Pure & Appl. Chem., Vol. 70, No. 9, pp. 1671–1683, 1998.
Printed in Great Britain.

Natural and anthropogenic environmental oestrogens: the scientific basis for risk assessment*

Principles of risk assessment

I. F. H. Purchase[1]† and G. L. P. Randall[2]

[1]*Director, Zeneca Central Toxicology Laboratory, Alderley Park, Macclesfield, Cheshire SK10 4TJ;* [2]*Director, Brixham Environmental Laboratory, Freshwater Quarry, Brixham, Devon TQ5 8BA, UK*

INTRODUCTION

Importance of risk assessment for risk management

There are a vast number of chemicals, both synthetic and natural, to which man and the environment are exposed. They may have beneficial effects—such as those conferred by vitamins—but all chemicals are toxic when given in a large enough dose. The range of toxicity is over 10^9 and thus there is a continuing need to define which chemicals represent a risk to health or the environment so that efforts can be directed to controlling those which present the greatest harm to the largest number of individuals.

In the particular context of chemicals which disrupt the endocrine system, there have been many observations of adverse effects occurring in humans or vertebrate species which have been linked to the presence of natural or synthetic chemicals in the environment, frequently without any evidence of causality. There are also observations of chemicals which have oestrogenic properties in human tissue, milk, food or the environment. It is often implied, frequently wrongly, that adverse effects are inevitable from the presence of these contaminants. The use of risk assessment can assist in resolving some of the uncertainties arising from observations of chemical contamination or adverse effects by providing a quantitative assessment of the relationship between the contaminant and the adverse effect.

For commercial organisations, the identification of chemicals which can be used safely is of paramount importance. For some chemicals, where the effect is dependant on the presence of the chemical (for example the treatment of disease with medicines), defining the conditions for safe use is important to protect the consumer. In other cases, the presence of the chemical in the consumer product is adventitious (for example the presence of chemicals in effluent, or the presence of pesticide residues, carcinogenic mycotoxins or oestrogenic compounds in food) and the issue is to what level must the chemical be controlled in order to preserve the environment and health.

The consumer is protected by the regulatory processes which are part of the legal framework for the registration of chemicals for certain uses (e.g. pharmaceuticals, pesticides, food additives, cosmetics or industrial chemicals). Equally the protection of public health requires the identification of toxic natural products and the introduction of the appropriate control measures to minimise risk. For most examples of regulatory action, the assessment of the risk posed by the presence of the chemical is one of the key steps in the decision making process. Eliminating chemicals completely, particularly natural products such as mycotoxins, is practically impossible and the development of high resolution analytical methods has demonstrated the presence of traces of such chemicals in virtually every product. The question is thus the same: To what concentration should the contaminant or natural product be limited in order to allow the benefits of the product (food, medicines etc.) to be realised without undue risk?

---

* *Pure & Appl. Chem.*, 1998, **70**(9)—an issue of special reports devoted to Environmental Oestrogens.

† *Current address*: School of Biological Sciences, Stopford Building, University of Manchester, Oxford Road, Manchester M13 9PT, UK

Definitions

Risk assessment is a tool used in risk management. It consists of several components—hazard identification, including knowledge of the dose response relationship, exposure assessment and finally the integration of these into a risk assessment.

Hazard is the inherent property of a chemical which gives it the potential for harm. The assessment of hazard includes a qualitative description of the type of adverse effect produced by the chemical (e.g. developmental toxicity, carcinogenicity etc.) and a quantitative description of the dose-response relationship in the species of concern. The second component of risk assessment is the assessment of **exposure**, which includes the level, frequency and duration of exposure. The assessment of risk places the potential for harm into context by asking the question 'How much harm is likely to occur when individuals or a population are exposed to a defined amount of a particular chemical?' Thus, **risk** can be considered as a quantitative statement of the likely harm which will be caused under particular circumstances of exposure to the chemical. It can be expressed as the likelihood that a harmful effect will occur in a population or the likelihood that a harmful effect will affect a single individual. Risk may also be expressed as a 'Margin of Safety' (MOS) derived from a comparison of the level of exposure and the No Observed Adverse Effect Level (NOAEL). Thus the MOS is the NOAEL divided by the exposure level. Ideally these expressions are in quantitative terms (such as 1 in a thousand individuals exposed will suffer the condition, or, the population of a particular organism will be reduced by a certain percentage). However, more frequently, because of the limitations in our knowledge, the risk is considered inconsequential below a certain exposure level or dose. In the latter case, the output of risk assessment is the setting of a standard which will protect the health of the individual or the environment.

The data available for risk assessment is obtained from laboratory experiments or from observations in the field for assessment of hazard to both man and the environment. The design of the laboratory experiments, field studies and epidemiological investigations is complex and requires particular expertise. This is even more important for the integration of hazard information and exposure data which is the crucial step in risk assessment. In many cases assumptions have to be made (e.g. that the results of the laboratory experiments can be used to inform decisions about risk to humans or the environment; that results from one length and frequency and duration of exposure are relevant to another set of circumstances). It is thus usual to rely on the judgement of experts or a committee of experts from the various fields in order to ensure that the best risk assessment is obtained.

EXPOSURE

The two key pieces of information required for a quantitative risk assessment are knowledge of the level of exposure and of the hazard posed by the chemical. The exposure in laboratory experiments is closely monitored and controlled. However, exposure of humans and other species is usually intermittent and of variable duration and frequency and different for each individual. Particularly for field studies, including human epidemiological studies, the least reliable information is often that of the level of exposure.

Consideration of dose, frequency and duration

The estimation of the dose, for the purposes of risk assessment of a chemical, is complicated by the different rates and routes of metabolism and differences in kinetics that are often seen between species and even between routes of administration in the same species. As a general rule the concentration of the chemical (or its metabolite, if that is the proximate toxin) with respect to time in the affected organ or cell (the tissue dose) is considered to be the critical measure of exposure. Such variables as the route, magnitude, frequency and duration of the administered dose and the species under consideration may all have a substantial effect on the tissue dose and hence may confound any attempt at risk assessment based on measured exposure. The area-under-the-curve (AUC) of the graphical representation of the concentration over time is considered to be the best representation of dose or exposure. Accurate

estimation of tissue dose for non-pharmaceutical chemicals is rarely available and thus simple estimation of exposure is often used in risk assessment.

Exposure and dose are often used synonymously in risk assessment. There is a tendency for the amount of a medicine administered to be considered as the dose, while the tissue level is described as the tissue exposure. For environmental chemicals, exposure is the concentration to which the organism, whether human or otherwise, experiences over a set period of time, while the dose is the amount reaching the tissue.

The methods of providing accurate estimations of exposure are particularly difficult for a problem as diffuse as 'environmental oestrogens' (Kavlock & Ankley, 1996). While it is known that there are natural oestrogens in, for example, food and synthetic chemicals in various environmental compartments, little work has been done to collect comprehensive data of actual exposure levels. Two examples of risk assessment are given in pp. 1845–1851 and pp. 1853–1860 of this book where methods of assessing exposure are very different. Some of the synthetic chemicals which are of particular interest as posing a potential risk because of their oestrogenic properties are persistent; exposure can be assessed by direct measurement or by estimation of the residues present in human tissues (e.g. PCBs or DDT metabolites in fat). For others which are less persistent exposure estimation, particularly retrospectively, may present particular problems.

Extrapolation from laboratory experiments to environmental exposure

Laboratory experiments usually use much larger doses than are encountered from environmental exposure. The reason for this practice is that many toxic events at environmental concentrations occur at a low frequency which would not be observed with the relatively small numbers of organisms (or animals) used in toxicity studies. An increase in dose is used to compensate for this by increasing the frequency of those events. In conducting mammalian toxicity studies a high dose is often employed to ensure that a toxic response is observed, at least at the top dose, so that the target organ and type of toxicity are identified.

A consequence of this practice is that toxicity has to be predicted at doses from 100 to 1,000,000 times lower than those used in the experiment. This is particularly important for toxic endpoints such as cancer, which occur at relatively low frequency at environmental exposure levels. It is usual to consider that the laboratory experiments will provide data which can be used for low dose extrapolation unless there is mechanistic evidence to the contrary.

Natural and synthetic chemicals

The majority of chemicals tested for toxicity are synthetic. However, the risk assessment of natural chemicals follows the same principles.

Mixtures

Environmental organisms and humans are always exposed to a mixture of chemicals, because they live in an environment rich in chemicals be they endogenous, of exogenous natural origin (i.e. a food) or man made. This raises the question of whether additive effects might occur or whether synergy could magnify the toxicity of the chemicals under study.

Additive effects occur when more than one chemical cause the same toxicity through the same mechanism of action. A good example is found with the isomers of dioxin, where the toxicity of the mixture of isomers is the sum of the amount and potency of each of them—all the isomers are assumed to act in a similar way and to produce the same type of toxicity. In general, compounds acting through a receptor or by inhibiting an enzyme would be expected to act in this way. Synergy occurs when the toxicity caused is greater than the sum of the toxicity of each chemical in the mixture. In principle, this may occur if one chemical increases target site concentration of another by increasing the absorption or decreasing the metabolic degradation. It may also occur if the chemicals act at different stages of the

same toxic pathway. In the majority of cases, the toxicity of two chemicals is independent, producing neither additivity nor synergy.

The administration of two oestrogenic chemicals, dieldrin and endosulfan, simultaneously was reported to result in synergistic activation of the oestrogen receptor (Arnold *et al.* 1996). If this were a general rule, it would have important implications for risk assessment as traces of many chemicals which might have oestrogenic properties are found in the environment and in human tissues. However, the results have not been confirmed (Ashby *et al.*, 1997a; Ramanoorthy *et al.*, 1997) and the authors of the original report have subsequently withdrawn it (McLachlan, 1997). Thus risk assessment of the oestrogenic effects of mixtures can reasonably rely on an additive model.

Special considerations for environmental organisms

The risk posed to an ecosystem is a very complex issue. The first factor that must be addressed is: 'which ecosystem are the assessments to be carried out for?'. They can be global as in the case of chloroflurocarbons or local as in the case of chemicals affecting terrestrial habitats and small streams.

The ecosystem focus then poses the next challenge: ' which organisms should be used for hazard assessment in the Laboratory?'. In a small pond ecosystem you would expect to find at least 200–2000 distinct organisms in addition to the inevitable microbial population. The issue would be similar in a terrestrial ecosystem such as a field/hedgerow combination. The individual organisms themselves would be in different parts of their life cycle according to the seasons. These life stages may have dramatically different exposure profiles, e.g. exposure of sediment dwelling larvae differs from that of flying gnats even in the case of the same species such as the chironomids. Individual weather events will change the relative proportions of organisms in any ecosystem and their location in the ecosystem, e.g. earthworms who inhabit deep soil profiles in drought are on the surface during heavy rainfall.

Every ecosystem will be an integrated food web. Thus a toxic effect in a lower level organism could have dramatic influence on the population at a higher level through starvation. It may be difficult to disentangle cause and effect in such integrated systems.

The actual dose that an individual organism receives will not be a simple factor of the measured concentration in a given environmental compartment. This is because the concentration at site of action will depend on the bioavailability of the substance from its matrix, e.g. organic matter in soils and sediments does have a significant impact on the availability of substances to organisms living in or ingesting these materials. In addition regular changes in some ecosystems modify the physiology of the organism with respect to its ability to absorb toxicants, e.g. crabs will have different susceptibility to heavy metals depending on the salinity of their environment, which will change with the tidal cycles and the organism's natural history/migration.

There are two other major factors that influence dose. These are bioaccumulation and biomagnification. The former is the ability of an organism to concentrated substances from its environment. The latter is the sequential concentration of a substance as organisms consume each other in a food web.

Special considerations for human exposure

Route of exposure: When considering the overall toxic burden of a chemical, the amount entering the body through any route should be summed. Thus a chemical may enter the body via oral ingestion, inhalation and skin absorption and these different amounts should be added together to give the total body burden. The human population is exposed to the majority of environmental chemicals by ingestion, in the food or drink. This generalisation appears to be true for environmental oestrogens also.

Risk assessment of exogenous synthetic oestrogenic chemicals must take into account that there are many naturally occurring chemicals which have oestrogenic properties (such as genistein in Soya) to which the population is likely to be exposed simultaneously. Endogenous oestrogen and oestrogens used for medicinal purposes, which will activate the same receptors, may confound the assessment.

Metabolism: In many cases, chemicals (both natural or endogenous and synthetic or exogenous) are metabolised in the body, particularly in the liver. The first phase of metabolism—an oxidative step—is catalysed by cytochrome P450 enzymes. The second—conjugation—produces a more water soluble product which is excreted. It is not uncommon for oxidative metabolism to increase the potency of the toxic effect, even converting a non-toxic chemical into a *proximate toxin*, responsible for the toxic effects of the ingested chemical.

The endogenous synthesis of steroid hormones is a complex multi-stage process converting cholesterol into di-hydro-testosterone or oestradiol, depending on the activity of the intermediate enzymes. There are many ways in which xenobiotics can affect the rate of synthesis in one or more steps, thereby producing an indirect oestrogenic or anti-oestrogenic effect (Fig. 1).

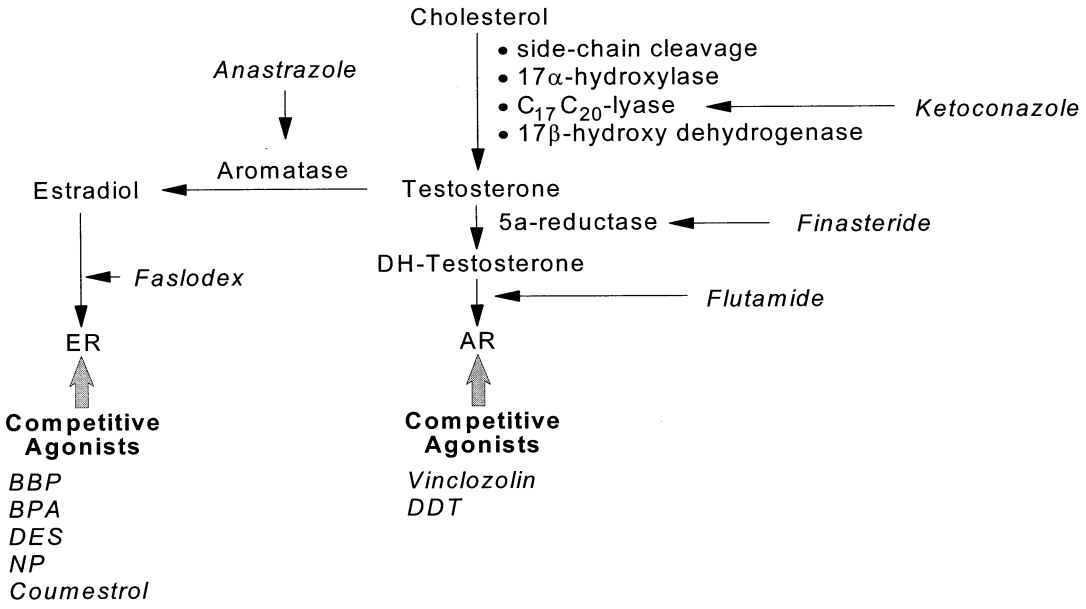

Fig. 1. Metabolic pathways for the biosynthesis of oestradiol and testosterone and the mechanism of action of various compounds.

A further important implication of the knowledge of metabolism of the xenobiotic is that *in vitro* cell based systems virtually always have different rates and sometimes routes of metabolism than is seen *in vivo*. This is no less important for oestrogenic activity than for other classes of chemicals (Ashby *et al.*, 1997c), with examples already identified where differences between *in vitro* and *in vivo* results can be attributed to the difference in metabolism.

Inter-species extrapolation of exposure: Using the simple rule that animals are equivalent to humans for the purpose of risk assessment often leads to large errors in the computed risk. The difference in body mass between animals and humans can contribute to these errors, particularly for chemicals for which the metabolism plays an important role in the potency of the toxic effects. One approach to overcoming this problem is to use allometric scaling, a method which takes account of the animal's body mass in adjusting the dose used in risk assessment. It is based on the observation that many biological parameters increase with 0.75 power of the body mass. Thus, when animal data are being used for human risk assessment, the margin of safety is multiplied by the scaling factor to accommodate the differences in body mass. The scaling factors are given in Table 1.

For chemicals which are not metabolised, simple equivalence of the dose in mg/kg may be used. When chemicals are given in the diet or drinking water, the consumption of the diet and drinking water increases with the 0.75 power of the body mass and adjustment is not required.

Table 1. Scaling factors for body weight (based on BW 0.75)

Species	BW (kg)	Scaling factor
Mouse	0.025	7.1
Mouse	0.050	6.0
Rat	0.200	4.3
Rat	0.250	4.0
Rat	0.300	3.8
Guinea-pig	0.500	3.4
Dog	10	1.6
Dog	15	1.4

The similarity of the oestrogen receptor and hormones activating it between species provides an indication that oestrogenic chemicals may have similar actions in various species of interest. However, there are now three oestrogen receptors identified (α, $\beta 1$ and $\beta 2$) and at least two response elements on the DNA. This may explain the diversity of responses between organs and species for some therapeutic oestrogens and anti-oestrogens. It also provides a caveat to the simple assumption that all oestrogens will act in the same way in all organs and all species.

Physiologically-based pharmaco-kinetic modelling: The most detailed method available for calculation of dosage for comparison between species is known as *physiologically-based pharmaco-kinetic* (PBPK) modelling. This method takes account of the route and rate of uptake, metabolism and excretion and the species-specific physiological parameters which are important in the distribution of the chemical around the body. PBPK modelling can provide good predictions of the concentration in tissues of interest of a chemical or its metabolite when enough data on the kinetics of its metabolism are available. In most cases there is insufficient information on the chemical of interest to allow the use of this method.

Measurement: In most cases, it is possible to estimate the concentration of the chemical of interest in the species of interest. Increasingly, human volunteer studies can be used for obtaining such data, provided they are carried out with appropriate ethical control. Because the analytical methods are so sensitive, such studies can be conducted with very low doses of chemical and thus without appreciable risk to the volunteers.

Biomarkers: For some chemicals, particularly those which are reactive, it is possible to estimate the body burden from measurement of reaction products within the body. For example, haemoglobin adducts of chemicals such as ethylene oxide can be readily measured, as can the levels in fat of chemicals which accumulate in lipid, such as dioxins. These methods for estimating dosage in epidemiological studies are used infrequently and may be difficult to interpret in retrospective studies where the observations are being made many years after exposure.

For some of the persistent organic lipophilic chemicals, such as DDT and dioxins, metabolites (DDE) or original chemicals (dioxins) with long half lives may be measured in the tissues of individuals, providing an index of average exposure over considerable periods of time. There is usually significant difficulty in assessing the dose which would be required to produce the tissue levels because of complex kinetics. When carrying out risk assessment using animal data, differences in kinetics can be critical; for example, the kinetics of dioxin in rats and humans are very different, presenting real difficulty in extrapolating the dosimetry between the species.

Estimation of exposure through food: As the majority of exposure to xeno-oestrogens occurs through food or water, methods of estimating exposure of human populations to food-borne chemicals will be important for risk assessment. In many countries there is excellent information on the dietary consumption of the population, based on extensive surveys. It is thus possible to estimate the distribution

of intake of particular foods, and their components, in the general population. Once the data on consumption are collected, it is relatively easy to calculate intakes based on analytical data for particular chemicals. An alternative approach is to collect 'food from the plate', which will include the effects of preparation. It is however a more cumbersome technique.

When dealing with the total intake of oestrogens, both naturally occurring and anthropomorphic chemicals must be taken into account. There is good evidence for substantial intake of naturally occurring oestrogenic chemicals in foods such a Soya and for considerable exposure of farm and wild animals to oestrogens such as those in clover (see pp. 1853–1860) or oestradiol found in UK rivers.

HAZARD ASSESSMENT

The next component of the risk assessment process is the assessment of hazard. As has been pointed out above, hazard can be considered as the inherent toxicity of a chemical. It is often described as 'potential toxicity' in the sense that any results from laboratory experiments only provide information that there is a potential for toxicity in the target species; there are many reasons, such as lack of susceptibility of the target species, different metabolism, inappropriate laboratory experimental design or instability of the chemical under field conditions which may result in no toxicity. It is important to distinguish hazard from risk, something which is often overlooked in media reports.

A more detailed description of the methods for assessing oestrogenic hazard are found in the section on risk assessment.

Toxicity and adaptation: The response of an organism to a chemical may represent an adaptive process or a toxic response. For example, some chemicals can increase the activity of enzymes in the liver with a resulting increase in liver weight, an adaptation to the enzyme induction. Similarly, environmental organisms will migrate away from some harmful influences, change their feeding preference to minimise harm or may even adopt a less vulnerable morphology, e.g. encysting in certain invertebrates.

Structure/function analysis

Various features of chemical structure, such as electrophilicity or the three-dimensional shape of the molecule, determine its mode of toxic action. It has always been an aspiration of toxicologists to predict the toxicity of chemicals by an analysis of their structure. For some types of toxicity, for example carcinogenicity and mutagenicity, it has proved possible to identify potentially active compounds by virtue of their electrophilicity. Even so, many issues such as metabolism and kinetics can be difficult to predict and there is thus a false prediction rate. Certain other aspects of toxicity, such as the ability to penetrate the skin, or the likely site of metabolism which can have an important effect on the expression of toxicity, can be predicted with reasonable success. Prediction in most cases is purely qualitative. It has prove to be very difficult to predict the site, type and potency of toxic responses on the basis of chemical structure alone (Dearden *et al.*, 1997). Greater success is obtained with homologous series of chemicals and single toxic endpoints (e.g. eye irritation, Barratt, 1997). For the particular case of oestrogenic chemicals, the prediction of activity on the basis of structure may be complicated by the presence of three different receptors and the fact that chemicals may have oestrogenic activity by virtue of their influence on the enzymes producing or degrading oestrogens or androgens. Multiple mechanisms will inevitably lead to multiple structure activity relationships (Ashby *et al.*, 1997c). A more detailed description of the complexity of the function of the endocrine system controlling reproduction is given in pp. 1633–1646, 1647–1656 and 1657–1669 and of the structure function relationships in pp. 1725–1733.

In the field of environmental hazard assessment, some success has been achieved by relating physico-chemical properties to structure: for example melting point, boiling point and to a lesser extent water solubility to partitioning behaviour. There are also a wide range of, in most cases, narrowly applicable relationships predicting ecotoxicity from structure or a physico-chemical property. The main concern about these relationships is the extent to which they can be applied to new and/or multifunctional chemicals. The latter, in particular, become harder to assess as unknown interactions alter the expected behaviour of the chemical.

Hazard identification: environmental studies

Epidemiology: The best subject for study of human diseases is the human population. Ultimately, results from epidemiological studies provide the critical information about the causal link between exposure and disease. These studies are expensive and time consuming to carry out and difficult to interpret, primarily because of the complexity of the causation of disease and the limited statistical power of all but the most extensive epidemiological studies. A further difficulty is the assessment of the exposure to the chemical of interest.

The simplest human studies are *case reports*, in which observations of particular diseases and the history of the patient are reported. They have been responsible for alerting us to a number of important causes of diseases from occupational or therapeutic exposure. Thus the early reports of babies with phocomelia which linked the condition to the administration of thalidomide provided the first evidence of the teratogenic effects of thalidomide in humans; subsequent formal studies confirmed that association. Similarly, the first reports of angiosarcoma of the liver in employees exposed to vinyl chloride (following earlier reports of angiosarcoma in rats exposed to vinyl chloride) provided the basis of formal cohort studies which demonstrated the carcinogenicity of vinyl chloride to humans.

There are two principle methods of studying the association between exposure to a chemical and the incidence of a disease. These methods begin with either the disease or the exposure. The *case controlled study* examines a number of cases of the disease and compares information on the potential causes of the disease with that obtained from a matched group of healthy individuals or patients suffering from an unrelated disease. Thus, one could compare the drug treatment of mothers of children with phocomelia with the treatment of mothers of healthy children or children with an unrelated disease. Case controlled studies have the advantage that it should be possible to obtain a good definition of the group of cases, particularly with respect to the accuracy of the diagnoses of the disease being studied. However, there may be real problems with identifying reliable indices of exposure and of avoiding bias or confounding influences. It is generally considered that case controlled studies are hypothesis forming studies and that the hypothesis will require to be confirmed by formal cohort studies.

Cohort studies compare the incidence of a disease (for example of lung cancer) in a cohort of individuals with a particular attribute (for example, a group people who have smoked) with a reference group matched for the variables known to affect the incidence of the disease (for example, matched for age, which is known to have a large effect on the incidence of cancer). Often national statistics of disease will be used for reference, which may raise the question of the relevance of the national disease pattern for the particular local cohort being studied. Usually cohort studies are *retrospective*, in that the exposure of the cohort occurred some time previously and the members of the cohort are identified from records. However, *prospective* studies have been made and have the advantage that appropriate diagnostic and exposure information can be collected. Well carried out studies of sufficient size are time consuming and difficult to conduct but provide the best information on the association between exposure and a disease. The main problems with these studies, particularly retrospective studies, arises from the difficulty in obtaining accurate information on exposure and in the follow up of the cohort with respect to accurate identification of the incidence of the disease in question. A further problem can arise from multiple comparisons, where statistically significant differences in the incidence of a disease may occur by chance if too many diseases are studied without there being an hypothesis to test (for example, studying the cancer incidence in a large cohort may identify 50 or more types of cancer which occur in the cohort, with a high probability that some will be significantly different by chance). Two of the most common criticisms of epidemiology study methodology are the influence of bias and of confounding.

Bradford Hill criteria. The interpretation of epidemiology studies is often controversial, because of the complexity of the causation of disease and the difficulty of studying isolated causes in the complex social environment. Bradford Hill described 9 criteria which could be used to test the interpretation of epidemiology studies. These criteria are: 1 Strength of association, i.e. the stronger the association, the more is the inference of causation justified; 2 Consistency, i.e. repeated observations made in different locations by different investigators; 3 Specificity, i.e. disease rare in the absence of exposure occurs frequently in the presence of exposure; 4 Temporality i.e. exposure precedes disease; 5 Biological

gradient i.e. dose response relationship observed; 6 Plausibility i.e. causation is biologically plausible; 7 Coherence i.e. epidemiological association does not conflict with known facts; 8 Experiment i.e. evidence of an effect of intervention, for example; 9 Analogy i.e. that a similar association has been seen before.

The Bradford Hill criteria have been applied to the overall evaluation of the relationship between exposure to oestrogenic chemicals and the presence of adverse effects in humans (Ashby *et al.*, 1997a). There is insufficient evidence to accept that such a relationship is causal. The use of these criteria for assessing the overall evidence of causation in the many examples of suspected oestrogenic effects is to be encouraged.

Problems of exposure assessment. Particularly for case controlled studies and for retrospective cohort studies, the accurate quantitative assessment of exposure presents significant problems. Usually, surrogates of exposure are used, such as job title or occupation, for example a process worker is assumed to have higher exposure that an office worker in the same company. Often the information related to exposure is obtained by questionnaire, sometimes from relatives. All of these methods suffer from problems of accuracy and consistency and the inability to confirm that the information used to study a particular association is reliable. However, there is sometimes the possibility of using biological markers of exposure, for example the presence of DDT and its metabolites in fat as an index of integrated exposure, and this provides a much more satisfactory estimate.

Environmental studies. Environmental field studies are fascinating challenges for biologists because it is impossible in a pragmatic study to monitor the effect on each organism. This needs the very careful selection of representatives as key indicators of change. It will rely on an understanding of the diurnal, temporal and seasonal variations of the chosen representatives. Controls are always used but their value relies on the often imperfect understanding of the base line variation in a given ecosystem.

The choice of ecosystem is also a challenge. For example, the species variety in a prairie slough (pond) in Canada and will be very different to a pond on a Swedish hill farm. How do you select a representative?

As with all field observations control is limited compared to the laboratory and incidents can cause major upsets. For example effects on a pair of blackbirds feeding selectively in a study of a field examining the effect of an agrochemical on soil invertebrates.

Experimental studies for human hazard assessment

The choice of experimental species is made on pragmatic grounds from amongst those which are commonly used and usually involves at least one rodent (usually the rat) and one non-rodent (usually the dog) species, with special studies relying on particular choices (e.g. sensitisation studies often use guinea-pigs). For discussion see Ashby *et al.*, 1997b.

The route duration and frequency of exposure during the experiment is based on the expected route and frequency of exposure of the exposed human population. For environmental chemicals either administration through the diet or drinking water or through inhalation administration would be selected based on the chemical and circumstances of exposure.

Experimental studies for environmental risk assessment

The choice of organism for hazard assessment is driven by several factors. The choice needs to relate to the ecosystem for which the eventual risk assessment seeks to protect. The organisms chosen need to be reliable surrogates of those under potential challenge. Also the very pragmatic issue of the ability to culture the organism reliably in the laboratory. Sometimes there is an interplay amongst these factors, e.g. in the aquatic testing of invertebrates *Daphnia magna* has been selected because, amongst other good features, it is available in clonal populations. This gives rise to excellent controls but a very poor ability to resolve issues associated with endocrine disruption and its impact on sexual reproduction.

On an international basis there is generally good agreement on the organisms that are useful surrogates. However, national and sub-national regulators may insist on studies that are carried out on indigenous species and in characteristic ecosystems.

The complexity of the matrix in an ecosystem is difficult to replicate in a reproducible way. For instance the use of 'standard' soils is favoured by some regulators and academic scientists. In these soils there is great debate about how the microbial population is represented if at all. In aquatic systems clean water, that is without sediments, particulates and high molecular weight components such as humic acids often give results that are very different to the natural environment for which the risk assessment is targeted.

Good laboratory practice (GLP)

All data produced for submission to regulatory authorities have to be carried out according to international GLP regulations. The purpose of these regulations is to ensure that the experiments are carried out reliably and that the reports of the work reflect the conduct of the experiment. Compliance involves careful recording of the details of the experimental conduct and protocol and quality assurance of the report. Differences of view about the importance of experimental design and conduct have been discussed recently (Ashby & Odum, 1998; Welshons, Nagel & vom Saal, 1998)

RISK ASSESSMENT

Numerical estimation of risk vs. setting standards: Laboratory experiments are usually carried out a doses much higher than those represented by the usual conditions of exposure of man or environmental organisms. When considering how to assess the risk associated with exposure much lower than that used in experiments, the way in which the severity or incidence of toxicity reduces with reducing dose level is critical. For most manifestations of toxicity, there is good evidence based on observation and studies of mechanisms of toxicity to conclude that there is a threshold below which no toxic effect will be seen; thus the method used for assessment of low dose risk relies on the concept of a threshold and provides 'safety factors' below the observed threshold to set 'safe' or acceptable standards of exposure. However, it has been argued that for cancer induced by a genotoxic chemical (i.e. a chemical which induces cancer as a consequence of damage to the genetic material leading to mutations which result in increased cell growth), there will be no threshold. This is because a single molecule of a genotoxic chemical could theoretically react with DNA to produce a growth enhancing mutation. It is also argued that, if the mechanism for the induction of cancer by the genotoxic chemical is the same as the mechanism for causing cancer in the control animals, than any dose of the chemical which adds to the mechanism will add to the incidence of cancer. What has not been factored into these arguments is that our knowledge of mechanisms of induction of cancer, although they have improved substantially in the last few years, are not sufficiently precise to establish whether the mechanisms of chemically induced cancer are identical with those of cancer in control animals. The repair mechanism in every cell can repair damage caused by genotoxic insults and this may also result in departure from a linear non-threshold dose response.

Many of the mechanisms which result in adverse effects (see Figure 1) associated with the endocrine system are similar to those which cause toxicity by other mechanisms. For example, the inhibition of enzymes involved in the metabolism or production of natural oestrogens can lead to an increase or decrease of the relevant hormone. In these examples the existing paradigm for low dose extrapolation should be adequate. However there remains considerable uncertainty about the appropriate methods for extrapolating effects mediated by receptor agonism or antagonism. From a biochemical point of view, it is likely that the interaction between a receptor and its ligand will be linear over a considerable dose range, particularly at low doses; but from a biological point of view it is clear that low levels of binding can be without significant biological effects. Thus the paradigm which is most likely to be effective in risk assessment of oestrogenic chemicals will encompass a threshold.

Hormesis: There is evidence that some dose-response curves are U-shaped, that is that the first response of the organism to the chemical stress is stimulatory, followed by an inhibitory, or toxic, response. This phenomenon may be observed in mammalian (Calabrese & Baldwin, 1997) or non-mammalian studies (Stebbing, 1997). The implications for risk assessment are considerable in that the presence of hormesis implies that at doses lower than the NOAEL beneficial results are manifest. Similarly, low dose extrapolation that relies on linearity will provide an incorrect assessment of risk if hormesis occurs. For hormone effects, such as oestrogenic effects, low doses in the physiological range are likely to be beneficial to the organism; however, higher doses may have adverse effects. Risk assessment will have to take these effects into account (Anon, 1998).

NOAEL/SF approach: For the majority of chemical toxicities, the method favoured for risk assessment relies on the establishment of a No Observed Adverse Effect Level (NOAEL) in the laboratory experiment and than applying a Safety Factor (SF) to establish the acceptable standard. A satisfactory safety factor is usually considered to be 100, derived from a 10-fold factor for inter-species extrapolation and a 10-fold factor for inter-individual variation. Larger safety factors may be used if the toxic events are considered severe (e.g. reproductive effects), if the experiments do not allow the NOAEL to be established unequivocally or if the exposed population is particularly susceptible.

Linear extrapolation methods: In these methods, the incidence of the toxic event (e.g. cancer) at low doses is calculated by using an equation and entering the data from the animal experiment. There are a number of different equations that can be used, but the most frequently favoured equation is the linear multistage equation. The result of these methods of risk calculation is to provide a point estimate of risk at a particular dose or exposure. It is generally accepted that cancer risks lower than 1 in a million do not warrant regulatory intervention.

A further development of risk assessment methods relies on the calculation of a 'benchmark' dose (i.e. the dose which is calculated to cause a particular incidence of the toxic event; often a 5% or 10% incidence as this is quantitatively similar to a NOAEL). A safety factor or linear extrapolation from this dose can then be carried out to provide the appropriate standard.

There are complex arguments for and against these methods; the most commonly used method internationally to set acceptable standards is the NOAEL/SF approach.

Environmental risk assessment: The usual approach for chemicals starts with hazard identification, often involving three tropic levels, viz. vertebrate, invertebrate and producer, followed by definition of the dose-response relationship. Depending on the type of studies used, safety factors are then applied to obtain a predicted no-effect concentration (PNEC). The longer the duration of the studies and the more they address sensitive parts of the organisms' life cycle, the lower the safety factors. The choice of the organisms would, if not prescribed, be very difficult. The use of *Daphnia magna* as a surrogate for all invertebrates is obviously a gross simplification. However, the alternative of trying to test all invertebrates is clearly inappropriate. Other issues that may need to be considered when assessing the PNEC of a chemical include the potential for environmental damage on species other than that under test and the scope for recruitment of organisms back into the environment being considered. This could be of particular interest when addressing the risk assessments associated with spills, when the major effect is frequently of a short term nature and once the chemicals are removed, organisms are able to migrate back into the previously disturbed area.

The assessment of the environmental exposure, known as the predicted environmental concentration (PEC), is frequently prone to assumptions and estimations. This is always the one area that needs improvement when evaluating the quality of a risk assessment.

Once both the PNEC and PEC are available, then a risk characterisation is possible. Although this may be a simplistic ratio of the two, there are other approaches, including the probabilistic assessment of effects and exposure, which can lead to the quantitative assessment of risk. Additional factors which have an impact on the risk assessment are the goals of risk management, including the type of environment under consideration, the extent to which the function or the structure of the environment needs to be assessed and whether populations or individuals are to be protected.

The need for the process to be thorough and take all information into account: One of the most frequent criticisms of risk assessment is that all the information on the toxicity, physico-chemical properties, mechanisms of toxicity, inter-species differences and other factors are not taken into account when considering a particular case. It is particularly important for the risk assessment to include a statement of the assumptions that are used, the reasons for the acceptance or otherwise of the data incorporated into the risk assessment and the arguments on which the integration of all the available data is based. It is equally important that methods should be selected to give equivalent safety for all the risks, so that the public and those responsible for risk management can make the right choices about how to deploy resources in risk management.

RISK MANAGEMENT

The purpose of risk assessment is to provide information for risk management. There is no defined method of risk management which applies to all chemical risks. However the principles of risk management include a number of common themes.

1. The first of these is risk evaluation or cost-benefit assessment, where the output of the risk assessment is evaluated in relation to the benefits. This can be a judgement of an expert group or it can be a formal calculation of the monetary value of risks and benefits using certain rules. Formal risk benefit analysis allows resources for risk management to be allocated to control the greatest risks.
2. There are several ways in which risk assessments are used;

a. Priority setting. When there are many chemicals involved or when there are many risks, risk assessment can provide information which helps to identify those risks, or those chemicals, which should be tackled first. A typical example is in the selection of a chemical among candidates for development in industry.

b. Defining use pattern. Information on risk is used to decide whether the risks associated with a particular use can be justified. For example, the use of propellants in medicinal aerosol inhalers versus their use in aerosol dispensers of household products.

c. Providing information for labelling. The user is expected to obtain information on the safe methods of use and of the hazards from use through information on the label. There are many classification schemes which help to provide clear information for users so that they can be aware of the hazards and risks of using a particular product.

d. Providing impetus for substitution. Where a chemical has a relatively high risk, substitution is one measure that can be taken. The US EPA safer pesticide policy relies on this paradigm.

e. Setting standards. The result of much of the activity under the broad definition of risk assessment results in the definition of an acceptable level of exposure, for example the Acceptable Daily Intake for pesticide residues.

f. Informing the risk manager and public of the magnitude of risks. The output of risk assessment is often used for this purpose.

The process of risk assessment is often enshrined in legal instruments giving statutory powers for the control of the use of various classes of chemicals. In addition, internal controls in various industries rely on risk assessment to achieve the standards expected by the companies.

REFERENCES

Anon (1998). Toxicological implications of hormesis. *BELLE Newsletter*, **6**, 1–33 and **7**, 1–35.

Arnold SF, Klotz DM, Collins BM, Vonier PM, Guliette, LJ, McLachlan JR & McLachlan JA (1996) Synergistic activation of estrogen receptor with combinations of environmental chemicals. *Science* **272**, 1489–1492.

Ashby J, Harris CA, Lefevre PA, Odum J, Routledge EJ & Sumpter J (1997a) Synergism between synthetic oestrogens? *Nature* **385**, 494.

Ashby J, Houthoff E, Kennedy SJ, Stevens J, Bars R, Jekat FW, Campbell P, Miller JV, Carpanini FM & Randall GLP (1997b) The challenge posed by endocrine-disrupting chemicals. *Environmental Health Perspectives*, **105**, 164–169.

Ashby J & Odum J (1998) The importance of protocol design and data reporting to research on endocrine disruption. *Environmental Health Perspectives*, **106**, 315–316.

Ashby J, Odum J, Tinwell H & Lefevre PA (1997c) Assessing the risks of adverse endocrine-mediated effects: where to from here? *Regulatory Toxicology and Pharmacology* **26**, 80–93.

Barratt MD (1997) QSAR's for the eye irritation potential of neutral organic chemicals. Van Zutphen LFM and Balls M, Eds. Animal alternatives, ealfare and ethics. *Proceedings of the 2nd World Congress on Alternatives and Animal Use in the Life Sciences*, pp. 731–736.

Van Zutphen LFM & Balls M Eds. Animal alternatives, welfare and ethics. Proceedings of the 2nd World Congress on Alternatives and Animal Use in the Life Sciences pp. 731–736.

Calabrese EJ & Baldwin L A (1997) The dose determines the stimulation (and poison): development of a chemical hormesis database. *International Journal of Toxicology*, **16**, 545–559.

Dearden J C, Barratt M D & 10 others (1997) The development and validation of expert systems for predicting toxicity. *ATLA* **25**, 223–252.

Ramamoorthy K, Wang F, Chen I-C, Norris JD, McDonnell DP, Leonard LS, Gaido KW, Bocchinfuso WP, Korach KS & Safe S (1997) Estrogenic activity of a dieldrin/toxaphene mixture in the mouse uterus, MCF-7 human breast cancer cells and yeast-based estrogen receptor assay. *Endocrinology* **138**, 1520–1527.

Stebbing ARD (1997) A theory for growth hormesis. *BELLE Newsletter*, **6**, 1–11.

Welshons WV, Nagel SC & vom Saal FS (1998) The importance of protocol design and data reporting to research on endocrine disruption—response. *Environmental Health Perspectives*, **106**, 316–317.