification. If available, replication will be useful, but a hypothesis may well be plausible even in the (temporary) absence of supportive evidence. Given the enormous diversity and complexity of environmental hazards, one implication of the PP is that research is primarily needed to document the extent of uncertainty and, when possible, to narrow this uncertainty to better inform decision-making and eventually to support more precise risk assessments as a basis for interventions that will no longer need to be precautionary. But rather than fine-tuning risk assessments for individual hazards, the vastly incomplete information on most environmental chemicals makes research into uncertainty a very urgent need.

26.10 Precautionary science

As already discussed, the PP does not specifically demand testing of a null hypothesis that an exposure may be without a discernible effect. Rather, information is required whether a hazard could potentially be serious. This point of view should inspire new ways of planning, conducting, and reporting environmental research. So the research question outlined in the beginning of this chapter in accordance with traditional scientific paradigms now needs to be rephrased (Neutra, 2002):

'PP-based question on an environmental hazard: Are we sufficiently confident that this exposure to a potential hazard leads to doses of a

magnitude that can result in adverse effects that are serious enough to initiate transparent and democratic procedures to decide on appropriate intervention?'

We must pay closer attention to variability and uncertainty when determining their possible magnitude. Unfortunately, standard statistical methods assume that an exposure is measured without imprecision, which is usually not true, although this problem is generally ignored, thus resulting in underestimation of a hazard (Table 26.4). Assessment of the imprecision and its implications is therefore crucial. While uncertainties may be erroneously thought to cause exaggeration of alleged risks, most often the opposite is true (Grandjean, 2008b). The extent of uncertainties can be expressed in terms of confidence limits (Figure 26.1), but the impact may often need to be explored by using sensitivity analyses. One or more worst-case scenarios deserve as careful scrutiny as the null hypothesis: How serious could the effects be; how large an effect can be reasonably ruled out?

The research evidence must be considered in light of both strengths and weaknesses. While a methodological failure may weaken the support for a particular association, the mere occurrence of some scientific weakness does not prove the absence of a risk. Unfortunate and erroneous rejection of warning signals has occurred in the past because of presumed confounding or other biases and uncertainties. As illustrated by the case chapters in this volume, inconsistencies in some methodological aspect have been used to derail conclusions otherwise adopted by the scientific community. Likewise, statistical acceptance of the null hypothesis has sometimes been interpreted as proof of safety. Further, effects within normal variability have been considered irrelevant, although a population-wide shift in the distribution may represent substantial harm. Focus on average effects may also be misleading, as populations at risk

may suffer much greater harm that can be diluted by the results of non-vulnerable groups.

By acknowledging the limitations to the research evidence, a different point of view needs to be considered, i.e. what could possibly be known, given the type of evidence available? Studies e.g. with imprecise estimates of the causative exposure and insensitive and nonspecific outcome measures, are likely to detect only the most serious risks and therefore should be interpreted in light of the weight of such evidence. The fact that the null hypothesis could not be rejected with confidence may be irrelevant in such cases.

In general, all conclusions must be accepted as being provisional and temporary. While a study of often-cited publications in major medical journals found that many of the conclusions were subsequently found to be wrong (Ioannidis, 2008), this does not mean that environmental hazards are exaggerated. While accepting that a tentative conclusion based on preliminary evidence may later turn out to be wrong, public health responsibility may still demand that a serious threat be taken seriously, even though a final proof is not at hand. Any actions would then need to be adjusted later on, as more definite evidence emerges. At the same time, we should not ignore that the majority of environmental chemicals are poorly documented (Gilbert, 2011; NRC, 1984; US EPA, 1998), and ignoring such potential risks is likely to involve a very large number of false negative conclusions.

The bibliometric analyses that we conducted assume, as to regulatory agencies, that research results are published. But the science publication industry has undergone substantial change due to the electronic potentials of the internet for low-cost distribution. However, the costs of science publication need to be covered, just like the subscribers paid for the print journals. Thus, the majority of science articles are not accessible to the public on the internet, unless an access toll is paid (although access may be free after an embargo period of 6–12 months). Thus, while a citizen may view the science journal at a public library, the internet favours the academic world despite the stiff subscription charges. Some journals are open access, where the author pays a fee for quality control, processing and maintenance of the website, and the published article is then free for everybody to see. A growing number of journals now use this model. The European Commission recommends that articles arising from EC-funded research must be available after no more than 12 months. Other funding agencies, such as the Wellcome Trust, have as a

requirement that the results of sponsored research must be published with open access. Groups of universities, e.g. in the Netherlands, have launched a repository, where their research publications can be accessed by anyone. So in regard to the Participatory aspect of the PATIO paradigm, access to information is improving.

Even preliminary data can facilitate PP-based decision-making. While early findings may provide only tentative conclusions, they can later be included in potential meta-analyses or provide a starting point for follow-up studies. This potential assumes that the data from previous studies are available, and that may not be true. Trade secrets may allegedly be involved, and numerous cases have occurred with suppression of information and withholding of evidence (Kurland, 2003). Some public funding agencies now demand that a data-sharing strategy be worked out for major projects, so that other researchers can carry out additional analyses, including meta-analyses. But there is also a risk that such further analyses are not entirely benevolent (Pearce and Smith, 2011). Hostile analyses have occurred, thus making researchers wary with whom they share their raw data.

Given the discussion on coverage of environmental hazards, attention to the needs of regulatory agencies, traditions of science publication and the impact of other players, we can now attempt to answer the four questions posed in the beginning. Stakeholder involvement, innovation, openness and transparency should become new, important assets in environmental research to serve better as documentation and inspiration for PP-based decision-making.

Ways to improve scientific evidence for robust and precautionary decisions on environmental hazards:

- 1) The choice of research topic should involve stakeholders and consider the societal needs for information on poorly known hazards;
- 2) The research should be innovative and complementary with the aim of extending current knowledge, rather than repetitive for verification purposes;
- 3) The findings should be communicated in such a way as to facilitate judgements concerning the possible magnitude of suspected environmental hazards;
- 4) The research should be openly available and independent of vested interests.

As argued elsewhere in this volume, science does not provide a prescription for the right decisions on environmental hazards. The emphasis on research will be different for those whose first priority is scientific exactitude and those who focus on making policy in the context of environmental protection and public health. When a precautionary perspective mandates action to prevent foreseeable harms, the evidence does not have to meet the most rigorous demands of science. However, world views, political and other preferences, technical and economic feasibility, and alternative options are crucial for decision-making. As illustrated by the case studies in both volumes of *Late lessons from early warnings*, science does not have a good track record for supporting decisions on improving environmental health. This chapter has highlighted some opportunities for environmental research to provide more relevant results, interpretation, and conclusions for prudent and timely decisions on environmental hazards.

References

- Axelsson, O., Balbus, J.M., Cohen, G., Davis, D., Donnay, A., Doolittle, R., Duran, B.M., Egilman, D., Epstein, S.S., Goldman, L., Grandjean, P., Hansen, E.S., Heltné, P., Huff, J., Infante, P., Jacobson, M.F., Joshi, T.K., LaDou, J., Landrigan, P.J., Lee, P.R., Lockwood, A.H., MacGregor, G., Melnick, R., Messing, K., Needleman, H., Ozonoff, D., Ravanesi, B., Richter, E.D., Sass, J., Schubert, D., Suzuki, D., Teitelbaum, D., Temple, N.J., Terracini, B., Thompson, A., Tickner, J., Tomatis, L., Upton, A.C., Whyatt, R.M., Wigmore, D., Wilson, T., Wing, S.B. and Sharpe VA., 2003, 'Re: Regulatory Toxicology and Pharmacology', *Int. J. Occup. Environ. Health*, (9/4) 386–389; author reply 389–390.
- Dann, A.B., Hontela, A., 2011, 'Triclosan: environmental exposure, toxicity and mechanisms of action', *J. Appl. Toxicol.*, (31/4) 285–311.
- EC, 2001, Communication from the Commission to the European Parliament, European Economic and Social Committee and the Committee of the Regions, 'Horizon 2020 — The Framework Programme for Research and Innovation', (COM(2011) 808 final of 30 November 2011).
- EEA, 2001, *Late lessons from early warnings: the precautionary principle 1896–2000*, Environmental issue report No 22, European Environment Agency.
- Eurostat, 2011, *R & D expenditure* (http://epp.eurostat.ec.europa.eu/statistics_explained/index.php/R_%26_D_expenditure) accessed 24 March 2012.
- Gee, D., 2009, 'Late Lessons from Early Warnings: Towards realism and precaution with EMF?', *Pathophysiology*, (16/2–3) 217–231.
- Gilbert, N., 2011, 'Data gaps threaten chemical safety law', *Nature*, (475/7355) 150–151.
- Goldstein, B.D., Carruth, R.S., 2004, 'Implications of the Precautionary Principle: is it a threat to science?', *Int. J. Occup. Med. Environ. Health*, (17/1) 153–161.
- Goodman, S., 2008, 'A dirty dozen: twelve p-value misconceptions', *Seminars in hematology*, (45/3) 135–140.
- Gori, G.B., 1996, 'Science, imaginable risks, and public policy: anatomy of a mirage', *Regul. Toxicol. Pharmacol.*, (23/3) 304–311.
- Grandjean, P., 2008a, 'Late insights into early origins of disease', *Basic Clin. Pharmacol Toxicol.*, (102/2) 94–99.
- Grandjean, P., 2008b, 'Seven deadly sins of environmental epidemiology and the virtues of precaution', *Epidemiology*, (19/1) 158–162.
- Grandjean, P., Andersen, E., Budtz Jørgensen, E., Nielsen, F., Mølbak, K., Weihe, P. and Heilmann, C., 2012, 'Serum Vaccine Antibody Concentrations in Children Exposed to Perfluorinated Compounds', *JAMA*, (307/4) 391–397.
- Grandjean, P., Bailar, J.C., Gee, D., Needleman, H.L., Ozonoff, D.M., Richter, E., Sofritti, M. and Soskolne, C.L., 2004, 'Implications of the Precautionary Principle in research and policy-making', *Am. J. Ind. Med.*, (45/4) 382–385.
- Grandjean, P., Bellinger, D., Bergman, A., Cordier, S., Davey Smith, G., Eskenazi, B., Gee, D., Gray, K., Hanson, M., van den Hazel, P., Heindel, J.J., Heinzow, B., Hertz-Picciotto, I., Hu, H., Huang, T.T., Jensen, T.K., Landrigan, P.J., McMillen, I.C., Murata, K., Ritz, B., Schoeters, G., Skakkebaek, N.E., Skerfving, S. and Weihe, P., 2008, 'The faroes statement: human health effects of developmental exposure to chemicals in our environment', *Basic Clin. Pharmacol. Toxicol.*, (102/2) 73–75.
- Grandjean, P., Eriksen, M.L., Ellegaard, O., Wallin, J.A., 2011, 'The Matthew effect in environmental science publication: A bibliometric analysis of

- chemical substances in journal articles', *Environ. Health*, (10/96) 96.
- Greenland, S., 2008, 'Bayesian interpretation and analysis of research results', *Seminars in hematology*, (45/3) 141–149.
- Hill, A.B., 1965, 'The environment and disease: Association or causation?', *Proc. R. Soc. Med.*, (58) 295–300.
- Holman, C.D., Arnold-Reed, D.E., de Klerk, N., McComb, C., English, D.R., 2001, 'A psychometric experiment in causal inference to estimate evidential weights used by epidemiologists', *Epidemiology*, (12/2) 246–255.
- Hyland, K., 1998, *Hedging in Scientific Research Articles*. Amsterdam, the Netherlands: John Benjamins.
- Ioannidis, J.P., 2008, 'Why most discovered true associations are inflated', *Epidemiology*, (19/5) 640–648.
- Jorgensen, A.W., Hilden, J., Gotzsche, P.C., 2006, 'Cochrane reviews compared with industry supported meta-analyses and other meta-analyses of the same drugs: systematic review', *BMJ*, (333/7572) 782.
- Krimsky, S., 2003, *Science in the private interest*. Lanham, MD: Rowman & Littlefield.
- Kurland, J., 2003, 'The heart of the precautionary principle in democracy', *Public Health Rep.*, (117/6) 498–500.
- Lang, J.M., Rothman, K.J., Cann, C.I., 1998, 'That confounded P-value', *Epidemiology*, (9/1) 7–8.
- Lindstrom, A.B., Strynar, M.J., Libelo, E.L., 2011, 'Polyfluorinated compounds: past, present, and future', *Environ. Sci. Technol.*, (45/19) 7954–7961.
- McGrayne, S.B., 2011, *The theory that would not die*. New Haven: Yale University Press.
- Merton, R.K., 1968, 'The Matthew effect in science. The reward and communication systems of science are considered', *Science*, (159/810) 56–63.
- Michaels, D., 2005, 'Doubt is their product', *Sci. Am.*, (292/6) 96–101.
- Michaels, D., 2008, *Doubt is Their Product: How Industry's Assault on Science Threatens Your Health*. New York: Oxford University Press.
- Myers, J.P., vom Saal, F.S., Akingbemi, B.T., Arizono, K., Belcher, S., Colborn, T., Chahoud, I., Crain, D.A., Farabollini, F., Guillette, L.J., Hassold, T., Ho, S.M., Hunt, P.A., Iguchi, T., Jobling, S., Kanno, J., Laufer, H., Marcus, M., McLachlan, J.A., Nadal, A., Oehlmann, J., Olea, N., Palanza, P., Parmigiani, S., Rubin, B.S., Schoenfelder, G., Sonnenschein, C., Soto, A.M., Talsness, C.E., Taylor, J.A., Vandenberg, L.N., Vandenberg, J.G., Vogel, S., Watson, C.S., Welshons, W.V. and Zoeller, R.T., 2009, 'Why public health agencies cannot depend on good laboratory practices as a criterion for selecting data: the case of bisphenol A', *Environ. Health Perspect.*, (117/3) 309–315.
- NRC, 1984, *Toxicity Testing: Strategies to Determine Needs and Priorities*, National Research Council.
- Needleman, H.L., 2000, 'The removal of lead from gasoline: historical and personal reflections', *Environ. Res.*, (84/1) 20–35.
- Neutra, R., DelPizzo, P., Geraldine, M., 2002, *An Evaluation of the Possible Risks from Electric and Magnetic Fields (EMFs) from Power Lines, Internal Wiring, Electrical Occupations, and Appliances*.
- Ong, E.K., Glantz, S.A., 2001, 'Constructing "sound science" and "good epidemiology": Tobacco, lawyers, and public relations firms', *Am. J. Public Health*, (91/11) 1749–1757.
- Oreskes, N., 2004, 'Beyond the ivory tower. The scientific consensus on climate change', *Science*, (306/5702) 1686.
- Oreskes, N., Conway, E.M., 2010, *Merchants of doubt: how a handful of scientists obscured the truth on issues from tobacco smoke to global warming*. 1st US. ed. New York: Bloomsbury Press.
- Pearce, N., Smith, A.H., 2011, 'Data sharing: not as simple as it seems', *Environ. Health*, (10) 107.
- Perneger, T.V., 1998, 'What's wrong with Bonferroni adjustments', *BMJ*, (316/7139) 1236–1238.
- Schwartz, S., Carpenter, K.M., 1999, 'The right answer for the wrong question: consequences of type III error for public health research', *Am. J. Public Health*, (89/8) 1175–1180.
- Stang, A., Poole, C., Kuss, O., 2010, 'The ongoing tyranny of statistical significance testing in biomedical research', *European journal of epidemiology*, (25/4) 225–230.

Taubes, G., 1995, 'Epidemiology faces its limits', *Science*, (269/5221) 164–169.

Taylor, B.L., Gerrodette, T., 1993, 'The Uses of Statistical Power in Conservation Biology — the Vaquita and Northern Spotted Owl', *Conservation Biology*, (7/3) 489–500.

Taylor, B.L., Martinez, M., Gerrodette, T., Barlow, J., Hrovat, Y.N., 2007, 'Lessons from monitoring trends in abundance of marine mammals', *Marine Mammal Science*, (23/1) 157–175.

Thompson, W.D., 1987, 'Statistical criteria in the interpretation of epidemiologic data', *Am. J. Public Health*, (77/2) 191–194.

Tong, E.K., England, L., Glantz, S.A., 2005, 'Changing conclusions on secondhand smoke in a sudden infant death syndrome review funded by the tobacco industry', *Pediatrics*, (115/3) e356–e366.

US EPA, 1998, *Chemical Hazard Data Availability Study. What do we really know about the safety of high production volume chemicals?* EPA's 1998 baseline of hazard information that is readily available to the public, U.S. Environmental Protection Agency.

US EPA, 2009, *Risk-Based Prioritization (RBP) Decisions Summary*, U.S. Environmental Protection Agency.

van den Hove, S., McGlade, J., Depledge, M.H., 2011, 'EU innovation must benefit society', *Nature*, (474/7350) 161.

Washburn, J., 2005, University, Inc. New York, Basic Books.

Ziman, J.M., 2000, *Real Science: what it is, and what it means.*, Cambridge: Cambridge University Press.