Pure Appl. Chem., Vol. 75, Nos. 11–12, pp. 2143–2149, 2003.
© 2003 IUPAC

Topic 3.12

Determination of acceptable exposure levels for humans for endocrine active substances: Use of animal models*,†

Penelope A. Fenner-Crisp‡

ILSI Risk Science Institute, One Thomas Circle, NW, Ninth Floor, Washington, DC 20005-5802, USA

Abstract: Regulatory authorities and other scientific organizations around the world have developed hazard/risk assessment practices that involve the derivation of numerical values said to represent "acceptable" or "safe" levels of human exposure to individual chemical substances. Human data would be preferable for this purpose since interspecies extrapolation would not be necessary. In most cases, however, these data are not available or are inadequate for this purpose. Therefore, results from studies conducted in non-human mammalian species are used as the principal or sole basis for hazard/risk assessment. Existing risk assessment frameworks developed for the evaluation of endpoints of toxicity for chemical substances are sufficiently flexible to incorporate knowledge concerning the mode(s)/mechanism(s) of action by which these endpoints occur, including those involving disturbance of normal endocrine status ("endocrine disruption"). Background on nomenclature, traditional practices using non-human animal models, and the nature and adequacy of data sets for deriving acceptable human exposures for chemicals, including endocrine active substances, are described, as is how/why the existing frameworks are adequate for the purpose of deriving numerical values for endocrine active substances. Projections on how assessment practices for these substances may evolve further in the future also are offered.

INTRODUCTION

Endocrine active substances that cause adverse effects often are characterized as "endocrine disruptors" when these effects can be linked by empirical evidence or inference to a disturbance of normal endocrine status. Endocrine disruption is not an adverse effect per se, but rather a mode or mechanism of action by which these substances may produce adverse effects [1]. Members of many different chemical classes and use categories (e.g., pesticides, human and veterinary drugs, and consumer products) may possess this capability. Up until the last few years, characterization or labeling of a chemical substance generally was only by the nature of the effect(s) it produced (e.g., "Chemical X is a developmental toxicant, a neurotoxicant, and a carcinogen") and not by the mode or mechanism of action by which the toxic effect is provoked. As the scientific community's understanding of the etiologies of toxic effects evolves, so does the practice of characterization of a chemical by effect *and/or* mode/mechanism of action. So, then, one might ask whether or not the long-standing hazard and risk assessment

*Report from a SCOPE/IUPAC project: Implication of Endocrine Active Substances for Human and Wildlife (J. Miyamoto and J. Burger, editors). Other reports are published in this issue, *Pure Appl. Chem.* **75**, 1617–2615 (2003).

†This paper represents the opinions only of its author and does not reflect opinions or policies of the International Life Sciences Institute or the ILSI Risk Science Institute.

‡Tel.: 202 659 3306; Fax: 202 659 3617; E-mail: pfennercrisp@ilsi.org

practices based primarily upon identification of an endpoint of toxicity require a major overhaul to accommodate the understanding of mode/mechanism of action. And, further, is "endocrine disruption" a special case? There is likely to be a range of opinions on both questions; however, this author would answer each question with a "no" in the belief that the existing assessment framework is flexible enough to accommodate knowledge of mode/mechanism and that "endocrine disruption" is just one of many modes/mechanisms that can be addressed within it.

NOMENCLATURE FOR "ACCEPTABLE" OR "SAFE" LEVELS OF HUMAN EXPOSURE

Over the past half-century or so, regulatory authorities and other scientific organizations around the world have developed procedures and practices that involve the application of mathematical techniques to representative studies from the body of available toxicity data on a chemical substance, with the intention of deriving numerical values that, as a matter of policy, are said to represent "acceptable" or "safe" levels of human exposure to that substance. This step of the risk assessment process occurs in the hazard characterization phase of the Risk Assessment Paradigm, with hazard characterization embracing both hazard identification and dose–response assessment. Exposure assessment and risk characterization round out the paradigm. Numerical values are known by a variety of names, depending, in part, upon the regulatory setting in which they are derived and, further, upon the organization developing the numbers. The key message with regard to nomenclature is that, even though different names may be given to these numerical values, they all represent essentially the same concept. Each represents, in the opinion of the originator of the number, an exposure dose that, when occurring over a specified time span, will be "acceptable" or "safe". The goal is to protect the human against adverse health consequences, should exposure to a chemical substance occur.

The longest-standing and most well-known numerical value developed to characterize "safety" or "acceptability" over a lifetime of oral exposure is the "acceptable daily intake" or ADI. This is a term historically associated with the hazard assessment that informs the setting of maximum residue levels for pesticides, veterinary drugs, and other additives in food (e.g., [2–4]). Some organizations use the term "tolerable daily intake" or "tolerable intake" when deriving numerical values for substances that are found inadvertently or unintentionally in food or elsewhere in the environment; that is, they have not been preapproved for a specific use (e.g., [5–7]). Examples of such substances are dichlorodiphenyltrichloroethane (DDT), dioxins, heavy metals, polychlorinated biphenyls (PCBs), or naturally occurring mycotoxins (e.g., aflatoxin on peanuts and grains). Alternatively, the U.S. Environmental Protection Agency (USEPA) prefers to use the term "reference dose" (RfD) for oral exposure scenarios and "reference concentration" (RfC) for inhalation exposure scenarios [8,9], and the U.S. Agency for Toxic Substances and Disease Registry (ATSDR) prefers to use the term "minimal risk level" to denote oral or inhalation exposure levels that agency deems to be without appreciable risk [10].

More recently, it has been realized that hazard and risk assessments were required for chemical exposure scenarios that were less-than-lifetime in duration. Perhaps the earliest implementation of such an approach is that of the drinking water program at the USEPA. While a number of formal drinking water standards were in place in the early 1980s, many additional substances were being found in drinking water supplies, and questions were being raised about their potential human health risks. Since standard-setting is a long and arduous process, and guidance was needed in the short term, the program developed the Health Advisory Program, which provided nonregulatory guidance to affected drinking water system and public health officials within hours or days of their initial inquiries. This guidance provided critical decision tools to determine if a "Do not drink" advisory should be issued, water from an outside source should be brought in, or no interventions were necessary. If the appropriate toxicological data were available, the program would develop Acute, Short-term, Longer-term, and Lifetime Health Advisories, each of which had the underpinnings of an RfD or its less-than-lifetime, exposure-duration-specific equivalent. In the late 1980s, the ATSDR began developing its series of toxicological profiles for substances found at hazardous waste sites. These profiles contain not only maximum

residue levels (MRLs) for chronic (≥365 days) oral and/or inhalation exposures, but, where the data allow, MRLs for acute (from 1 to 14 days) and intermediate (from 14 to 364 days) exposure durations, as well [10]. In the early 1990s, the Pesticide Program at the USEPA began deriving acute RfDs for acute dietary exposure scenarios (single serving/meal or single day exposures) for pesticides with acute effects (e.g., cholinesterase-inhibiting insecticides and developmental toxicants) and intermediate RfDs for residential and occupational risk assessments. Within the past five years, the World Health Organization (WHO) Expert Panel of the (FAO)/WHO Joint Meeting on Pesticide Residues, an expert committee that provides technical advice to the Codex Committee on Pesticide Residues, has begun deriving acute RfDs for acutely toxic pesticides, as well [11].

Recently, a technical workgroup of the USEPA's Risk Assessment Forum, in its reevaluation of the agency's RfD/RfC development process, recommended that acute, intermediate, and long-term RfDs and RfCs be derived for all environmental substances within their regulatory purview [9]. This recommendation is tempered by the caveat that RfDs/RfCs for these multiple durations of exposure should be derived *only* if adequate and appropriate data are available to do so.

"TRADITIONAL" HAZARD/RISK ASSESSMENT FRAMEWORK

Traditionally, numerical values have been developed using toxicity data that provide two kinds of information. In the first instance, these data allow *hazard identification*; that is, the determination of the inherent toxic characteristics of the substance (e.g., whether it causes liver toxicity, infertility, or neurotoxicity). Secondly, data from studies that include multiple groups, including both exposed and unexposed subjects, allow *dose–response assessment*. Historically, data developed following chemical exposures in humans and in laboratory animal *mammalian* species have served as the basis for the development of numerical values. The desirability of using human data is obvious; no inter-species extrapolation need be done. However, adequate and relevant chemical-specific human data rarely are available for this purpose. Thus, we have come to depend upon a small number of non-human mammalian test species to provide the empirical base for hazard and risk assessment. These species have been chosen for two principal reasons: (1) assumed biological similarity to humans, and (2) economics and feasibility (e.g., relative ease of handling in a laboratory setting). The most commonly used species in regulatory testing paradigms are rats and mice, along with rabbits and dogs. Occasionally, guinea pigs, hamsters, and non-human primates also may be employed for testing. These seven species and/or others may be used in the research setting.

National and multinational regulatory agencies engaged in chemicals management generally have the authority to require the generation of a body of toxicological data for certain use categories prior to the introduction of those chemicals into commerce. This is particularly true for pesticides, human and veterinary drugs, and food additives. Some agencies also have the authority to require pre-market data on commodity chemicals destined for industrial, consumer, or other uses. Not all of these required data may be suitable for use in deriving numerical values. One criterion of suitability would be how well the study design can predict the potential for effects, or lack thereof (i.e., hazard characterization), in the exposure scenario for which the assessment is being conducted. A second criterion would be whether the study design includes multiple treatment groups and at least one control group so that *dose–response assessment* can be conducted. Historically, key assumptions involved when deriving tolerable daily intakes (TDIs)/ ADIs/RfDs/RfCs/MRLs are that the dose response for the parameter used as the basis for quantitation is monotonic and that a point of departure (POD) for quantitative assessment can be identified. This POD may be an observed no/lowest-observed-adverse-effect level (NOAEL or LOAEL) or a modeled benchmark dose representing some percentage of occurrence of effect (e.g., 1, 5, 10 %). Traditionally, safety or uncertainty factors then are applied to the selected POD to yield a "number". Selection of the factor(s) is informed by a modestly complex set of decision criteria. Detailed discussion of this component of hazard assessment can be found in Topics 3.13 and 3.14.

Pesticides are a use category of chemicals for which a robust set of toxicity data is generated in order for regulatory authorities to grant pre-market approval before allowing their introduction into commerce and the environment. As such, the resulting data are likely to be amenable to the development of "acceptable" or "safe" levels of human exposure for lifetime and less-than-lifetime exposure durations.

At the present time, the standard set of first-tier toxicology regulatory requirements focuses on identification of effects and studies are carried out in conformance with standardized consensus testing protocols. Roughly identical first-tier requirements exist for all Organization for Economic Cooperation and Development (OECD) member countries, including the members of the European Union. Many other non-OECD countries apply the OECD testing paradigm in their own regulatory programs.

An example of the standard first-tier set for food-use pesticides of conventional chemistry can be found in the U.S. Code of Federal Regulations [12,13]. Routinely required (i.e., first-tier) studies for food use pesticides of conventional chemistry include a battery of six tests to characterize acute toxicity (one each by the oral, dermal, and inhalation routes plus eye and skin irritation and dermal sensitization), subchronic feeding studies in rodent and non-rodent species, chronic feeding studies in rodent and non-rodent species, carcinogenicity studies in two rodent species, prenatal developmental toxicity studies in rodents and non-rodents, and a two-generation reproduction study in rodents. Second-tier studies may include dermal penetration, 21-day dermal, subchronic dermal, subchronic inhalation, and acute and subchronic neurotoxicity studies in rodents; acute and subchronic delayed neurotoxicity in hens; or a developmental neurotoxicity (DNT) study in rodents. These are triggered by some special characteristic of the pesticide (e.g., its chemical class) or by potential use and exposure patterns (e.g., residential uses) or by the results of routinely required studies. These study types have been found to have value in the derivation of numerical values, except for the six acute studies, which, by virtue of their study design and measurement parameters, are best suited only for hazard identification, but not for dose–response assessment. Embedded in all of the study protocols is the selection of specific parameters to be measured, which are thought to represent key manifestations of particular endpoint of toxicity. Over the years, as the state-of-the-science has evolved, so have the kind and number of endpoints measured. Many of these endpoints have come to be understood as events in endocrine disruptor mode/mechanisms of action.

Several key observations can be made about the traditional hazard/risk assessment framework for derivation of numerical values in human health assessment:

- Toxicological data development is focused primarily on toxic endpoint identification.
- The data used for derivation of numerical values come only from human exposure situations or studies conducted in other mammalian species.
- As a default, the dose responses for these endpoints are assumed to be monotonic and reflective of a threshold for the effect.

AN EVOLVING HAZARD/RISK ASSESSMENT FRAMEWORK FOR ENDOCRINE DISRUPTORS

Generally, in the evaluation of most chemical substances, pharmaceuticals being the principal exception, only after effects of concern have been identified are follow-up studies done to characterize the mode/mechanism of action by which these effects arise. Currently, there are no standardized consensus guidelines available for this purpose for any specific mode/mechanism, other than the mutagenicity/genotoxicity studies that help to inform the characteristics of a carcinogenic or heritable mutation effect. As discussed elsewhere in this conference, a significant effort is underway on a national (e.g., in the United States, in response to legislative mandates in the 1996 Food Quality Protection and Safe Drinking Water Acts [14]) and international scale (i.e., through the OECD [15]) to develop testing paradigms specifically to identify endocrine disruptors. These will be supported by the development

and implementation of new and updated standardized and validated consensus protocols. Some of these study designs will define a toxicological profile, as do the traditionally conducted studies; others will assist in characterizing mode/mechanism of action. Worthy of note here is that, as the mode(s)/mechanism(s) of action (MOAs) of potential/identified endocrine active substances have been characterized, it has become apparent that many of the parameters that were measured originally only to identify effects of concern in the traditional studies also represent a key event or stage in the MOAs of the effect(s) of concern. Thus, as one component of a strategy to identify and characterize endocrine disruptors, it is reasonable and appropriate to build upon, and modify, traditional study designs for endpoint identification as the MOAs become better understood, rather than to start over.

How, then, might the traditional hazard/risk human health assessment framework used in deriving numerical values evolve to accommodate mode/mechanism of action information in the identification and assessment of endocrine active substances? In the near term, it should capitalize on the efforts underway to develop the screening and testing programs. In the first instance, continued improvements to traditional apical study designs in whole animal mammalian systems will contribute to more accurate understanding of endpoints of toxicity and also their MOAs. Other kinds of studies will be conducted in in vitro mammalian systems or in in vivo or in vitro nonmammalian systems. Some of the in vivo studies in nonmammalian systems include a relatively broad spectrum of apical endpoints, and, therefore, are structurally similar to the traditional apical mammalian studies currently used for number development. Others include few or no apical endpoints and better serve for elucidation of the MOA(s). Data from the non-apical mammalian and/or the nonmammalian studies still could have a role in both *hazard identification* and *dose–response assessment*, as follows:

Hazard identification

Currently, many endpoints are evaluated in the apical mammalian studies that are used for development of numerical values. Absent an understanding of the underlying MOA for the production of these endpoints, it is more difficult to sort out the effects of concern from the ones of lesser importance (i.e., the effects that are not adverse). Data from these other types of studies can help refine and focus the endpoint selection process. Instead of selecting the endpoint with the lowest NOAEL or lowest 1, 5, or 10 % benchmark dose (BMD) simply because it is the lowest, the important endpoints can be selected based upon their involvement in potentially adverse consequences. These important endpoints may or may not have the lowest NOAEL or BMD, but they would represent a more accurate assessment of potential hazard.

Dose–response assessment

As noted earlier, the default assumption for the shape of the dose–response curve for most toxic effects is monotonic, nonlinear. (There continue to be differences of opinion globally concerning the appropriate default assumption for carcinogenic responses.) The extrapolation techniques, then, were designed with the monotonic, nonlinear default in mind, but also could easily accommodate a linear response. However, as has been noted in many scientific papers in the literature, some of the dose responses for some effects observed following exposure to endocrine active substances exhibit neither a linear nor a monotonic, nonlinear dose response. Bimodal and "U"-shaped dose responses have been observed. Others may exist. This is not a problem if these effects are not selected as the critical ones for use in development of numerical values. It does pose a new challenge if they are. Obviously, then, extrapolation techniques that account for these other shapes must developed and applied. Defaults should be abandoned when credible data are available to accurately describe the biology. The data may validate the default or prove it incorrect. Empirical evidence should drive the assessment process.

Hazard characterization

As noted earlier, hazard characterization encompasses both hazard identification and dose–response assessment, but also is the phase of risk assessment in which the quantitative assessment, if done, is completed and described. Therefore, it is at this point in the process that the application of the procedure for deriving numerical values would be synthesized and summarized. In the future, it may be that MOA data play a role in further refinement of the numerical value, beyond those already described above in the hazard identification and dose–response assessment sections. No specific existing examples are known to the author and no hypothetical examples come to mind at this time. We, however, should remain alert to this possibility.

SUMMARY

The derivation of numerical values is often a key element in the assessment of human health hazard/risk from exposures to chemical substances. Traditional risk assessment practices typically employ toxicity data only from mammalian species (although this could be human data). These toxicity data generally identify only endpoints of effect, and inform little with regard to mode/mechanism of action. Absent empirical data, dose responses are generally assumed to be monotonic, nonlinear, although linear responses can be easily accommodated in the extrapolation process. An evolving framework for risk assessment of endocrine active substances/endocrine disruptors should incorporate relevant information on mode/mechanism of action, even if developed in nonmammalian systems. Empirical evidence indicates that sometimes the dose response for effects resulting as a consequence of disturbance of normal endocrine status is neither monotonic, nonlinear nor linear. If these effects are selected as critical to the development of a numerical value, extrapolation techniques that accommodate these different dose responses must be employed.

REFERENCES

1. U.S. EPA. *Special Report on Environmental Endocrine Disruption: An Effects Assessment and Analysis*, February, EPA/630/R-96/012, U.S. Environmental Protection Agency, Washington, DC (1997). <http://www.epa.gov/ORD/WebPubs/endocrine>
2. WHO JECFA. *Principles for the Safety Assessment of Food Additives and Contaminants in Food*, WHO Environmental Health Criteria No. 70, World Health Organization Expert Committee of the Joint FAO/WHO Expert Committee on Food Additives, World Health Organization, Geneva (1987). <http://www.inchem.org/documents/ehc/ehc/ehc70.htm>
3. WHO JMPR. *Principles for the Toxicological Assessment of Pesticide Residues in Food*, WHO Environmental Health Criteria, No. 104. Geneva, World Health Organization Expert Group on Pesticide Residues of the Joint FAO/WHO Meeting on Pesticide Residues, World Health Organization (1990). <http://www.inchem.org/documents/ehc/ehc/ehc104.htm>
4. U.S. FDA. *Food Contact Substance Notification Program's Cumulative Estimated Daily Intake/Acceptable Daily Intake Database*, U.S. Food and Drug Administration, Center for Food Safety & Applied Nutrition, Office of Food Additive Safety, August, Washington, DC (2002). <http://www.cfsan.fda.gov/~dms/opa-edi.html>
5. U.S. FDA. *Guidance Document for Cadmium in Shellfish*, U.S. Food and Drug Administration, Center for Food Safety & Applied Nutrition, Office of Food Additive Safety, August, Washington, DC (1993). <http://www.cfsan.fda.gov/~frf/guid-cd.html>
6. WHO IPCS. *Assessing Human Health Risks of Chemicals: Derivation of Guidance Values for Health-based Exposure Limits*, Environmental Health Criteria No. 170. World Health Organization International Program on Chemical Safety, World Health Organization, Geneva (1994). <http://www.inchem.org/documents/ehc/ehc/ehc170.htm>

7. European Commission Scientific Committee on Food. Opinion of the Scientific Committee on Food on Bisphenol A. European Commission Health and Consumer Protection Directorate-General, April, Brussels (2002). (http://europa.eu.int/foods/fs/sc/scf/index_en.html>
8. U.S. EPA. *Reference Dose (RfD): Description and Use in Health Assessments. Integrated Risk Information System (IRIS) Background Document 1A*, U.S. Environmental Protection Agency, Washington, DC (1994). <http://www.epa.gov.iris/rfd.htm>
9. U.S. EPA. *A Review of the Reference Dose and Reference Concentration Processes. External Review Draft*, May, EPA/630/P-02/002A, U.S. Environmental Protection Agency, Washington, DC (2002a). <http://www.epa.gov/ncea/raf/pdfs/RfDRfC/rfdrfcextrevdrft.pdf>
10. ATSDR. *Minimal Risk Levels for Hazardous Substances*, Agency for Toxic Substances and Disease Registry (2002). <http://www.atsdr.cdc.gov/mrls.html>
11. WHO JMPR. "Procedures for estimating acute reference dose", In *Pesticide Residues in Food–1997*, World Health Organization Expert Group on Pesticide Residues of the Joint FAO/WHO Meeting on Pesticide Residues, Rome (1998). <http://www.fao.org/waicent/faoinfo/agricult/agp/agpp/pesticid/jmpr/download/98/cont-intro.pdf>
12. 40 CFR Part 158. 2002. Code of Federal Regulations. Part 158.340 Toxicology data requirements. <http://www.access.gpo.gov/nara/cfr/index.html>
13. U.S. EPA. *Determination of the Sppropriate FQPA Safety Factor(s) in Tolerance Assessment*, U.S. Environmental Protection Agency, Office of Pesticide Programs, Washington, DC (2002b). <http://www.epa.gov/oppfead1/trac/science/determ.pdf>
14. U.S. EPA. Endocrine Disruptor Screening Program Web site, U.S. Environmental Protection Agency (2002c). <http://www.epa.gov/scipoly/oscpendo/index.htm>
15. OECD. Special Activity of the Test Guidelines Programme on Endocrine Disruptor Testing and Assessment, Organization of Economic Cooperation and Development (2002). <http://www1.oecd.org/ehs/endocrin/Subpage/Reports.htm>