1

The scientific appraisal of hazardous substances in the environment

Olwenn V. Martin[1*] and Jane A. Plant[2]

[1]*Institute for the Environment, Brunel University, Kingston Lane, Uxbridge, Middlesex, UB8 3PH*
[2]*Centre for Environmental Policy and Department of Earth Science and Engineering, Imperial College London, Prince Consort Road, London SW7 2AZ*
[*]*Corresponding author, email olwenn.martin@brunel.ac.uk*

1.1 Introduction

This book describes a wide range of non-living toxins that are present in the environment and are potentially harmful to human health or the environment. It covers both organic and inorganic substances, natural as well as manufactured substances. It deals mostly with substances that are hazardous because of their chemical properties, but also includes some where the hazard derives from their physical properties, for example, particulates and nanoparticles. This chapter summarises some important concepts of basic toxicology and environmental epidemiology relevant to an understanding of the possible effects of pollutants in the environment.

A common misconception is that chemicals made by nature are intrinsically good and, conversely, those manufactured by man are bad (Ottoboni, 1991). However, there are many examples of toxic compounds produced by algae or other micro-organisms, venomous animals and plants. There are even examples of environmental harm resulting from the presence of relatively benign natural compounds, either in unexpected places or in unexpected quantities. It is therefore of prime importance to define what is meant by 'chemical' when referring to chemical hazards in this chapter and the rest of this book. The correct term for a chemical compound to which an organism may be exposed, whether of natural or synthetic origins, is xenobiotic, i.e. a substance foreign to an organism (the term has also been used for transplants). A xenobiotic can be defined as a chemical which is found in an organism but which is not normally produced or expected to be present in it. It can also cover substances that are present in much higher concentrations than are usual.

1.2 Fundamental concepts of toxicology

Toxicology is the science of poisons. A poison is commonly defined as 'any substance that can cause an adverse effect as a result of a physicochemical interaction with living tissue' (Duffus, 2006). The use of poisons is as old as the human race, as a method of hunting or warfare as well as murder, suicide or execution. The evolution of this scientific discipline cannot be separated from the evolution of pharmacology, or the science of cures. Theophrastus Phillippus Aureolus Bombastus von Hohenheim, more commonly known as Paracelsus (1493–1541), a physician contemporary of Copernicus, Martin Luther and da Vinci, is widely considered as the father of toxicology. He challenged the ancient concepts of medicine based on the balance of the four humours (blood, phlegm, yellow and black bile) associated with the four elements and believed that illness occurred when an organ failed and poisons accumulated. This use of chemistry and chemical analogies was particularly offensive to the contemporary medical establishment. He is famously credited with the quotation that still underlies present-day toxicology.

Pollutants, Human Health and the Environment: A Risk Based Approach, First Edition. Edited by Jane A. Plant, Nikolaos Voulvoulis and K. Vala Ragnarsdottir.

> 'All substances are poisons; there is none which is not a poison. The right dose differentiates a poison from a remedy.'
> *Paracelsus*

In other words, all substances are potential poisons, since all can cause injury or death following excessive exposure. Conversely, this statement implies that all chemicals can be used safely if handled with appropriate precautions and exposure is kept below a defined limit below which risk is considered tolerable (Duffus, 2006). The concepts of tolerable risk and adverse effect illustrate the value judgements embedded in an otherwise scientific discipline relying on observable, measurable empirical evidence. What is considered abnormal or undesirable is dictated by society rather than science. Any change from the normal state is not necessarily an adverse effect even if statistically significant. An effect may be considered harmful if it causes damage, irreversible change or increased susceptibility to other stresses, including infectious disease. The stage of development or state of health of the organism may also have an influence on the degree of harm.

1.2.1 Routes of exposure

Toxicity will vary depending on the route of exposure. There are three routes by which exposure to environmental contaminants may occur:

- Ingestion.
- Inhalation.
- Skin adsorption.

In addition, direct injection may be used in testing for toxicity. Toxic and pharmaceutical agents generally produce the most rapid response and greatest effect when given intravenously, directly into the bloodstream. A descending order of effectiveness for environmental exposure routes would be inhalation, ingestion and skin adsorption.

Oral toxicity is most relevant for substances that might be ingested with food or drinks. It could be argued that this is generally under an individual's control, but people often don't know what chemicals there are in their food or water and are not well informed about the current state of knowledge about their harmful effects.

Inhalation of gases, vapours, dusts and other airborne particles is generally involuntarily (with the notable exception of smoking). The destination of inhaled solid particles depends upon their size and shape. In general, the smaller the particle, the further into the respiratory tract it can go. A large proportion of airborne particles breathed through the mouth or cleared by the cilia of the lungs can enter the gut.

Dermal exposure generally requires direct and prolonged contact with the skin. The skin acts as a very effective barrier against many external toxicants, but because of its large surface area (1.5–2 m^2) some of the many and diverse substances it comes in contact with may elicit topical or systemic effects (Williams and Roberts, 2000). If dermal exposure is often most relevant in occupational settings, it may nonetheless be pertinent in relation to bathing waters (ingestion is also an important route of exposure in this context). The use of cosmetics raises the same questions regarding the adequate communication of current knowledge about potential effects as those related to food.

1.2.2 Duration of exposure

The toxic response will also depend on the duration and frequency of exposure. The effect of a single dose of a chemical may be severe whilst the same total dose given at several intervals may have little or no effect: the effect of drinking four beers in one evening, for example, is very different from that of drinking four beers in four days. Exposure duration is generally divided into four broad categories: acute, sub-acute, sub-chronic and chronic. Acute exposure to a chemical usually refers to a single exposure event or repeated exposures over a duration of less than 24 hours. Sub-acute exposure to a chemical refers to repeated exposures for 1 month or less, sub-chronic exposure to continuous or repeated exposures for 1 to 3 months or approximately 10 per cent of the lifetime of an experimental species, and chronic exposure to continuous or repeated exposures for more than 3 months, usually 6 months to 2 years in rodents (Eaton and Klaassen, 2001). Chronic exposure studies are designed to assess the cumulative toxicity of chemicals with potential lifetime exposure in humans. The same terms are used in real-life situations, though it is generally very difficult to ascertain with any certainty the frequency and duration of exposure.

For acute effects, the time component of the dose is not important, as it is the high dose that is responsible for the effects. However, the fact that acute exposure to agents that are rapidly absorbed is likely to induce immediate toxic effects does not rule out the possibility of delayed effects, and these are not necessarily similar to those associated with chronic exposure (e.g. latency between the onset of certain cancers and exposure to a carcinogenic substance). The effect of exposure to a toxic agent may depend on the timing of exposure. In other words, long-term effects as a result of exposure to a toxic agent during a critically sensitive stage of development may differ markedly from those seen if an adult organism is exposed to the same substance. Acute effects are almost always the result of accidents, or, less commonly, criminal poisoning or self-poisoning (suicide). Chronic exposure to a toxic agent is generally associated with long-term low-level chronic effects, but this does not preclude the possibility of some immediate (acute) effects after each administration. These concepts are closely related to the mechanisms of metabolic degradation and excretion of ingested substances, as illustrated in Figure 1.1.

1.2.3 Mechanisms of toxicity

The interaction of a foreign compound with a biological system is two-fold: there is the effect of the organism on the compound

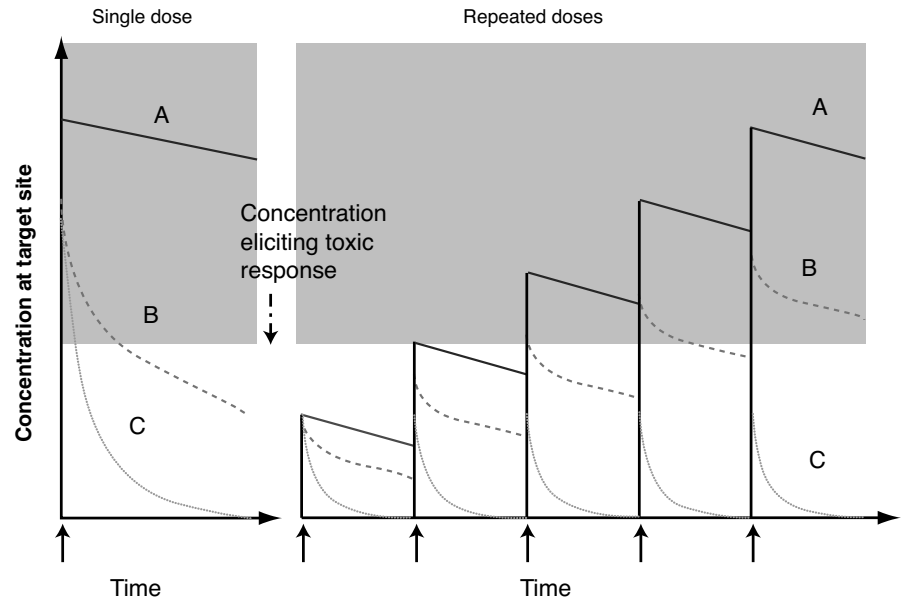


Figure 1.1 Relationship between dose and concentration at the target site under different conditions of dose frequency and elimination rate (reproduced from Eaton and Klaassen, 2001). *Line A.* Chemical with very slow elimination. *Line B.* Chemical with a rate of elimination slower than frequency of dosing. *Line C.* Rate of elimination faster than the dosing frequency. Shaded area represents the concentration range at the target site exhibiting a toxic response

(toxicokinetics) and the effect of the compound on the organism (toxicodynamics).

Toxicokinetics relate to the delivery of the compound to its site of action, including absorption (transfer from the site of administration into the general circulation), distribution (via the general circulation into and out of the tissues), and elimination (from general circulation by metabolism or excretion). The target tissue refers to the tissue where a toxicant exerts its effect, and is not necessarily where the concentration of the toxic substance is highest. Many halogenated compounds such as polychlorinated biphenyls (PCBs) or flame retardants such as polybrominated diphenyl ethers (PBDEs) are known to bioaccumulate in body fat. Whether such sequestration processes are actually protective to the individual organisms (by lowering the concentration of the toxicant at the site of action) is not clear (O'Flaherty, 2000). In an ecological context however, such bioaccumulation may serve as an indirect route of exposure for organisms at higher trophic levels, thereby potentially contributing to biomagnification through the food chain.

Absorption of any compound that has not been intravenously injected will entail transfer across membrane barriers before it reaches the systemic circulation, and the efficiency of absorption processes is highly dependent on the route of exposure.

It is also important to note that distribution and elimination, although often considered separately, take place simultaneously. Elimination itself comprises two kinds of processes, excretion and biotransformation, which also take place simultaneously. Elimination and distribution are not independent of each other, as effective elimination of a compound will prevent its distribution in peripheral tissues, whilst, conversely, wide distribution of a compound will impede its excretion (O'Flaherty, 2000). Kinetic models attempt to predict the concentration of a toxicant at the target site from the administered dose. The ultimate toxicant, i.e. the chemical species that induces structural or functional alterations resulting in toxicity, may be the compound administered (parent compound), but it can also be a metabolite of the parent compound generated by biotransformation processes, i.e. toxication rather than detoxication (Timbrell, 2000; Gregus and Klaassen, 2001). The liver and kidneys are the most important excretory organs for non-volatile substances, whilst the lungs excrete volatile compounds and gases. Other routes of excretion include the skin, hair, sweat, nails and milk. Milk may be a major route of excretion for lipophilic chemicals due to its high fat content (O'Flaherty, 2000).

Toxicodynamics is the study of toxic response at the site of action, including the reactions with and binding to cell constituents, and the biochemical and physiological consequences of these actions. Such consequences may therefore be manifested and observed at the molecular or cellular levels, at the target organ or on the whole organism. Therefore, although toxic responses have a biochemical basis, the study of toxic response is generally subdivided, either depending on the organ on which toxicity is observed, including hepatotoxicity (liver), nephrotoxicity (kidney), neurotoxicity (nervous system), pulmonotoxicity (lung) or depending on the type of toxic response, including teratogenicity (abnormalities of physiological development), immunotoxicity (immune system impairment), mutagenicity (damage of genetic material), carcinogenicity (cancer causation

or promotion). The choice of the toxicity endpoint to observe in experimental toxicity testing is therefore of critical importance. In recent years, rapid advances of biochemical sciences and technology have resulted in the development of bioassay techniques that can contribute invaluable information regarding toxicity mechanisms at the cellular and molecular level. However, the extrapolation of such information to predict effects in an intact organism for the purpose of risk assessment is still in its infancy (Gundert-Remy *et al.*, 2005).

1.2.4 Dose–response relationships

The theory of dose–response relationships is based on the assumptions that (1) the activity of a substance depends on the dose an organism is exposed to (i.e. all substances are inactive below a certain threshold and active over that threshold), and (2) dose–response relationships are monotonic (i.e. the response rises with the dose). Toxicity may be detected either as an all-or-nothing phenomenon such as the death of the organism or as a graded response such as the hypertrophy of a specific organ. Dose–response relationships for all-or-nothing (quantal) responses are typically S-shaped and this reflects the fact that sensitivity of individuals in a population generally exhibits a normal or Gaussian distribution (bell-shaped curve). When plotted as a cumulative frequency distribution, a sigmoid dose–response curve is observed (Figure 1.2).

Studying dose response and developing dose–response models are central to determining 'safe' and 'hazardous' levels.

The simplest measure of toxicity is lethality, and determination of the median lethal dose, the LD_{50} is usually the first toxicological test performed with new substances. The LD_{50} is the dose at which a substance is expected to cause the death of half of the experimental animals and it is derived statistically from dose–response curves (Eaton and Klaassen, 2001). LD_{50} values are the standard for comparison of acute toxicity between chemical compounds and between species. Some values are given in Table 1.1. It is important to note that the higher the LD_{50}, the less toxic the substance.

Table 1.1 Acute LD_{50} of some well-known substances (adapted from Eaton and Klaassen, 2001)

Agent	LD_{50}, mg/kg body weight
Ethyl alcohol	10,000
Sodium chloride	4,000
Ferrous sulphate	1,500
Morphine sulphate	900
Phenobarbital sodium	150
Strychnine sulphate	2
Nicotine	1
Dioxin (TCDD)	0.001
Botulimum toxin	0.00001

Similarly, the EC_{50}, the median effective dose, is the quantity of the chemical that is estimated to have an effect in 50 per cent of the organisms. However, median doses alone are not very informative, as they do not convey any information on the shape of the dose–response curve. This is best illustrated by Figure 1.3. While toxicant A seems (always) more toxic than toxicant B on the basis of its lower LD_{50}, toxicant B will start affecting organisms at lower doses (lower threshold) while the steeper slope for the dose–response curve for toxicant A means that once individuals become overexposed (exceed the threshold dose) the increase in response occurs over much smaller increments in dose.

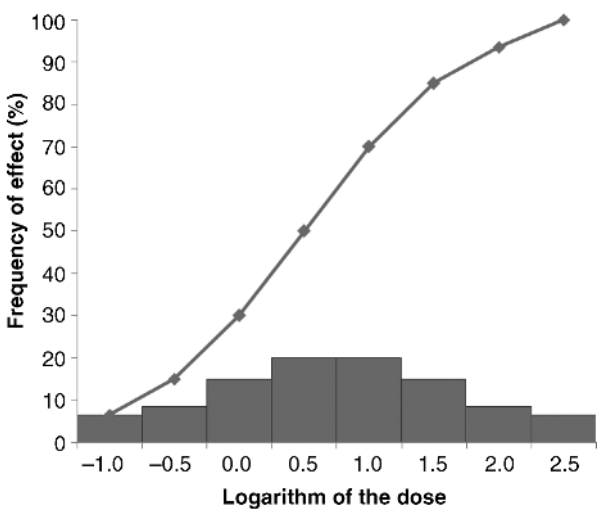


Figure 1.2 Quantal dose–response relationship. Bar chart shows the proportion of individuals affected at each dose and the line shows the cumulative frequency

1.2.4.1 Low dose responses

The classic paradigm for extrapolating dose–response relationships at low doses is based on the concept of threshold for non-carcinogens, whereas for carcinogens it is assumed that there is no threshold and a linear relationship is hypothesised (Figures 1.4 and 1.5).

The NOAEL (No Observed Adverse Effect Level) is the exposure level at which there is no statistically or biologically significant increase in the frequency or severity of adverse effects between exposed population and its appropriate control. The NOEL for the most sensitive test species and the most sensitive indicator of toxicity is usually employed for regulatory purposes. The LOAEL (Lowest Observed Adverse Effect Level) is the lowest exposure level at which there is a statistically or biologically significant increase in the frequency or severity of adverse effects between exposed population and its appropriate control. The main criticism of NOAEL and LOAEL is that they are dependent on study design, i.e. the dose groups selected and the number of individuals in each group. Statistical methods of deriving the concentration that produces a specific effect EC_x,

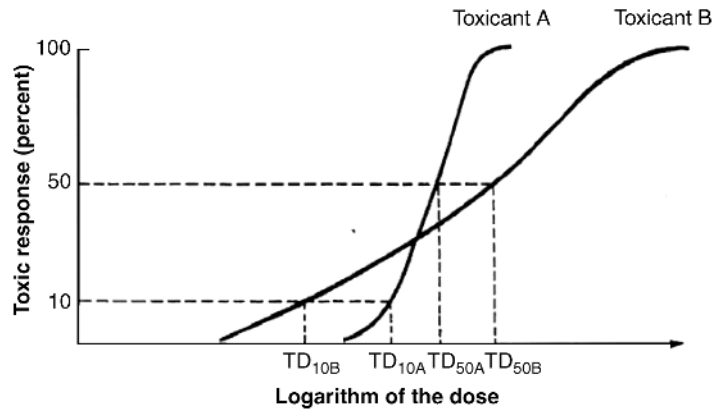


Figure 1.3 Importance of the dose–response relationship (Reproduced from *Principles of Toxicology: environmental and industrial applications*, R. C. James, © 2000 by John Wiley & Sons, with permission from John Wiley & Sons Inc.)

or a benchmark dose (BMD), the statistical lower confidence limit on the dose that produces a defined response (the benchmark response or BMR), are increasingly preferred.

To understand the risk that environmental contaminants pose to human health requires the extrapolation of limited data from animal experimental studies to the low doses critically encountered in the environment. Such extrapolation of dose–response relationships at low doses is the source of much controversy. Recent advances in the statistical analysis of very large populations exposed to ambient concentrations of environmental pollutants have not observed thresholds for cancer or non-cancer outcomes (White *et al.*,2009). The actions of chemical agents are triggered by complex molecular and cellular events that may lead to cancer and non-cancer outcomes in an organism. These processes may be linear or non-linear at an individual level. A thorough understanding of critical steps in a toxic process may help refine current assumptions about thresholds (Boobis *et al.*, 2009). The dose–response curve, however, describes the response or variation in sensitivity of a population. Biological and statistical attributes such as population variability, additivity to pre-existing conditions or diseases induced at background exposure will tend to smooth and linearise the dose–response relationship, obscuring individual thresholds.

1.2.4.2 Hormesis

Dose–response relationships for substances that are essential for normal physiological function and survival are actually U-shaped. At very low doses, adverse effects are observed due to a deficiency. As the dose of such an essential nutrient is increased, the adverse effect is no longer detected and the organism can function normally in a state of homeostasis. Abnormally high doses, however, can give rise to a toxic

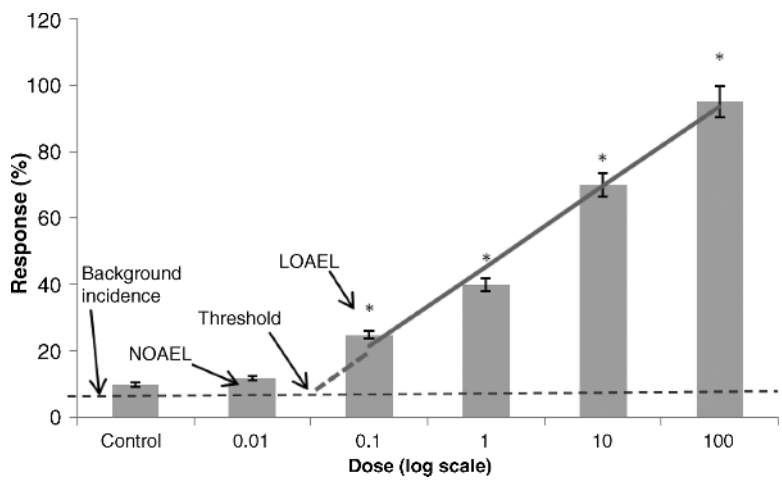


Figure 1.4 Extrapolation of the dose–response relationship at low doses for non-carcinogens (threshold concept). Vertical lines represent the standard error and * denotes statistical significance

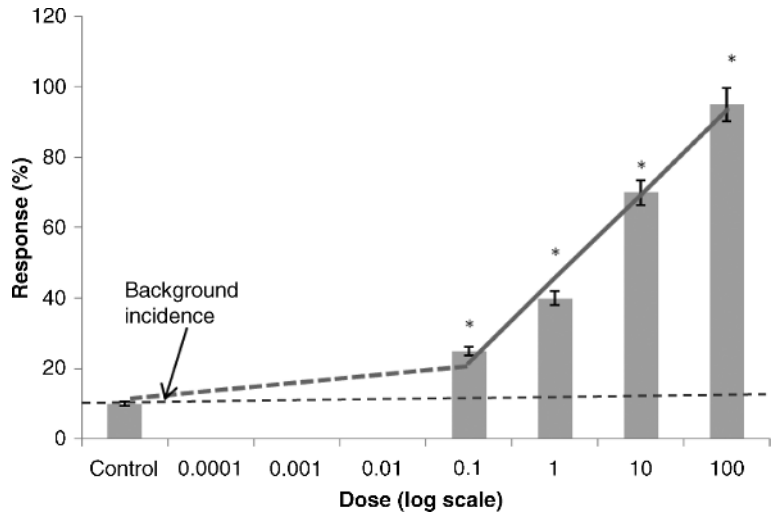


Figure 1.5 Extrapolation of the dose–response relationship at low doses for carcinogens (assumes no threshold). Vertical lines represent the standard error and * denotes statistical significance

response. This response may be qualitatively different and the toxic endpoints measured at very low and very high doses are not necessarily the same.

There is evidence that non-essential substances may also impart an effect at very low doses (Figure 1.6). Some authors have argued that hormesis, defined as beneficial stimulatory effects of toxins at low doses, ought to be the default assumption in the risk assessment of toxic substances (Calabrese and Baldwin, 2003). Whether such low dose effects should be considered stimulatory or beneficial is controversial. Further, potential implications of the concept of hormesis for the risk management of the combinations of the wide variety of environmental contaminants present at low doses that individuals with variable sensitivity may be exposed to are at best unclear.

1.2.5 Chemical interactions

In regulatory hazard assessment, chemical hazards are typically considered on a compound-by-compound basis, the possibility of chemical interactions being accounted for by the use of safety or uncertainty factors. Mixture effects still represent a challenge for the risk management of chemicals in the environment, as the presence of one chemical may alter the response to another chemical. The simplest interaction is *additivity*: the effect of two or more chemicals acting together is equivalent to the sum of the effects of each chemical in the mixture when acting independently. *Synergism* is more complex and describes a situation when the presence of both chemicals causes an effect that is greater than the sum of their effects when acting alone. In *potentiation,* a substance that does not produce specific toxicity on its own increases the toxicity of another substance when both are present. *Antagonism* is the principle upon which antidotes

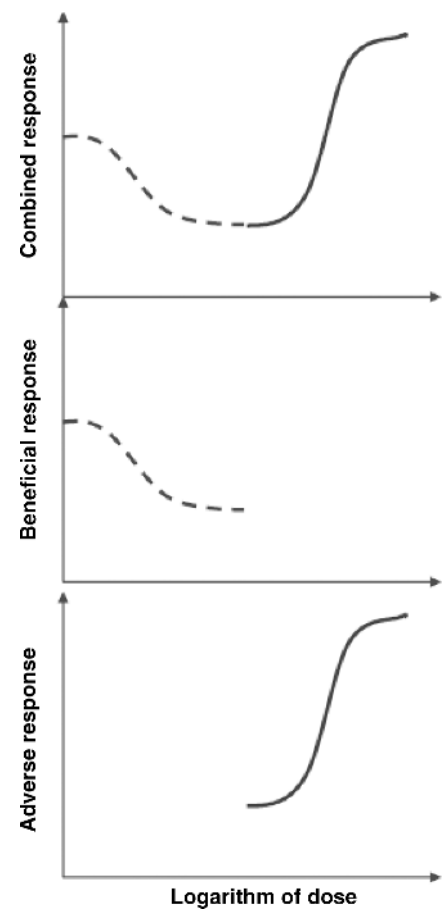


Figure 1.6 Hypothetical hormetic dose–response

Table 1.2 Mathematical representations of chemical interactions (reproduced from James *et al.*, 2000)

Effect	Hypothetical mathematical illustration	Example
Additive	$2 + 3 = 5$	Organophosphate pesticides
Synergistic	$2 + 3 = 20$	Cigarette smoking + asbestos
Potentiation	$2 + 0 = 10$	Alcohol + carbon tetrachloride
Antagonism	$6 + 6 = 8$ or	Toluene + benzene
	$5 + (-5) = 0$ or	Caffeine + alcohol
	$10 + 0 = 2$	Dimercaprol + mercury

are based, whereby a chemical can reduce the harm caused by a toxicant (James *et al*., 2000; Duffus, 2006). Mathematical illustrations and examples of known chemical interactions are given in Table 1.2.

There are four main ways in which chemicals may interact (James *et al*., 2000);

1. Functional: both chemicals have an effect on the same physiological function.
2. Chemical: a chemical reaction between the two compounds affects the toxicity of one or both compounds.
3. Dispositional: the absorption, metabolism, distribution or excretion of one substance is increased or decreased by the presence of the other.
4. Receptor-mediated: when two chemicals have differing affinity and activity for the same receptor, competition for the receptor will modify the overall effect.

1.2.6 Relevance of animal models

A further complication in the extrapolation of the results of toxicological experimental studies to humans, or indeed other untested species, is related to the anatomical, physiological and biochemical differences between species. This requires some previous knowledge of the mechanism of toxicity of a chemical and the comparative physiology of different test species. When adverse effects are detected in screening tests, these should be interpreted with the relevance of the chosen animal model in mind. For the derivation of safe levels, safety or uncertainty factors are again usually applied to account for the uncertainty surrounding inter-species differences (James *et al*., 2000; Sullivan, 2006).

1.2.7 A few words about doses

When discussing dose–response, it is also important to understand which dose is being referred to and to differentiate between concentrations measured in environmental media and the concentration that will elicit an adverse effect at the target organ or tissue. The exposure dose in a toxicological testing setting is generally known or can be readily derived or measured from concentrations in media and average consumption (of food or water for example) (Figure 1.7). Whilst toxicokinetics help to develop an understanding of the relationship between the internal dose and a known exposure dose, relating concentrations in environmental media to the actual exposure dose, often via multiple pathways, is in the realm of exposure assessment.

1.2.8 Other hazard characterisation criteria

It is important to understand the difference between hazard and risk. Hazard is defined as the potential to produce harm; it is therefore an inherent qualitative attribute of a given chemical substance. Risk, on the other hand, is a quantitative measure of the magnitude of the hazard and the probability of it being realised. Hazard assessment is therefore the first step of risk assessment, followed by exposure assessment and finally risk characterisation.

'Carcinogenic, mutagenic or reprotoxic' (CMR) is a designation applied by manufacturers and legislative bodies in the EU to chemicals identified as hazardous substances capable of: initiating cancer; increasing the frequency of changes in an organism's genetic material above their natural background level; and/or harm the ability of organisms to successfully reproduce. Chemical producers wishing to produce or import chemicals in quantities greater than 1 t per year in the EU are now obliged to carry out standardised toxicological tests to identify

Figure 1.7 Relationships between environmental concentration, exposure dose and internal dose

CMR properties. Formerly, this information was forwarded to EU designated member-state authorities to be classified by the European Chemicals Bureau (Langezaal, 2002). Since 2007, however, regulatory functions have been passed over to the European Chemicals Agency, which has replaced the European Chemicals Bureau and plays a central role in coordinating the implementation of the REACH Directive. REACH is the legislation now in place for the **R**egistration, **E**valuation, **A**uthorisation and Restriction of **Ch**emicals. Substances which are found to have CMR properties are classified alongside persistent, bioaccumulative and toxic substances as Substances of Very High Concern. Their sale and use are strictly regulated (ECHA, 2007).

Toxicity is not the sole criterion evaluated for hazard characterisation purposes. Some chemicals have been found in the tissues of animals in the arctic, for example, where they have never been used or produced. This realisation that some persistent pollutants are able to travel considerable distances and bioaccumulate through the food web has led researchers to take account of such inherent properties of organic compounds, as well as their toxicity, for the purpose of hazard characterisation.

Persistence is the result of resistance to environmental degradation mechanisms such as hydrolysis, photodegradation and biodegradation. Hydrolysis only occurs in the presence of water, photodegradation in the presence of UV light and biodegradation is primarily carried out by micro-organisms. Degradation is related to water solubility, itself inversely related to lipid solubility, so persistence tends to be correlated with lipid solubility (Francis, 1994). The persistence of inorganic substances has proved more difficult to define as they cannot be degraded to carbon and water.

Chemicals may accumulate in environmental compartments and constitute environmental sinks that could be remobilised and lead to toxic effects on organisms. Further, some substances can accumulate in one species without adverse effects but be toxic to its predator(s). *Bioconcentration* refers to accumulation of a chemical from its surrounding environment rather than specifically through food uptake. *Biomagnification* refers to uptake from food without consideration for uptake through the body surface. *Bioaccumulation* integrates both paths, surrounding medium and food. *Ecological magnification* refers to an increase in concentration through the food web from lower to higher trophic levels. Accumulation of organic compounds generally involves transfer from a hydrophilic to a hydrophobic phase and correlates well with the n-octanol/water partition coefficient (Herrchen, 2006).

Persistence and bioaccumulation of a substance is evaluated by standardised OECD tests. Criteria for the identification of persistent, bioaccumulative and toxic substances (PBT), and very persistent and very bioaccumulative substances (vPvB) as defined in Annex XIII of the European Directive on the Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH) (European Union, 2006) are given in Table 1.3. To be classified as a PBT or vPvB, a given substance must fulfil each criterion.

Table 1.3 REACH criteria for identifying PBT and vPvB chemicals

Criterion	PBT criteria	vPvB criteria
Persistence	Either: Half-life > 60 days in marine water Half-life > 60 days in fresh or estuarine water Half-life > 180 days in marine sediment Half-life > 120 days in fresh or estuarine sediment Half-life > 120 days in soil	Either: Half-life > 60 days in marine, fresh or estuarine water Half-life > 180 days in marine, fresh or estuarine sediment Half-life > 180 days in soil
Bioaccumulation	Bioconcentration factor (BCF) > 2000	Bioconcentration factor (BCF) > 2000
Toxicity	Either: Chronic no-observed effect concentration (NOEC) < 0.01 mg/l substance is classified as carcinogenic (category 1 or 2), mutagenic (category 1 or 2), or toxic for reproduction (category 1, 2 or 3) there is other evidence of endocrine-disrupting effects	

Key points

- Toxicological studies generate experimental data to further our understanding of the mode of action and/or toxicity of a xenobiotic.
- Dose–response relationships are central to the prediction of toxicity in regulatory risk assessment.
- There are many uncertainties in the extrapolation of such results to the prediction of actual risks from environmental exposure to low levels of numerous pollutants via multiple routes.

1.3 Some notions of environmental epidemiology

Epidemiology is an observational approach to the study of associations between environment and disease. It can be defined as 'the study of how often diseases occur and why, based on the measurement of disease outcome in a study sample in relation to a population at risk.' (Coggon *et al*., 2003). Environmental epidemiology refers to the study of distribution patterns of disease and health related to exposures that are exogenous and involuntary. Such exposures generally occur in the air, water, diet or soil and include physical, chemical and biological agents. The extent to which environmental epidemiology is considered to include social, political, cultural and engineering or architectural factors affecting human contact with such agents varies according to authors. In some contexts, the environment can refer to all non-genetic factors, although dietary habits are generally excluded, despite the facts that some deficiency diseases are environmentally determined and nutritional status may also modify the impact of an environmental exposure (Steenland and Savitz, 1997; Hertz-Picciotto, 1998).

Most of environmental epidemiology is concerned with endemics (acute or chronic disease occurring at relatively low frequency in the general population due partly to a common and often unsuspected exposure) rather than epidemics (acute outbreaks of disease affecting a limited population shortly after the introduction of an unusual known or unknown agent). Measuring such low-level exposure to the general public may be difficult – if not impossible – particularly when seeking historical estimates of exposure to predict future disease. Estimating very small changes in the incidence of health effects of low-level common multiple exposure on common diseases with multifactorial aetiologies is particularly difficult, because often greater variability may be expected for other reasons and environmental epidemiology has to rely on natural experiments which, unlike controlled experiments, are subject to other, often unknown, risk factors and variables. However, environmental epidemiology may still be important from a public-health perspective, as small effects in a large population can have large attributable risks if the disease is common (Steenland and Savitz, 1997; Coggon *et al*., 2003).

1.3.1 Definitions

1.3.1.1 What is a case?

The definition of a case generally requires a dichotomy, i.e. for a given condition people can be divided into two discrete classes – the affected and the non-affected. It increasingly appears that diseases exist in a continuum of severity within a population rather than an all-or-nothing phenomenon. For practical reasons, a cut-off point to divide the diagnostic continuum into 'cases' and 'non-cases' is therefore required. This can be done on a statistical, clinical, prognostic or operational basis. On a statistical basis, the 'norm' is often defined as within two standard deviations of the age-specific mean, thereby arbitrarily fixing the frequency of abnormal values at around 5 per cent in every population. Moreover, it should be noted that what is usual is not necessarily good. A clinical case may be defined by the level of a variable above which symptoms and complications have been found to become more frequent. On a prognostic basis, some clinical findings may carry an adverse prognosis, yet be symptomless. When none of the other approaches is satisfactory, an operational threshold will need to be defined, e.g. based on a threshold for treatment (Coggon *et al*., 2003).

1.3.1.2 Incidence, prevalence and mortality

The *incidence* of a disease is the rate at which new cases occur in a population during a specified period or frequency of incidents.

$$\text{Incidence} = \frac{\text{Number of new cases}}{\text{Population at risk} \times \text{time during which cases were ascertained}}$$

The *prevalence* of a disease is the proportion of the population that are cases at a given point in time. This measure is appropriate only in relatively stable conditions and is unsuitable for acute disorders. Even in a chronic disease, the manifestations are often intermittent and a point prevalence will tend to underestimate the frequency of the condition. A better measure when possible is the period prevalence, defined as the proportion of a population that are cases at any time within a stated period.

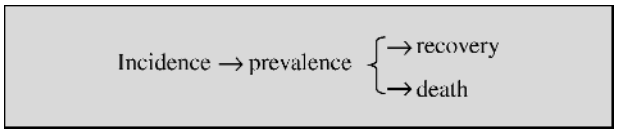
$$\text{Prevalence} = \text{incidence} \times \text{average duration}$$

In studies of aetiology, incidence is the most appropriate measure of disease frequency, as different prevalences result from differences in survival and recovery as well as incidence.

Mortality is the incidence of death from a disease (Coggon *et al*., 2003).

1.3.1.3 Interrelation of incidence, prevalence and mortality

Each incident case enters a prevalence pool and remains there until either recovery or death:



A chronic condition will be characterised by both low recovery and low death rates, and even a low incidence will produce a high prevalence (Coggon *et al*., 2003).

1.3.1.4 Crude and specific rates

A crude incidence, prevalence or mortality is one that relates to results for a population taken as a whole, without subdivisions or refinement. To compare populations or samples, it may be helpful to break down results for the whole population to give rates specific for age and sex (Coggon *et al*., 2003).

1.3.1.5 Measures of association

Several measures are commonly used to summarise association between exposure and disease.

Attributable risk is most relevant when making decisions for individuals and corresponds to the difference between the disease rate in exposed persons and that in unexposed persons. The *population attributable risk* is the difference between the rate of disease in a population and the rate that would apply if all of the population were unexposed. It can be used to estimate the potential impact of control measures in a population.

Population attributable risk

= attributable risk × prevalence of exposure to risk factor

The *attributable proportion* is the proportion of disease that would be eliminated in a population if its disease rate were reduced to that of unexposed persons. It is used to compare the potential impact of different public-health strategies.

The *relative risk* is the ratio of the disease rate in exposed persons to that in people who are unexposed.

Attributable risk

= rate of disease in unexposed persons × (relative risk − 1)

Relative risk is less relevant to risk management but is nevertheless the measure of association most commonly used because it can be estimated by a wider range of study designs. Additionally, where two risk factors for a disease act in concert, their relative risks have often been observed empirically to come close to multiplying.

The *odds ratio* is defined as the odds of disease in exposed persons divided by the odds of disease in unexposed persons (Coggon *et al*., 2003).

1.3.1.6 Confounding

Environmental epidemiological studies are observational, not experimental, and compare people who differ in various ways, known and unknown. If such differences happen to determine risk of disease independently of the exposure under investigation, they are said to confound its association with the disease and the extent to which observed association are causal. It may equally give rise to spurious associations or obscure the effects of a true cause (Coggon *et al*., 2003). A confounding factor can be defined as a variable which is both a risk factor for the disease of interest, even in the absence of exposure (either causal or in association with other causal factors), and is associated with the exposure but not a direct consequence of the exposure (Rushton, 2000).

In environmental epidemiology, nutritional status suggests potential confounders and effect modifiers of associations between environment and disease. Exposure to environmental agents is also frequently determined by social factors: where one lives, works, socialises or buys food, and some argue that socio-economic context is integral to most environmental epidemiology problems (Hertz-Picciotto, 1998).

Standardisation is usually used to adjust for age and sex, although it can be applied to account for other confounders. Other methods include mathematical modelling techniques such as logistic regression and are readily available. They should be used with caution, however, as the mathematical assumptions in the model may not always reflect the realities of biology (Coggon *et al*., 2003).

1.3.1.7 Standardisation

Direct standardisation is suitable only for large studies, and entails the comparison of weighted averages of age and sex-specific disease rates, the weights being equal to the proportion of people in each age and sex group in a reference population.

In most surveys the indirect method yields more stable risk estimates. Indirect standardisation requires a suitable reference population for which the class-specific rates are known for comparison with the rates obtained for the study sample (Coggon *et al*., 2003).

1.3.2 Measurement error and bias

1.3.2.1 Bias

Bias is a systematic tendency to underestimate or overestimate a parameter of interest because of a deficiency in the design or

execution of a study. In epidemiology, bias results in a difference between the estimated association between exposure and disease and the true association. Three general types of bias can be identified: selection bias, information bias, and confounding bias. *Information bias* arises from errors in measuring exposure or disease, and the information is wrong to the extent that the relationship between the two can no longer be correctly estimated. *Selection bias* occurs when the subjects studied are not representative of the target population about which conclusions are to be drawn. It generally arises because of the way subjects are recruited or the way cases are defined (Bertollini *et al*., 1996; Coggon *et al*., 2003).

1.3.2.2 Measurement error

Errors in exposure assessment or disease diagnosis can be important sources of bias in epidemiological studies, and it is therefore important to assess the quality of measurements. Errors may be differential (different for cases and controls) or non-differential. Non-differential errors are more likely to occur than differential errors and have until recently been assumed to tend to diminish risk estimates and dilute exposure-response gradients (Steenland and Savitz, 1997). Non-differential misclassification is related to both the precision and the magnitude of the differences in exposure or diagnosis within the population. If these differences are substantial, even a fairly imprecise measurement would not lead to much misclassification. A systematic investigation of the relative precision of the measurement of the exposure variable should ideally precede any study in environmental epidemiology (Bertollini *et al*., 1996; Coggon *et al*., 2003).

1.3.2.3 Validity

The validity of a measurement refers to the agreement between this measure and the truth. It is potentially a more serious problem than a systematic error, because in the latter case the power of a study to detect a relationship between exposure and disease is not compromised. When a technique or test is used to dichotomise subjects, its validity may be analysed by comparison with results from a standard reference test. Such analysis will yield four important statistics: sensitivity, specificity, systematic error and predictive value. It should be noted that both systematic error and predictive value depend on the relative frequency of true positives and true negatives in the study sample (prevalence of the disease or exposure being measured) (Bertollini *et al*., 1996; Coggon *et al*., 2003).

1.3.2.4 Repeatability

When there is no satisfactory standard against which to assess the validity of a measurement technique, then examining the repeatability of measurements within and between observers can offer useful information. Whilst consistent findings do not necessarily imply that a technique is valid, poor repeatability does indicate either poor validity or that the measured parameter varies over time. When measured repeatedly in the same subject, physiological or other variables tend to show a roughly normal distribution around the subject's mean. Misinterpretation can be avoided by repeat examinations to establish an adequate baseline, or by including a control group. Conversely, conditions and timing of an investigation may systematically bias subjects' response and studies should be designed to control for this.

The repeatability of measurements of continuous variables can be summarised by the standard deviation of replicate measurements or by their coefficient of variation. Within-observer variation is considered to be largely random, whilst between-observer variation adds a systematic component due to individual differences in techniques and criteria to the random element. This problem can be circumvented by using a single observer or, alternatively, allocating subjects to observers randomly. Subsequent analysis of results by observers should highlight any problem and may permit statistical correction for bias (Coggon *et al*., 2003).

1.3.3 Exposure assessment

The quality of exposure measurement underpins the validity of an environmental epidemiology study. Assessing exposure on an ever/never basis is often inadequate because the certainty of exposure may be low and a large range of exposure levels with potentially non-homogeneous risks are grouped together. Ordinal categories provide the opportunity to assess dose–response relations, whilst quantified measures, where possible, also allow researchers to assess comparability across studies and can provide the basis for regulatory decision making. Instruments for exposure assessment include (Hertz-Picciotto, 1998):

- interviews, questionnaires, and structured diaries;
- measurement in the macro-environment, either conducted directly or obtained from historical records;
- concentration in the personal micro-environment;
- biomarkers of physiological effect in human tissues or metabolic products.

All questionnaires and interview techniques rely on human knowledge and memory, and hence are subject to error and recall bias. Cases tend to report exposure more accurately than controls and this biases risk estimates upwards and could lead to false positive results. There are techniques that can be applied to detect this bias, such as including individuals with a disease unrelated to the exposure of interest, probing subjects about the understanding of the relationship between the disease and exposure under study, or attempting to corroborate information given by a sample of the cases and controls through records, interviews, or environmental or biological monitoring. Interviews either face-to-face or on the phone

may also elicit under-reporting of many phenomena subject to the 'desirability' of the activity being reported. Self-administered questionnaires or diaries can avoid interviewer influences but typically have lower response rates and do not permit the collection of complex information (Bertollini *et al.*, 1996; Hertz-Picciotto, 1998).

A distinction has been made between exposure measured in the external environment, at the point of contact between the subject and its environment, and measurements made in human tissue or sera. Measurements in external media yield an ecological measure and are useful when group differences outweigh inter-individual differences. Macro-environment measures are also more relevant to the exposure context rather than to individual pollutants. Sometimes, the duration of contact (or potential contact) can be used as a surrogate quantitative measure, the implicit assumption being that duration correlates with cumulative exposure. When external measurements are available, they can be combined with duration and timing of residence and activity-pattern information to assign quantitative exposure estimates for individuals. Moreover, many pollutants are so dispersed in the environment that they can reach the body through a variety of environmental pathways (Bertollini *et al.*, 1996; Hertz-Picciotto, 1998).

The realisation that human exposure to pollutants in micro-environments may differ greatly from those in the general environment was a major advance in environmental epidemiology. It has led to the development of instrumentation suitable for micro-environmental and personal monitoring and sophisticated exposure models. Nonetheless, these estimates of individual absorbed doses still do not account for inter-individual differences due to breathing rate, age, sex, medical conditions, and so on (Bertollini *et al.*, 1996; Hertz-Picciotto, 1998).

The pertinent dose at the target tissue depends on toxicokinetics, metabolic rates and pathways that could either produce the active compound or detoxify it, as well as storage and retention times, and elimination. Measuring and modelling of integrated exposure to such substances are difficult at best, and, when available, the measurement of biomarkers of internal doses will be the preferred approach. Whilst biomarkers can account for individual differences in pharmacokinetics, they do not tell us which environmental sources and pathways are dominating exposure and in some situations they could be poor indicators of past exposure. Moreover, many pollutants are so dispersed in the environment that they can reach the body through a variety of environmental pathways (Bertollini *et al.*, 1996; Hertz-Picciotto, 1998).

To study diseases with long latency periods, such as cancer, or those resulting from long-term chronic insults, exposures or residences at times in the past are more appropriate. Unfortunately, reconstruction of past exposures is often fraught with problems of recall, incomplete measurements of external media, or inaccurate records that can no longer be validated, and retrospective environmental exposure assessment techniques are still in their infancy (Bertollini *et al.*, 1996; Hertz-Picciotto, 1998).

1.3.4 Types of studies

1.3.4.1 Ecological studies

In ecological studies, the unit of observation is the group, a population or a community, rather than the individual. The relation between disease rates and exposures in each of a series of populations is examined. Often the information about disease and exposure is abstracted from published statistics such as those published by the World Health Organisation (WHO) on a country-by-country basis. The populations compared may be defined in various ways (Steenland and Savitz, 1997; Coggon *et al.*, 2003):

- Geographically. Care is needed in the interpretation of results, due to potential confounding effects and differences in ascertainment of disease or exposure.
- Time trends or time series. Like geographical studies, analysis of secular trends may be biased by differences in the ascertainment of disease. However, validating secular changes is more difficult as it depends on observations made and often scantily recorded many years ago.
- Migrants studies. These offer a way of discriminating genetic from environmental causes of geographical variation in disease, and may also indicate the age at which an environmental cause exerts its effect. However, the migrants may themselves be unrepresentative of the population they leave, and their health may have been affected by the process of migration.
- By occupation or social class. Statistics on disease incidence and mortality may be readily available for socio-economic or occupational groups. However, occupational data may not include data on those who left this employment, whether on health grounds or not, and different socio-economic groups may have different access to healthcare.

1.3.4.2 Longitudinal or cohort studies

In a longitudinal study subjects are identified and then followed over time with continuous or repeated monitoring of risk factors and known or suspected causes of disease and subsequent morbidity or mortality. In the simplest design, a sample or cohort of subjects exposed to a risk factor is identified along with a sample of unexposed controls. By comparing the incidence rates in the two groups, attributable and relative risks can be estimated. Case-response bias is entirely avoided in cohort studies where exposure is evaluated before diagnosis. Allowance can be made for suspected confounding factors, either by matching the controls to the exposed subjects so that they have similar patterns of exposure to the confounder, or by measuring exposure to the confounder in each group and adjusting for any difference in statistical analysis. One of the main limitations of this method is that when it is applied to the study of chronic diseases a large number of people must be followed up for long periods before sufficient cases accrue to give statistically meaningful results.

When feasible, the follow-up could be carried out retrospectively, as long as the selection of exposed people is not influenced by factors related to their subsequent morbidity. It can also be legitimate to use the recorded disease rates in the national or regional population for control purposes, when exposure to the hazard in the general population is negligible (Bertollini *et al*., 1996; Coggon *et al*., 2003).

1.3.4.3 Case-control studies

In a case-control study, patients who have developed a disease are identified and their past exposure to suspected aetiological factors is compared with that of controls or referents that do not have the disease. This allows the estimation of odds ratio but not of attributable risks. Allowance is made for confounding factors by measuring them and making appropriate adjustments in the analysis. This adjustment may be rendered more efficient by matching cases and controls for exposure to confounders, either on an individual basis or in groups. Unlike a cohort study, however, matching does not on its own eliminate confounding, and statistical adjustment is still required (Coggon *et al*., 2003).

1.3.4.3.1 Selection of cases and controls In general, selecting incident rather than prevalent cases is preferred. The exposure to risk factors and confounders should be representative of the population of interest within the constraints of any matching criteria. It often proves impossible to satisfy both those aims. The exposure of controls selected from the general population is likely to be representative of those at risk of becoming cases, but assessment of their exposure may not be comparable with that of cases due to recall bias, and studies will tend to overestimate risk. Recall bias can be addressed by including a control group composed of patients with other diseases, but their exposure may be unrepresentative, and studies will tend to underestimate risk if the risk factor under investigation is involved in other pathologies. It is therefore safer to adopt a range of control diagnoses rather than a single disease group. Interpretation can also be helped by having two sets of controls with different possible sources of bias. Selecting equal numbers of cases and controls generally makes a study most efficient, but the number of cases available can be limited by the rarity of the disease of interest. In this circumstance, statistical confidence can be increased by taking more than one control per case. There is, however, a law of diminishing returns, and it is usually not worth going beyond a ratio of four or five controls to one case (Coggon *et al*., 2003).

1.3.4.3.2 Exposure assessment Many case-control studies ascertain exposure from personal recall, using either a self-administered questionnaire or an interview. Exposure can sometimes be established from existing records such General Practice notes. Occasionally, long-term biological markers of exposure can be exploited, but they are only useful if not altered by the subsequent disease process (Coggon *et al*., 2003).

1.3.4.4 Cross-sectional studies

A cross-sectional study measures the prevalence of health outcomes or determinants of health, or both, in a population at a point in time or over a short period. The risk obtained is disease prevalence rather than incidence. Such information can be used to explore aetiology, but associations must be interpreted with caution. Bias may arise because of selection into or out of the study population, giving rise to effects similar to the healthy-worker effect encountered in occupational epidemiology. A cross-sectional design may also make it difficult to establish what is cause and what is effect. Because of these difficulties, cross-sectional studies of aetiology are best suited to non-fatal degenerative diseases with no clear point of onset and to the pre-symptomatic phases of more serious disorders (Rushton, 2000; Coggon *et al*., 2003).

1.3.5 Critical appraisal of epidemiological reports

1.3.5.1 Design

A well-designed study should state precisely formulated, written objectives and the null hypothesis to be tested. This should in turn demonstrate the appropriateness of the study design for the hypothesis to be evaluated. Ideally, a literature search of relevant background publications should be carried out in order to explore the biological plausibility of the hypothesis (Elwood, 1998; Rushton, 2000; Coggon *et al*., 2003).

In order to be able to appraise the selection of subjects, each study should first describe the target population that the study participants are meant to represent. The selection of study participants affects not only how widely the results can be applied but also, more importantly, their validity. The internal validity of a study relates to how well a difference between the two groups being compared can be attributed to the effects of exposure rather than to chance or confounding bias. In contrast, the external validity of a study refers to how well the results can be applied to the general population. Whilst both are desirable, design considerations that help increase the internal validity of a study may decrease its external validity. However, the external validity of a study is only useful if the internal validity is acceptable. The selection criteria should therefore be appraised by considering the effects of potential selection bias on the hypothesis being tested and the external and internal validity of the study population. The selection process itself should be effectively random (Elwood, 1998; Rushton, 2000; Coggon *et al*., 2003).

The sample size should allow the primary purpose of the study, formulated in precise statistical terms, to be achieved, and its adequacy should be assessed. If it is of particular interest that certain subgroups are relatively over-represented, a stratified random sample can be chosen by dividing the study population into strata and then drawing a separate random sample from each. Two-stage sampling may be adequate when the study

population is large and widely scattered but there is some loss of statistical efficiency, especially if only a few units are selected at the first stage (Rushton, 2000; Coggon *et al*., 2003).

To be able to appraise a study, a clear description of how the main variables were measured should be given. The choice of method needs to allow a representative sample of adequate size to be examined in a standardised and sufficiently valid way. Ideally, observers should be allocated to subjects in a random manner to minimise bias due to observer differences. Importantly, methods and observers should allow rigorous standardisation (Rushton, 2000; Coggon *et al*., 2003).

1.3.5.2 Bias

Almost all epidemiological studies are subject to bias, and it is important to allow for the probable impact of biases in drawing conclusions. In a well-reported study, this question would already have been addressed by the authors, who may even have collected data to help quantify bias (Coggon *et al*., 2003).

Selection bias, information bias and confounding have all been discussed in some detail in previous sections, but it is worth mentioning the importance of accurately reporting response rates, as selection bias can also result if participants differ from non-participants. The likely bias resulting from incomplete response can be assessed in different ways: subjects who respond with and without a reminder could be compared, or a small random sample can be drawn from the non-responders and particularly vigorous efforts made to collect some of the information that was originally sought and findings then compared with those of the earlier responders; or differences based on available information about the study population such as age, sex and residence could give an indication of the possibility of bias, and making extreme assumptions about the non-responders can help to put boundaries on the uncertainty arising from non-response (Elwood, 1998).

1.3.5.3 Statistical analysis

Even after biases have been taken into account, study samples may be unrepresentative just by chance. An indication of the potential for such chance effects is provided by statistical analysis and hypothesis testing. There are two kinds of errors that one seeks to minimise. A type-I error is the mistake of concluding that a phenomenon or association exists when in truth it does not; by convention, the rate of such errors is usually set at 5 per cent. A result is therefore called statistically significant, when there is a less than 5 per cent probability to have observed an association in the experiment when such an association does not actually exist. A type-II error, failing to detect an association that actually does exist, is, also by convention, often set at 20 per cent, although this is in fact often determined by practical limitations of sample size (Armitage and Berry, 1994). It is important to note that failure to reject the null hypothesis (i.e. no association) does not equate with its acceptance but only provides reasonable confidence that if any association exists it would be smaller than an effect size determined by the power of the study. The issues surrounding power and effect size should normally be addressed at the design stage of a study, although this is rarely reported (Rushton, 2000).

1.3.5.4 Confounding versus causality

If an association is found and not explained by bias or chance, the possibility of unrecognised residual confounding still remains. Assessment of whether an observed association is causal depends in part on the biological plausibility of the relation. Certain characteristics of the association, such as an exposure–response gradient, may encourage causal interpretation, though in theory it may still arise from confounding. Also important is the magnitude of the association as measured by the relative risk or odds ratio. The evaluation of possible pathogenic mechanisms and the importance attached to exposure–response relations and evidence of latency are also a matter of judgement (Coggon *et al*., 2003).

1.3.6 Future directions

Some progress has been made in the area of exposure assessment, but more work is needed in integrating biological indicators into exposure assessment, and much remains to be done with respect to timing of exposures as they relate to induction and latency issues.

An obstacle to analysis of multiple exposures is the near impossibility of separating induction periods, dose–response, and interactive effects from one another. These multiple exposures include not only the traditional chemical and physical agents, but should also be extended to social factors as potential effect modifiers.

An emerging issue for environmental epidemiologists is that of variation in susceptibility. This concept is not new: it constitutes the element of the 'host' in an old paradigm of epidemiology that divided causes of disease into environment, host and agent. It has, however, taken on a new dimension with the current technology that permits identification of genes implicated in many diseases. The study of gene–environment interactions as a mean of identifying susceptible subgroups can lead to studies with a higher degree of specificity and precision in estimating effects of exposures.

Key points

- The type of studies required to obtain reliable estimates of long-term risks following chronic exposures are expensive and time-consuming.
- Again, assessing environmental exposure to low levels of pollutants via multiple routes is an issue.
- Inter-individual variability and interactions between environment and disease can also obscure results.

1.4 Scientific evidence and the precautionary principle

1.4.1 Association between environment and disease

Scientific evidence on associations between exogenous agents and health effects is derived from epidemiological and toxicological studies. As discussed previously, both types of methods have advantages and disadvantages, and much uncertainty and controversy stem from the relative weights attributed to different types of evidence. Environmental epidemiology requires the estimation of often very small changes in the incidence of common diseases with multifactorial aetiologies following low-level multiple exposures. For ethical reasons, it is necessarily observational, and natural experiments are subject to confounding and to other, often unknown, risk factors (Steenland and Savitz, 1997; Coggon *et al*., 2003). Some progress has been made in the development of specific biomarkers, but this is still hindered by issues surrounding the timing of exposures as they relate to induction and latency. Toxicology, on the other hand, allows the direct study of the relationship between the quantity of chemical to which an organism is exposed and the nature and degree of consequent harmful effect. Controlled conditions, however, limit the interpretation of toxicity data, as they generally differ considerably from those prevailing in the natural environment.

1.4.1.1 Bradford-Hill criteria

Since 1965, evaluations of the association between environment and disease have often been based on the nine 'Bradford-Hill criteria' (Hill, 1965).

Bradford-Hill criteria

- Strength
- Consistency
- Specificity
- Temporality
- Biological gradient
- Plausibility
- Coherence
- Experiment
- Analogy

Results from cohort, cross-sectional or case-control studies of not only environmental but also accidental, occupational, nutritional or pharmacological exposure, as well as toxicological studies, can inform all the Bradford-Hill tenets of association between environment and disease. Such studies often include some measure of the *strength of the association* under investigation and its statistical significance. Geographical studies and migrant studies provide some insights into the *consistency of observations.* Consistency of observations between studies of different chemicals exhibiting similar properties, or between studies of different species, should also be considered. Whilst *specificity* provides evidence of specific environment–disease association, the lack of it, or association with multiple endpoints, does not constitute proof against a potential association. Time-trend analyses are directly related to the *temporality* aspect of a putative association, whether trends in environmental release of the chemical agents of interest precedes similar trends in the incidence of disease. This is also particularly relevant in the context of the application of the precautionary principle, as the observation of intergenerational effects in laboratory animals (Newbold *et al*., 1998, 2000) may raise concerns of 'threats of irreversible damage'. Occasionally, studies are designed to investigate the existence of a *biological gradient* or dose–response. *Plausibility* is related to the state of mechanistic knowledge underlying a putative association, while *coherence* can be related to what is known of the aetiology of the disease. *Experimental evidence* can be derived both from toxicological studies and from natural epidemiological experiments following occupational or accidental exposure. Finally, *analogy,* where an association has been shown for analogous exposure and outcomes, should also be considered.

1.4.2 Precautionary principle

A common rationale for the precautionary principle is that increasing industrialisation and the accompanying pace of technological development and widespread use of an ever-increasing number of chemicals exceed the time needed to test those chemicals adequately and collect sufficient data to form a clear consensus among scientists as to their potential to do harm (Burger, 2003).

The precautionary principle became European Law in 1992 when the Maastricht Treaty modified Article 130r of the treaty establishing the European Economic Community, and in just over a decade it has also been included in several international environmental agreements (Marchant, 2003). The precautionary principle is nonetheless still controversial and lacks a definitive formulation. This is best illustrated by the differences between two well-known definitions of the principle, the Rio Declaration produced in 1992 by the United Nations Conference on Environment and Development and the Wingspread Statement formulated by proponents of the precautionary principle in 1998 (Marchant, 2003). One interpretation of the principle is therefore that uncertainty is not justification to delay the prevention of a potentially harmful action, whilst the other implies that no action should be taken unless it is certain that it will do no harm

(Rogers, 2003). Definitions also differ in the level of harm necessary to trigger action, from 'threats of serious or irreversible damage' to 'possible risks' (Marchant, 2003). Whilst there are situations where risks clearly exceed benefits and vice versa, there is a large grey area in which science alone cannot decide policy (Kriebel *et al*., 2001), and the proponents of a strong precautionary principle advocate public participation as a means to make environmental decision-making more transparent. This will require the characterisation and efficient communication of scientific uncertainty to policy-makers and the wider public, and scientific uncertainty is a well-known 'dread factor', increasing the public's perception of risk (Slovic, 1987).

> 'When there are threats of serious and irreversible damage, lack of full scientific certainty shall not be used as a reason for postponing cost-effective measures to prevent environmental degradation'
>
> *Rio Declaration*

> 'When an activity raises threats of harms to human health or the environment, precautionary measures should be taken even if some cause and effect relationships are not fully established scientifically'
>
> *Wingspread Statement*

1.4.2.1 Sufficiency of evidence

The European Environment Agency *Late Lessons* report (2001) provided a working definition of the precautionary principle, and this was improved following further discussions and legal developments.

> 'The Precautionary Principle provides justification for public policy actions in situations of scientific complexity, uncertainty and ignorance, where there may be a need to act in order to avoid, or reduce, potentially serious or irreversible threats to health or the environment, using an appropriate level of scientific evidence, and taking into account the likely pros and cons of action and inaction.'
>
> (*Gee, 2006*)

It specifies complexity, uncertainty and ignorance as contexts where the principle may be applicable and makes explicit mention that precautionary actions need to be justified by a sufficiency of scientific evidence. The report also offers a clarification of the terms Risk, Uncertainty and Ignorance and corresponding states of knowledge with some examples of proportionate actions (Table 1.4).

1.5 Uncertainty and controversy: the endocrine disruption example

More than 10 years after the publication of Theo Colborn's *Our Stolen Future* (Colborn *et al*., 1996), endocrine disruption probably remains one of the most controversial current environmental issues. News stories about the potential effects of 'gender-bending chemicals' on unborn male fetuses are still being printed in some sections of the general media, while by virtue of the precautionary principle, the term 'endocrine disrupters' can be found in emerging European environmental legislation, such as the Water Framework Directive or the REACH proposal (European Community, 2000; Commission

Table 1.4 Examples of precautionary actions and the scientific evidence justifying them (reproduced from European Environment Agency, 2001)

Situation	State and dates of knowledge	Examples of action
Risk	'Known impacts'; 'known probabilities'; e.g. asbestos causing respiratory disease, lung and mesothelioma cancer, 1965–present	Prevention: action taken to reduce known hazards; e.g. eliminate exposure to asbestos dust
Uncertainty	'Known' impacts; 'unknown' probabilities; e.g. antibiotics in animal feed and associated human resistance to those antibiotics, 1969–present	Precautionary prevention: action taken to reduce potential risks; e.g. reduce/eliminate human exposure to antibiotics in animal feed
Ignorance	'Unknown' impacts and therefore 'unknown'probabilities; e.g. the 'surprises' of chlorofluorocarbons (CFCs) and ozone layer damage prior to 1974; asbestos mesothelioma cancer prior to 1959	Precaution: action taken to anticipate, identify and reduce the impact of 'surprises'; e.g. use of properties of chemicals such as persistence or bioaccumulation as 'predictors' of potential harm; use of the broadest possible sources of information, including long-term monitoring; promotion of robust, diverse and adaptable technologies and social arrangements to meet needs, with fewer technological 'monopolies' such as asbestos and CFCs

of the European Communities, 2003). It is therefore of interest to consider here what makes endocrine disruption such a challenging topic for environmental toxicologists.

1.5.1 Emergence of the 'endocrine disruption' hypothesis

The realisation that human and animal hormonal function could be modulated by synthetic variants of endogenous hormones is generally attributed to a British scientist, Sir Edward Charles Dodds (1889–1973), a professor of biochemistry at the Middlesex Hospital Medical School at the University of London, who had won international acclaim for his synthesis of the oestrogen diethylstilbestrol (DES) in 1938, subsequently prescribed for a variety of gynaecologic conditions, including some associated with pregnancy (Krimsky, 2000). By then, it was also known that the sexual development of both male and female rodents could be disrupted by prenatal exposure to sex hormones (Greene *et al*., 1938). It was not until 1971, however, that an association was made between DES exposure *in utero* and a cluster of vaginal clear-cell adenocarcinoma in women under 20, an extremely rare type of cancer for this age group (Herbst *et al*., 1971). It took another 10 years to link DES prescription to pregnant women to other genital-tract abnormalities in their progeny.

Meanwhile, Rachel Carson famously associated the oestrogenic pesticide *o,p*-dichlorodiphenyltrichloroethane (DDT) with eggshell thinning in her book *Silent Spring* (Carson, 1962). Until then, overexploitation and habitat destruction were considered the most significant causes of declining wildlife populations. Pesticides were subsequently found in tissues of wildlife from remote parts of the world, and Carson's observation that these concentrations increased with trophic levels, a process called biomagnification, was verified. Nevertheless, it took a further 30 years for the endocrine disruption hypothesis to emerge as a result of the convergence of several separate lines of enquiry.

In 1987, Theo Colborn began an extensive literature search on toxic chemicals in the Great Lakes. Wildlife toxicology had previously concentrated on acute toxicity and cancer, but Colborn found that reproductive and developmental abnormalities were more common than cancer and effects were often observed in the offspring of exposed wildlife (Colborn *et al*., 1996). Another path to the generalised endocrine hypothesis originated from studies of male infertility and testicular cancer. The advent of artificial insemination was accompanied by the development of techniques to assess sperm quality and such information began to be recorded. In the early 1970s, Skakkebaek, a Danish paediatric endocrinologist, noticed a group of cells resembling fetal cells in the testes of men diagnosed with testicular cancer and began to suspect that testicular cancer had its origin in fetal development. A study of 'normal' subjects in the mid-1980s found that 50 per cent of these males had abnormal sperm; it was suggested that environmental factors may be at work, and oestrogenic compounds were suspected (Carlsen *et al*., 1992).

Although the concept of endocrine disruption first developed when it was observed that some environmental chemicals were able to mimic the action of the sex hormones (oestrogens and androgens), it has now evolved to encompass a range of mechanisms involving the many hormones secreted directly into the blood circulatory system by the glands of the endocrine system and their specific receptors and associated enzymes (Harvey *et al*., 1999).

1.5.2 Definitions

Many national and international agencies have proposed their definitions for endocrine disrupters. One of the most commonly used definitions is referred to as the 'Weybridge' definition and was drafted at a major European Workshop in December 1996.

> 'The Weybridge definition: An endocrine disrupter is an exogenous substance that causes adverse health effects in an intact organism, or its progeny, subsequent to changes in endocrine function.'
>
> (*MRC Institute for Environment and Health, 1997*)

> 'An endocrine disruptor is an exogenous agent that interferes with the synthesis, secretion, transport, binding, action, or elimination of natural hormones in the body that are responsible for the maintenance of homeostasis, reproduction, development and/or behaviour.'
>
> (*EPA, 1997*)

> 'An endocrine disruptor is an exogenous substance or mixture that alters function(s) of the endocrine system and consequently causes adverse health effects in an intact organism or its progeny or (sub)populations.'
>
> (*IPCS, 2002*)

A major issue with the Weybridge definition is the use of the term 'adverse'. For a chemical to be considered an endocrine disrupter, its biological effect must amount to an adverse effect on the individual or population and not just a change that falls within the normal range of physiological variation (Barker, 1999).

The US Environmental Protection Agency Risk Assessment Forum's definition focuses more on any biological change regardless of amplitude.

The International Programme on Chemical Safety (IPCS) modified the Weybridge definition to clarify the fact that endocrine disruption is a mechanism that explains a biological effect.

1.5.3 Modes of action

1.5.3.1 The endocrine system

There are two main systems by which cells of metazoan organisms communicate with each other.

The nervous system serves for rapid communication using chains of interconnected neurones transmitting transient impulses and also producing chemicals called neurotransmitters, which are rapidly destroyed at the synapses. Such responses are generally associated with sensory stimuli.

The endocrine system uses circulating body fluids such as blood to carry chemical messengers secreted by ductless glands to specific receptors non-uniformly distributed on target organs or tissues that are physicochemically programmed to react and respond to them (Highnam and Hill, 1977; Bentley, 1998; Hale *et al*., 2002). These messengers, referred to as hormones, have a longer biological life and are therefore suited to control long-term processes within the body, such as growth, development, reproduction and homeostasis. Recently, the number of endogenous chemicals found to have hormonal activity has increased dramatically. Many are local hormones (paracrine or autocrine), delivered to their target organ by non-endocrine routes (Harvey *et al*., 1999).

Nerves and hormones are often mutually interdependent. Central nervous activity in most animals is likely to be strongly affected by hormones, and hormone production and release are dependent on nervous activity (Highnam and Hill, 1977; Bentley, 1998). Similarly, the endocrine system is known to influence and be influenced by the immune system.

1.5.3.2 Levels of effect

To understand the significance of endocrine disruption, it is necessary to determine whether there is a causal relationship between an environmental factor and an observed effect. Endocrine disruption is not a toxicological endpoint *per se* but a functional change that may or may not lead to adverse effects. Endocrine disruption can be observed at different levels, and each level of observation gives a different insight into the mode of action of an endocrine disrupter. At the cellular level, the information is gained regarding the potential mechanism of action of a contaminant, whilst at the population level, a greater understanding of the ecological significance of such mechanism is gained. A classification of the different levels at which endocrine disruption can be observed is proposed in Figure 1.8.

It is then clear that any effect observed at any one level cannot constitute evidence of endocrine disruption in itself.

Harvey *et al*. have suggested a classification scheme to cover the main types of endocrine and hormonally modulated toxicity (Harvey *et al*., 1999):

- Primary endocrine toxicity involves the direct effect of a chemical on an endocrine gland, manifested by hyperfunction or hypofunction. Because of the interactions between endocrine glands and their hormones and non-endocrine target tissues, direct endocrine toxicity often results in secondary responses.
- Secondary endocrine toxicity occurs when effects are detected in an endocrine gland as a result of toxicity elsewhere in the endocrine axis. An example would be castration cells that develop in the pituitary as a result of testicular toxicity.
- Indirect toxicity involves either toxicity within a non-endocrine organ, such as the liver, resulting in an effect on the endocrine system or the modulation of endocrine physiology as a result of the stress response to a toxicant.

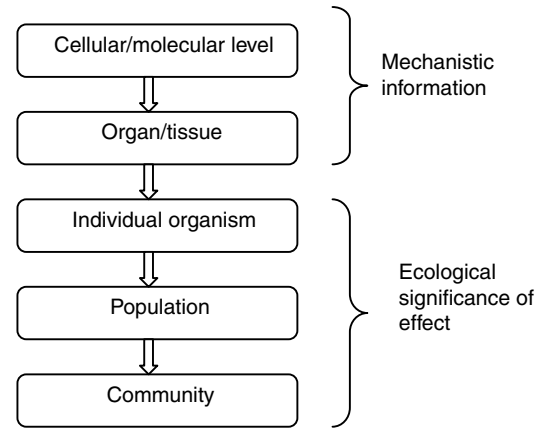


Figure 1.8 Hierachy of levels of observation for endocrine disruption effects

There is a general consensus that indirect endocrine toxicity should not be described as endocrine disruption and that the term itself may have been sometimes misused to include toxicological effects better described in terms of classical toxicology (Eggen *et al*., 2003).

1.5.4 Mechanisms

Endocrine disruption was first recognised when it was found that certain environmental contaminants were able to mimic the actions of endogenous hormones. Some chemicals were subsequently shown to be able to block such actions, and other mechanisms involved in the control of circulating hormone levels were identified.

Contaminants have been shown (Cheek *et al*., 1998; Folmar *et al*., 2001; Guillette and Gunderson, 2001) to:

- act as hormone receptor agonists or antagonists;
- alter hormone production at its endocrine source;
- alter the release of stimulatory or inhibitory hormones from the pituitary or hypothalamus;
- alter hepatic enzymatic biotransformation of hormones;
- alter the concentration or functioning of serum-binding proteins, altering free hormone concentrations in the serum; and
- alter other catabolic pathways of clearance of hormones.

Receptor-mediated mechanisms have received the most attention, but other mechanisms have been shown to be equally important.

1.5.4.1 Hormone–receptor agonism and antagonism

The current focus for concerns about endocrine-mediated toxicity has mostly been on chemicals interacting with the steroid hormone receptor superfamily, receptors for oestrogens, androgens, thyroid hormones, etc. These receptors are predominantly involved in changing gene transcription (Barton and Andersen, 1997). According to the accepted paradigm for receptor-mediated mechanisms, a compound binds to a receptor forming a ligand-receptor complex with high binding affinity for specific DNA sequences or responsive elements. Once bound to this responsive element, the ligand-receptor complex induces gene transcription followed by translation into specific proteins which are the ultimate effectors of observed responses (Zacharewski, 1997).

Whilst hormone agonists are not only able to bind to the receptor under consideration, but also induce gene transcription thereby amplifying endogenous hormonal response, antagonists bind to the receptor but are unable to effect increased gene transcription, but rather competitively inhibit it by their occupancy of receptor binding sites.

1.5.5 Dose–response relationships

The theory of dose–response relationships of xenobiotics generally assumes that they are monotonic, the response rising with the dose. However, endogenous hormones are already present at physiological concentrations and are therefore already beyond the threshold (Andersen *et al*., 1999b). Additionally, most endocrine processes are regulated by feedback controls such as receptor autoregulation and control of enzymes involved in synthesis of high-affinity ligands. This is expected to give rise to highly non-linear dose–response characteristics and abrupt changes from one biological condition to another over a very small change in concentration. While many of these non-linear switching mechanisms are expected to produce non-linear dose–response curves for the action of endogenous hormones, the dose response for effects of exogenous compounds still depends on the combination of effects of the native ligand and the endocrine-disrupting chemical (Andersen *et al*., 1999a). Evidence of low-dose effects has proved very controversial, mainly due to lack of reproducibility, and it has been suggested that this may be related to the natural variability between individuals (Ashby *et al*., 2004).

There are also important time-dependent variations in normal endogenous hormone levels such as circadian rhythms, puberty, oestrous or menstrual cycles, and reproductive senescence and aging. This introduces the additional problems of critical life stages: exposure at sensitive developmental stages can result in irreversible changes and latency (the time between exposure and the observed effects) as was exemplified by DES exposure *in utero* (Barlow *et al*., 2002).

1.6 Concluding remarks

This chapter intends to illustrate the complex issues surrounding attempts to predict the effects of environmental contaminants and equip the reader with some basic concepts that may aid a critical understanding of evidence for such effects. It should encourage rather than deter the reader from reading and referring to the authoritative works cited.

References

Andersen, H. R., A. M. Andersson, S. F. Arnold, H. Autrup, M. Barfoed, N. A. Beresford, P. Bjerregaard, L. B. Christiansen, B. Gissel and R. Hummel *et al.* (1999a). Comparison of short-term estrogenicity tests for identification of hormone-disrupting chemicals. *Environmental Health Perspectives* **107** (Supplement 1): 89–108.

Andersen, M. E., R. B. Conolly, E. M. Faustman, R. J. Kavlock, C. J. Portier, D. M. Sheehan, P. J. Wier and L. Ziese (1999b). Quantitative mechanistically based dose–response modeling with endocrine-active compounds. *Environmental Health Perspectives* **107**: 631–638.

Armitage, P. and G. Berry (1994). *Statistical Methods in Medical Research*. Cambridge, Cambridge University Press.

Ashby, J., H. Tinwell, J. Odum and P. Lefevre (2004). Natural variability and the influence of concurrent control values on the detection and interpretation of low-dose or weak endocrine toxicities. *Environmental Health Perspectives* **112**(8): 847–853.

Barker, J. (1999). Liquid chromatography-mass spectrometry (LC/MS) *Mass Spectrometry*. D. J. Ando. Chichester, John Wiley & Sons: 320–334.

Barlow, S. M., J. B. Greig, J. W. Bridges, A. Carere, A. J. M. Carpy, G. L. Galli, J. Kleiner, I. Knudsen, H. B. W. M. Koeter, L. S. Levy, C. Madsen, S. Mayer, J. F. Narbonne, F. Pfannkuch, M. G. Prodanchuk, M. R. Smith and P. Steinberg (2002). Hazard identification by methods of animal-based toxicology. *Food and Chemical Toxicology* **40**(2–3): 145–191.

Barton, H. A. and M. E. Andersen (1997). Dose–response assessment strategies for endocrine-active compounds. *Regulatory Toxicology and Pharmacology* **25**(3): 292–305.

Bentley, P. J. (1998). *Comparative Vertebrate Endocrinology*. Cambridge, Cambridge University Press.

Bertollini, R., M. D. Lebowitz, R. Saracci and D. A. Savitz, Eds. (1996). *Environmental Epidemiology – Exposure and Disease*. Copenhagen, Lewis Publishers on behalf of the World Health Organisation.

Boobis, A., G. Daston, R. Preston and S. Olin (2009). Application of key events analysis to chemical carcinogens and noncarcinogens. *Critical Reviews in Food Science and Nutrition* **49**(8): 690–707.

Burger, J. (2003). Making decisions in the 21(st) century: Scientific data, weight of evidence, and the precautionary principle. *Pure and Applied Chemistry* **75**(11–12): 2505–2513.

Calabrese, E. and L. Baldwin (2003). Hormesis: The dose–response revolution. *Annual Review of Pharmacology and Toxicology* **43**: 175–197.

Carlsen, E., A. Giwercman, N. Keiding and N. E. Skakkebaek (1992). Evidence for decreasing quality of semen during the past 50 years. *British Medical Journal* **305**: 609–613.

Carson, R. (1962). *Silent Spring*. New York, Penguin Books.

Cheek, A. O., P. M. Vonier, E. Oberdorster, B. C. Burow and J. A. McLachlan (1998). Environmental signaling: A biological context for endocrine disruption. *Environmental Health Perspectives* **106**: 5–10.

Coggon, D., G. Rose and D. J. P. Barker (2003). *Epidemiology for the Unititiated*. BMJ Publishing Group.

Colborn, T., D. Dumanoski and M. J. Peterson (1996). *Our Stolen Future*. London, Abacus.

Commission of the European Communities (2003). Proposal for a Regulation of the European Parliament and of the Council concerning the Registration, Evaluation, Authorisation and Restriction of Chemicals (Reach), establishing a European Chemicals Agency and amending Directive 1999/45/EC and Regulation (EC) {on Persistent Organic Pollutants} {SEC(2003 1171}/* COM/2003/0644 final—COD 2003/0256 */**I**.

Duffus, J. H. (2006). Introduction to Toxicology. *Fundamental Toxicology*. J. H. Duffus and H. G. J. Worth. Cambridge, Royal Society of Chemistry: 1–17.

Eaton, D. L. and C. D. Klaassen (2001). Principles of Toxicology. *Casarett & Doull's Toxicology: the basic science of poisons*. C. D. Klaassen, McGraw-Hill: 11–34.

Eggen, R. I. L., B. E. Bengtsson, C. T. Bowmer, A. A. M. Gerritsen, M. Gibert, K. Hylland, A. C. Johnson, P. Leonards, T. Nakari, L. Norrgren, J. P. Sumpter, M. J. F. Suter, A. Svenson and A. D. Pickering (2003). Search for the evidence of endocrine disruption in the aquatic environment: Lessons to be learned from joint biological and chemical monitoring in the European Project COMPREHEND. *Pure and Applied Chemistry* **75**(11–12): 2445–2450.

Elwood, M. (1998). *Critical Appraisal of Epidemiological Studies and Clinical Trials*. Oxford, Oxford University Press.

ECHA (2007) Guidance for the Preparation of an Annex XV Dossier on the Identification of Substances of Very High Concern. Guidance for the Implementation of REACH. Helsinki, European Chemicals Agency: 58.

EPA (1997). Special report on environmental endocrine disruption: an effects assessment and analysis. Washington D.C., United States Environmental Protection Agency (EPA).

European Community (2000). Directive 200/60/EC of the European Parliament and of the Council of 23 October 2000 establishing a framework for Community action in the field of water policy: 72.

European Environment Agency (2001). Late Lessons from Early Warnings: The Precautionary Principle 1896–2000. Copenhagen, European Environment Agency.

European Union (2006). Regulation (EC) No 1907/2006 of the European Parliament and of the Council of 18 December 2006 concerning the Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH), establishing a European Chemicals Agency, amending Directive 1999/45/EC and repealing Council Regulation (EEC) No 793/93 and Commission Regulation (EC) No 1488/94 as well as Council Directive 76/769/EEC and Commission Directives 91/155/EEC, 93/67/EEC, 93/105/EC and 2000/21/EC. *Regulation (EC) No 1907/2006.* Brussels: 278.

Folmar, L. C., N. D. Denslow, K. Kroll, E. F. Orlando, J. Enblom, J. Marcino, C. Metcalfe and L. J. J. Guillette (2001). Altered serum sex steroids and vitellogenin induction in Walleye (Stizostedion vitreum) collected near a metroplitan sewage treatment plant. *Archives of Environmental Contamination and Toxicology* **40**: 392–398.

Francis, B. M. (1994). Water pollution, persistence, and bioaccumulation. *Toxic Substances in the Environment*. New York, John Wiley & Sons: 93–120.

Gee, D. (2006). Late lessons from early warnings: toward realism and precaution with endocrine-disrupting substances. *Environmental Health Perspectives* **114** (Suppl. 1): 152–160.

Greene, R. R., M. W. Burrill, and A. C. Ivy (1938) Experimental intersexuality: the production of feminized male rats by antenatal treatment with estrogens. *Science* **188**(2275): 130–131.

Gregus, Z. and C. D. Klaassen (2001). Mechanisms of toxicity. *Casarett & Doull's Toxicology: the Basic Science of Poisons*. C. D. Klaassen, McGraw-Hill: 35–82.

Guillette, L. J. and M. P. Gunderson (2001). Alterations in development of reproductive and endocrine systems of wildlife populations exposed to endocrine-disrupting contaminants. *Reproduction* **122**(6): 857–864.

Gundert-Remy, U., P. Kremers, A. Renwick, A. Kopp-Schneider, S. G. Dahl, A. Boobis, A. Oberemm and O. Pelkonen (2005). Molecular approaches to the identification of biomarkers of exposure and effect – report of an expert meeting organized by COST Action B15. November 28, 2003. *Toxicology Letters* **156**(2): 227–40.

Hale, R. C., M. J. La Guardia, E. Harvey and T. M. Mainor (2002). Potential role of fire retardant-treated polyurethane foam as a source of brominated diphenyl ethers to the US environment. *Chemosphere* **46**(5): 729–35.

Harvey, P. W., K. C. Rush and A. Cockburn, Eds. (1999). *Endocrine and Hormonal Toxicology*. Chichester, John Wiley & Sons.

Herbst, A. L., H. Ulfelder and D. C. Peskanzer (1971). Adenocarcinoma of the vagina: association of maternal stilbestrol therapy with tumor appearances in young women. *New England Journal of Medicine* **284**: 878–881.

Herrchen, M. (2006). Pathways and behaviour of chemicals in the environment. *Fundamental Toxicology*. J. H. Duffus and H. G. J. Worth. Cambridge, The Royal Society of Chemistry: 238–256.

Hertz-Picciotto, I. (1998). Environmental epidemiology. *Modern Epidemiology*. K. J. Rothman and S. Greenland. Philadelphia, Lippincott-Raven Publishers: 555–583.

Highnam, K. C. and L. Hill (1977). *The Comparative Endocrinology of the Invertebrates*. London, William Clowes & Sons.

Hill, A. B. (1965). The environment and disease: association or causation? *Proceedings of the Royal Society of Medicine* **58**: 295–300.

International Programme on Chemical Safety (2002). *Global Assessment of the State-of-the-Science of Endocrine Disruptors*. World Health Organisation: 180.

James, R. C., S. M. Roberts and P. L. Williams (2000). General principles of toxicology. *Principles of Toxicology: Environmental and Industrial Applications*. R. C. James, S. M. Roberts and P. L. Williams. New York, John Wiley & Sons: 3–34.

Kriebel, D., J. Tickner, P. Epstein, J. Lemons, R. Levins *et al*. (2001). The precautionary principle in environmental science. *Environmental Health Perspectives* **109**(9): 871–876.

Krimsky, S. (2000). *Hormonal Chaos. The Scientific and Social Origins of the Environmental Endocrine Hypothesis*. Baltimore, The Johns Hopkins University Press.

Langezaal, I. (2002) *The Classification and Labelling of Carcinogenic, Mutagenic, Reprotoxic and Sensitising Substances*. Ispra, European Chemicals Bureau, Joint Research Centre: 193.

Marchant, G. (2003). From general policy to legal rule: aspirations and limitations of the precautionary principle. *Environmental Health Perspectives* **111**(14): 1799–803.

MRC Institute for Environment and Health (Editor) (1997). *European Workshop on the Impact of Endocrine Disrupters on Human Health and Wildlife*. Brussels, European Commission: 125.

Newbold, R. R., R. B. Hanson, W. N. Jefferson, B. C. Bullock, J. Haseman and J. A. McLachlan (1998). Increased tumors but uncompromised fertility in the female descendants of mice exposed developmentally to diethylstilbestrol. *Carcinogenesis* **19**(9): 1655–63.

Newbold, R. R., R. B. Hanson, W. N. Jefferson, B. C. Bullock, J. Haseman and J. A. McLachlan (2000). Proliferative lesions and reproductive tract tumors in male descendants of mice exposed developmentally to diethylstilbestrol. *Carcinogenesis* **21**(7): 1355–63.

O'Flaherty, E. J. (2000). Absorption, Distribution, and Elimination of Toxic Agents. *Principles of Toxicology: Environmental and Industrial Applications*. P. L. Williams, R. C. James and S. M. Roberts. New York, John Wiley & Sons: 35–56.

Ottoboni, M. A. (1991). *What are chemicals? The dose makes the poison—A plain language guide to toxicology*. New York, Van Nostrand Reinhold: 5–18.

Rogers, M. D. (2003). Risk analysis under uncertainty, the Precautionary Principle, and the new EU chemicals strategy. *Regulatory Toxicology and Pharmacology* **37**(3): 370–381.

Rushton, L. (2000). Reporting of occupational and environmental research: use and misuse of statistical and epidemiological methods. *Occupational and Environmental Medicine* **57**(1): 1–9.

Slovic, P. (1987). Perception of risk. *Science* **236**(4799): 280–5.

Steenland, K. and D. A. Savitz, Eds. (1997). *Topics in Environmental Epidemiology*. New York, Oxford University Press.

Sullivan, F. M. (2006). Reproductive toxicity. *Fundamental Toxicology*. J. H. Duffus and H. G. J. Worth. Cambridge, Royal Society of Chemistry: 142–153.

Timbrell, J. (2000). Factors affecting toxic responses: metabolism. *Principles of Biochemical Toxicology*. London, Taylor & Francis: 65–112.

White, R., I. Cote, L. Zeise, M. Fox, F. Dominici and T. Burke (2009). State-of-the-Science Workshop Report: issues and approaches in low-dose–response extrapolation for environmental health risk assessment. *Environmental Health Perspectives* **117**(2): 283–287.

Williams, F. S. and S. M. Roberts (2000). Dermal and ocular toxicology: toxic effects of the skin and eyes. *Principles of Toxicology: Environmental and industrial Applications*. P. L. Williams, R. C. James and S. M. Roberts. New York, John Wiley & Sons: 157–168.

Zacharewski, T. (1997). In vitro bioassays for assessing estrogenic substances. *Environmental Science and Technology* **31**(3): 613–623.