# IPCS

INTERNATIONAL PROGRAMME ON CHEMICAL SAFETY

WHO ILO UNEP

Environmental Health Criteria 240

# Principles and Methods for the Risk Assessment of Chemicals in Food

## Chapter 5
## DOSE-RESPONSE ASSESSMENT AND DERIVATION OF HEALTH-BASED GUIDANCE VALUES

A joint publication of the Food and Agriculture Organization of the United Nations and the World Health Organization

Food and Agriculture Organization of the United Nations

World Health Organization

## Environmental Health Criteria 240

# PRINCIPLES AND METHODS FOR THE RISK ASSESSMENT OF CHEMICALS IN FOOD

A joint publication of the Food and Agriculture Organization of the United Nations and the World Health Organization

Published under the joint sponsorship of the United Nations Environment Programme, the International Labour Organization and the World Health Organization, and produced within the framework of the Inter-Organization Programme for the Sound Management of Chemicals.


Food and Agriculture Organization of the United Nations

World Health Organization

The **International Programme on Chemical Safety (IPCS)**, established in 1980, is a joint venture of the United Nations Environment Programme (UNEP), the International Labour Organization (ILO) and the World Health Organization (WHO). The overall objectives of the IPCS are to establish the scientific basis for assessment of the risk to human health and the environment from exposure to chemicals, through international peer review processes, as a prerequisite for the promotion of chemical safety, and to provide technical assistance in strengthening national capacities for the sound management of chemicals.

The **Inter-Organization Programme for the Sound Management of Chemicals (IOMC)** was established in 1995 by UNEP, ILO, the Food and Agriculture Organization of the United Nations, WHO, the United Nations Industrial Development Organization, the United Nations Institute for Training and Research and the Organisation for Economic Co-operation and Development (Participating Organizations), following recommendations made by the 1992 UN Conference on Environment and Development to strengthen cooperation and increase coordination in the field of chemical safety. The purpose of the IOMC is to promote coordination of the policies and activities pursued by the Participating Organizations, jointly or separately, to achieve the sound management of chemicals in relation to human health and the environment.

WHO Library Cataloguing-in-Publication Data

Principles and methods for the risk assessment of chemicals in food.

(Environmental health criteria ; 240)

1. Risk assessment. 2. Hazard assessment. 3. Exposure assessment. 4. Dose-response assessment. 5. Chemicals. 6. Food safety. 7. Food additives. 8. Contaminants. 9. Pesticide residues. 10. Veterinary drug residues. I.World Health Organization. II.Food and Agriculture Organization of the United Nations.

ISBN 978 92 4 157240 8 (NLM classification: WA 712)
ISSN 0250-863X



This document was technically and linguistically edited by Marla Sheffer, Ottawa, Canada.

Printed by Wissenchaftliche Verlagsgesellschaft mbH, Stuttgart, Germany.

# 5. DOSE–RESPONSE ASSESSMENT AND DERIVATION OF HEALTH-BASED GUIDANCE VALUES



---

For acronyms and abbreviations used in the text, the reader may refer to the list of acronyms and abbreviations at the front of this monograph. Definitions of select terms may be found in the glossary at the end of the monograph.

## 5.1 Dose–response assessment

### *5.1.1 Basic concepts of dose–response assessment*

Dose–response assessment approaches generally take one of two forms: 1) analyses that provide a quantitative (or sometimes just qualitative) estimation of risk and 2) analyses that establish health-based guidance values, such as an acceptable daily intake (ADI) or tolerable daily intake (TDI), which are levels of human exposure considered to be without appreciable health risk. The latter approach, which is often described as "safety assessment", is used more often in cases where exposure can be controlled, such as for food additives and residues of pesticides and veterinary drugs in foods.

One of the primary criteria of a risk assessment is determination of the presence or absence of a cause–effect relationship. If there is sufficient plausibility for the presence of such a relationship, then dose–response data are essential, and dose–response analysis is a major part of the hazard characterization within the risk assessment paradigm.

Dose–response data may be derived from in vivo studies in laboratory animals or humans, which usually provide the basis for risk characterization, and in vitro studies, which are often related to investigations of mode of action. In each case, interpretation of the data on effects usually requires recognition of the levels of exposure that do not produce a measurable effect and the relationship between the increase in incidence, severity or nature of the effect with increase in exposure.

Toxicological or epidemiological data have been used in hazard characterization by the Joint Food and Agriculture Organization of the United Nations (FAO)/World Health Organization (WHO) Expert Committee on Food Additives (JECFA) and the Joint FAO/WHO Meeting on Pesticide Residues (JMPR) in three main ways (see chapter 7):

1) derivation of a health-based guidance value, such as an ADI, TDI or acute reference dose (ARfD);
2) estimation of the margin of exposure (MOE) between a defined point on the dose–response curve and the level of human exposure; and
3) quantification of the magnitude of the risk at specified levels of human exposure.

In addition, it is possible to use dose–response data to define the exposure that theoretically would be associated with some specified level of risk, such as a 1 in a million increase in lifetime risk of cancer.

Having established that there is a statistically significant treatment-related or exposure-related effect that is relevant to human health, the calculation of a health-based guidance value or MOE requires definition of a reference point or point of departure (POD) on the dose–response curve. There have been two basic approaches to dose–response assessment applied to data from studies in animals:

1) Pairwise comparisons of the findings in different groups in order to define experimental doses that cause statistically significant effects and the highest experimental dose that does not produce an observed adverse effect in that study, the no-observed-adverse-effect level (NOAEL). The NOAEL is then used as the POD to estimate a health-based guidance value, after allowing for uncertainties such as species differences and human variability.

2) Fitting a model or models to the dose–response data for all groups in order to define the relationship in the observed range; the model can then be used to define the exposure associated with a specified level of response. This value can then be used as the POD to estimate a health-based guidance value or calculate an MOE or extrapolated to estimate the risk at the levels of human exposure that are relevant to problem formulation and risk characterization.

These approaches and variants on them are discussed in this chapter, which is based on an Environmental Health Criteria (EHC) document on Principles for Modelling Dose–Response for the Risk Assessment of Chemicals, developed as part of the International Programme on Chemical Safety (IPCS) Harmonization

Project on Approaches to the Assessment of Risk from Exposure to Chemicals. EHC 239 (IPCS, 2009) covers toxicants with threshold effects and those for which there may be no practical threshold, such as substances that are genotoxic and carcinogenic. It focuses primarily on experimental animal studies, but dose–response relationships are also critical to the assessment of human experimental studies and epidemiological data. Dose–response assessment is also important for studies that attempt to define the relationships of different steps in a postulated mode of action. EHC 239 also includes areas that are not of direct relevance to this chapter, such as the basic risk analysis paradigm and the consequences of dose–response modelling (DRM) for the advice provided by risk assessors to risk managers.

#### 5.1.1.1 *Dose*

It is critical when performing dose–response analyses to have a clear concept of what type of "dose" has been used in the available dose–response data. There are three basic types of dose that arise from scientific investigations; they are inter-related, and each of them can be used to express dose–response relationships. They are 1) the administered or external dose, 2) the internal (absorbed) dose and 3) the target or tissue dose.

External dose denotes the amount of an agent or chemical administered to an experimental animal or human in a controlled experimental setting by some specific route at some specific frequency. In the terminology used by JECFA, the external dose is often referred to as exposure or intake (see chapter 6). External dose, or external exposure, is frequently the dose metric that is used in observational epidemiological studies.

Internal dose is the amount that is systemically available and can be regarded as the fraction of the external dose that is absorbed and enters the general circulation. It is affected by absorption, metabolism and excretion of the chemical and can be derived from suitable toxicokinetic mass balance studies. The analytical method used in the toxicokinetic studies will determine whether the dose refers to the parent compound alone or the parent compound plus first-pass metabolites (see chapter 4, section 4.2). Biomarkers of body burden, such as plasma concentrations or urinary excretion, are sometimes available in epidemiological studies.

The tissue dose is the amount that is distributed to and present in a specific tissue of interest. As for internal dose, the analytical method used in the toxicokinetic studies will determine whether the dose refers to the toxic entity, whether it be the parent compound alone or the parent compound plus first-pass metabolites (see chapter 4, section 4.2). An additional consideration for tissue dose is whether the dose metric is the peak concentration or a time-weighted average, such as the area under the concentration–time curve (AUC).

Two temporal parameters are important determinants of dose: the dose frequency and the duration of dosing. Dosing can be acute, subchronic or chronic; the term dose can apply to any of these, and the principles of dose–response assessment apply to all three forms. The description of dose should reflect the magnitude, frequency and duration over which it applies. Dose can be expressed in a variety of metrics, including a simple single external dose (e.g. mg/kg body weight), daily intake (e.g. mg/kg body weight per day),[1] peak body burden or body burden averaged over a given period of time (e.g. ng/kg body weight) or tissue concentration (e.g. ng/kg).

In epidemiological studies, exposure (the external dose) is rarely known precisely, and its estimation often requires various assumptions. Sometimes exposure is measured by the biomonitoring of blood or tissue concentrations; dose–response assessment for such data usually raises the issue of conversion of the biomarker of internal exposure into an external dose. An additional problem that has arisen (e.g. with the dioxin database) is that measurements of the biomarker were made many years after what was believed to be the period of highest exposure (FAO/WHO, 2002a).

Sometimes the doses used in an experimental animal study are transformed to the equivalent human exposures prior to DRM. In this situation, models of internal exposure linked to the response data may be used to develop a dose–response model. However, such models need knowledge, for both experimental animals and humans, of the events controlling absorption, tissue distribution, metabolism, excretion and the other molecular and biochemical processes that ultimately

[1] In animal studies, exposure is often measured as concentration in feed only. For conversion from feed concentration to external dose, refer to Annex 2.

lead to particular responses. Interspecies extrapolation of such a dose metric may be possible by the use of a physiologically based toxicokinetic (PBTK) model. Although this more sophisticated approach can refine DRM, incomplete data will add uncertainty to the output of the modelling. The issue of interspecies extrapolation is usually addressed separately and subsequent to DRM using the unadjusted animal data and application of an uncertainty factor (section 5.2.3).

#### 5.1.1.2 *Response*

Response, in this context, generally relates to an observation or effect seen following exposure in vivo or in vitro. Possible end-points cover a broad range of observations, from early responses such as biochemical alterations to more complicated responses such as cancer and developmental defects.

Responses can be either adaptive or adverse. Adverse effects are defined as a change in the morphology, physiology, growth, development, reproduction or lifespan of an organism or subsystem (e.g. subpopulation of cells) that results in an impairment of functional capacity, an impairment of the capacity to compensate for additional stress or an increase in susceptibility to other influences (IPCS, 2004). The responses are sometimes species or tissue specific and have different degrees of variation across individuals. DRM can address each response, provide insight into their quantitative similarities across species and tissues and link responses in a mechanistically reasonable manner.

Response is generally considered to vary across experimental units (experimental animals, humans, cell cultures) in the same dose group in a random fashion. This random variation is usually assumed to follow some statistical distribution describing the frequency of any given response for a population. In general, statistical distributions are characterized by their central tendency (usually the mean or median) and their effective range (usually based on the standard deviation or geometric standard deviation).

Most responses of interest in the context of dose–response assessment fall into one of four basic categories:

1) *Quantal responses*: Also referred to as binary or categorical responses, these generally relate to an effect that is either observed

or not observed in each individual subject (laboratory animal or human); for each dose, the number of subjects responding out of the number of subjects available is reported (e.g. the proportion of animals with a tumour in a cancer bioassay).

2) *Counts*: These generally relate to a discrete number of items measured in a single experimental unit (e.g. number of papillomas on the skin).

3) *Continuous measures*: These generally relate to a quantitative measurement that is associated with each individual subject and can take on any value within a defined range (e.g. body weight).

4) *Ordinal categorical measures*: These generally take on one value from a small set of ordered values (e.g. tumour severity grades); ordinal data are an intermediate type of data and reflect (ordered) severity categories—i.e. they are qualitative data but with a rank order (e.g. histopathological severity data) in each individual. When the categories are non-ordered, they are called categorical data, but these are rare for response data.

Sometimes it is useful for DRM purposes to convert continuous data into proportions (e.g. number of animals outside a clinically relevant range for an immune system marker) or categories (e.g. measured degree of liver necrosis converted to minimal, moderate or extensive).

There are some differences in how each of these different types of data are handled for DRM, but as a general rule, the goal of DRM is to describe the mean and variance of the response as a function of exposure or time.

### 5.1.2 Dose–response modelling (DRM)

#### 5.1.2.1 Overview

DRM can be described by six basic steps, with a variety of options at each step (Table 5.1). The first four steps relate to the analysis of the dose–response data, which is referred to as dose–response analysis (IPCS, 2009). Dose–response analysis provides the linkage of a model to dose–response data for the purposes of predicting response to a

**Table 5.1. Basic steps in dose–response assessment/modelling (adapted from IPCS, 2009)**

| Step | Description | Options |
|---|---|---|
| 1. Data selection | Determine the response to be modelled, and select appropriate data | End-point, quality, sample size, utility, availability |
| 2. Model selection | Choose the type of model to be applied to the data | End-point, data availability, purpose |
| 3. Statistical linkage | Choose statistical distributions to describe the variability in response | End-point, data type, model choice, software availability |
| 4. Parameter estimation | Combine the first three steps in an appropriate computer program to obtain estimates of the model parameters | Linkage function, software availability, variance |
| 5. Implementation | Use the estimated model parameters and the model formula to predict response/dose as needed | Outputs, target selection, model predictions, BMD, direct extrapolation |
| 6. Evaluation | Examine the sensitivity of the resulting predictions to the assumptions used in the analysis ("model validation") | Model comparison, uncertainty |

BMD, benchmark dose.

given dose or predicting the dose causing a given level of response. The last two steps deal with implementation and evaluation of the results of the analysis.

Step 1 involves selection of appropriate data for dose–response assessment. The criteria applied to assess whether the data are suitable for risk characterization purposes are similar whether hazard characterization is based on pairwise analyses of groups or modelling using all dose groups.

Step 2 involves the choice of an appropriate model. The type of data available can have a marked impact on the complexity of the model that can be used. For example, whereas two points can be used to identify the slope of a line, it takes at least three points to identify the shape of a more complex dose–response relationship. The issue of whether there are enough data to support a given model is

complex (see IPCS, 2009). Models may be divided into two categories: empirical and biologically based models. Most DRM to date has used empirical models—i.e. mathematical descriptions of the data that are not based on a mechanism of action. Biologically based models are generally based on basic principles about the onset and progression of disease in a biological system, are functionally complex and have far greater data requirements than do empirical models.

Step 3 requires the choice of a statistical linkage between the data and the model. The most common linkage method is to assume a statistical distribution for the response and use that distribution to derive a mathematical function describing the quality of the fit of the model to the data. The advantage of choosing a formal statistical linkage is the ability to test hypotheses and derive confidence intervals for model predictions.

Step 4 is the fitting of the selected model to the data. As the primary components of a model are the parameters that define the model, curve fitting simply involves choosing values for the parameters in the model. If a formal statistical function has been developed for linking the data to the model, then the parameters are chosen such that they "optimize" the value of the linkage function. A common choice is to link the data to the model by minimizing the sum of the squares of the differences between the predicted value from the model and the observed value. Simpler methods can also be used to estimate model parameters. Formal optimization is a better choice for modelling than ad hoc procedures, which lack transparency.

Step 5 is to make the inferences necessary to address the risk assessment questions developed at the problem formulation stage. The different types of data (quantal, count, continuous, categorical) require different methods for predicting changes in response beyond the normal response. In general, treatment-related responses may be described by added response (treated minus control response), relative response (fold change relative to control response) and extra response (added response scaled to range from zero to the maximum possible response). Each of these choices can have an impact on the final decision, so care should be taken to understand why a specific choice is made. Development of risk assessment advice usually requires extrapolation of results from the specific responses seen for the experiment

being modelled to other exposure scenarios and other doses. This step can also involve an extrapolation from a laboratory species to humans.

Step 6, uncertainty analysis, can be used to show the impact of sampling error and model selection on the model estimates. Sensitivity analysis can be used to evaluate the impact of a particular model choice on the estimate.

Dose–response assessment may be used to develop risk assessment advice in a variety of ways:

1) Simple pairwise comparisons of the data for different dose levels can be used to define the NOAEL or sometimes a lowest-observed-adverse-effect level (LOAEL), which is used as a POD for the observed dose–response data.

2) The dose–response model may be used to identify a dose with a known level of response at or slightly below the observable range. A specified response or level of effect for quantal and continuous data, respectively, is known as the benchmark response (BMR), and the dose associated with that response, the benchmark dose (BMD). The lower one-sided confidence limit of the BMD (the BMDL) can be used as the POD for the derivation of a health-based guidance value or for calculation of an MOE. Alternatively, the BMDL may be the starting point for linear low-dose extrapolation (see below).

3) The model may be used to find the dose associated with a negligible (e.g. 1 in a million) response over control. In general, this requires extrapolation far beyond the range of the data, which creates considerable uncertainty.

In addition, the model may be used to estimate the magnitude of effect associated with current levels of exposure for chemicals where exposure is ongoing and the dose–response data are derived from human studies.

Approach 1 is currently used by JECFA and JMPR to derive health-based guidance values in order to protect against effects that are considered to show a threshold.

Approach 2 was used by JECFA at its sixty-fourth meeting (FAO/WHO, 2006) to define MOEs for a number of genotoxic carcinogens. The same meeting also considered the use of linear extrapolation from the BMDL to estimate the risk of cancer at relevant levels of human exposure and concluded that

> calculation of the intake associated with an incidence of 1 in 1 million from the BMDL for a 10% incidence using linear extrapolation is simply equivalent to dividing the BMDL by 100 000, and this approach is therefore no more informative than calculation of a MOE.

Approach 3 was considered by JECFA at its sixty-fourth meeting (FAO/WHO, 2006), and the Committee concluded that:

> In order to provide realistic estimates of the possible carcinogenic effect at the estimated exposure for humans, mathematical modelling would need to take into account the shape of the dose–response relationship for the high doses used in the bioassay for cancer and for the much lower intakes by humans. Such information cannot be derived from the available data on cancer incidence from studies in animals. In the future, it may be possible to incorporate data on dose–response or concentration–response relationships for the critical biological activities involved in the generation of cancer (e.g. metabolic bioactivation and detoxification processes, DNA [deoxyribonucleic acid] binding, DNA repair, rates of cell proliferation and apoptosis) into a biologically based dose–response model for cancer that would also incorporate data on species differences in these processes. However, such data are not currently available. At present, any estimate of the possible incidence of cancer in experimental animals at intakes equal to those for humans has to be based on empirical mathematical equations that may not reflect the complexity of the underlying biology. A number of mathematical equations have been proposed for extrapolation to low doses. The resulting risk estimates are dependent on the mathematical model used; the divergence increases as the dose decreases and the output by different equations can differ by orders of magnitude at very low incidences.

In step 6, the basic steps of DRM shown in Table 5.1 are repeated to consider other options in the process in order to understand the impact of choices on the health-based measures derived from DRM. This final step is aimed at understanding the sensitivity of the analysis to specific choices and to judge the overall quality of the final predictions. Depending on the degree of difference between choices, there could be value in performing a formal analysis of the quality of the fit of the model to the data. Other methods can also be used to assess the

impact of choices used in the modelling on the eventual outcome, such as uncertainty analysis and Bayesian mixing.

#### 5.1.2.2 *Mathematical models*

A number of mathematical models have been or can be used to describe dose–response data. Their application and interpretation require specialized expertise. The main models are outlined below, and further details are provided in the report of the sixty-fourth meeting of JECFA (FAO/WHO, 2006) and in EHC 239 (IPCS, 2009).

Dose–response models are mathematical expressions fitted to scientific data that characterize the relationship between dose and response. Mathematical models consist of three basic components: 1) assumptions used to derive the model, 2) a functional form for the model and 3) parameters that are components of the functional form.

Dose–response models range from very simple models, such as the linear model described above, to extremely complicated models for which the eventual functional form cannot easily be expressed as a single equation (e.g. biologically based dose–response models).

Models can also be linked, meaning that one model could describe part of the dose–response process while another describes the remainder of the process. For example, for chemical carcinogenesis, in most cases tissue concentration is more closely linked to cancer risk than is administered dose. Given data on dose, tissue concentration and tumour response, a toxicokinetic model may be able to relate external dose to tissue concentration, and a multistage cancer model may be able to relate tissue concentration to response. The two models need to be combined in order to describe the dose–response relationship.

Dose–response models may incorporate other information into the model form. Age and time on study are commonly used in DRM, but other factors, such as species/strain/human ethnicity, sex and body weight, have also been used to expand the utility of dose–response models.

#### 5.1.2.3 *Dose–response models for continuous data*

The models listed in Table 5.2 are some of the forms that may be used to describe the relationship between dose and the magnitude of

**Table 5.2. Dose–response models for continuous data**

| Name | Notes | Equation for response | Parameter explanations |
|---|---|---|---|
| Hill equation log-logistic | A modification of the Michaelis-Menten equation that supposes that the occupation of multiple sites or receptors is required for the production of an effect. | $= \text{RMax} \dfrac{D^n}{K_D^n + D^n}$ | RMax is the maximum response, $D$ is the dose, $K_D$ is the reaction constant for the drug–receptor interaction and $n$ is the number of (hypothetical) binding sites. |
| Exponential | If the interaction between a chemical and a target site is irreversible, then the rate of the reaction is determined by the rate of association (ka) only. | $= \text{RMax}\,(1 - e^{-rD})$<br>The above is an equation for a first-order exponential model. | RMax is the maximum response, $D$ is the dose and $r$ is the exponential rate constant. |
| Power | Simple exponential model. | $= \beta D^{\alpha}$ | $D$ is the dose, α is the shape parameter and β is the scale parameter. |
| Linear | Although there is usually no biological theory to suggest it, linear models are often justified by their simplicity; linear models have but a single parameter. | $= mD$ | $D$ is the dose, and $m$ is the slope. |

a response on a continuous scale in an individual. When combined with a statistical distribution (e.g. normal or lognormal), these equations can also be used to describe the relationship between dose and a continuous response in a population, where the continuous model corresponds to the central estimate.

Dose–response data are often adjusted by subtracting the (mean) control value from each individual observation. However, this procedure does not account for the fact that the background response level in the controls is, as in the experimental groups, subject to sampling error and individual variability. A better approach is to account for the background response in the model with a parameter that needs to be estimated from the data (see IPCS, 2009).

#### 5.1.2.4 *Dose–response models for quantal data*

Quantal dose–response functions describe the relationship between dose and the frequency of a particular outcome in a population (see Table 5.3). For a group of homogeneous or nearly identical individuals, the relationship between dose and frequency could be described with a step function, where all subjects either respond or fail to respond at any given dose. However, because variability is ubiquitous in living organisms, quantal dose–response data typically show gradually increasing incidence with dose. One interpretation of this is that individual subjects differ in tolerance to the agent, which can be described by a statistical tolerance distribution. Hence, any cumulative distribution function (CDF) may be used as a quantal dose–response function. Other models have been derived from statistical assumptions about how the agent might exert its effect in an organism, such as the gamma multi-hit model.

Background response rates should be accounted for by incorporating an additional parameter in the dose–response model (see IPCS, 2009).

#### 5.1.2.5 *Model fitting and estimation of parameters*

Two basic methodologies are available for model fitting: conventional, in which parameters are selected to minimize or maximize an objective function, and Bayesian, in which information in a data set is

**Table 5.3. Dose–response models for quantal data**

| Name | Theoretical basis | Equation for frequency (*F*) | Parameter explanations |
|---|---|---|---|
| Step function | No variability. | If $D < T$, $F = 0$<br>If $D \geq T$, $F = 1$ | *D* is the dose, and *T* is the threshold parameter. |
| One-hit (single-hit) | Hit theory models employ the use of a rate to describe the interaction between a group of causal agents (e.g. molecules) and a group of targets (e.g. a human population). | $= 1 - e^{-(\alpha + D^*\beta)}$ | *D* is the dose, α is a location parameter and β is the slope parameter. |
| Gamma multi-hit | An expansion of the one-hit model, which is based on the notion that multiple hits or events are required to produce a particular effect. | = Γ(gamma**D*, *k*) | Γ() is the incomplete gamma CDF, *D* is the dose, gamma is a rate parameter and *k* is the number of hits required to produce the effect. |
| Probit normal | A descriptive model based on a normal or Gaussian distribution. | $= \Phi(\alpha + D^*\beta)$ | Φ() is the normal CDF, *D* is the dose, α is a location parameter and β is the slope parameter. |
| Log-probit | An expansion of the probit model. | $= \Phi(\alpha + \ln D^*\beta)$ | Φ() is the normal CDF, *D* is the dose, α is a location parameter and β is the slope parameter. |
| Logistic | The statistical logistic model is also a descriptive tool with no theoretical basis. | $= 1/(1 + e^{-(\alpha + D^*\beta)})$ | *D* is the dose, α is a location parameter and β is the slope parameter. |
| Log-logistic | An expansion of the logistic model. | $= 1/(1+ e^{-(\alpha + \ln D^*\beta)})$ | *D* is the dose, α is a location parameter and β is the slope parameter. |
| Weibull | A flexible descriptive model originally developed to describe survival data in demography. | $= e^{-(\alpha + (D^*\beta)^{\gamma})}$ | *D* is the dose, α is the background parameter, β is the slope parameter and γ is an exponent. |

combined with prior information about model parameters, resulting in a posterior distribution for those parameters that reflects the degree of uncertainty about the parameters. For historical and computational reasons, "user-friendly" software designed for carrying out dose–response analysis and non-linear modelling in general has been restricted to using conventional methodologies, whereas Bayesian methods are implemented in packages that require more extensive programming and substantially greater understanding of the statistical details (for further details on Bayesian approaches, see Hasselblad & Jarabek, 1995; Gelman et al., 2004). Whereas current software requires substantial statistical understanding for successful use of Bayesian methods and is thus beyond the reach of this document, even conventional methods require an understanding of some basic principles before outcomes from applying the software can be properly interpreted. Some general remarks may be helpful here.

The general approach of fitting a model is to find parameter values for the model that optimize the fit of the model to the data. To that end, a criterion function is defined, reflecting what is considered to be a good fit of the model. The goal is to find the parameter values that optimize the value of the criterion. For many models typically used, this can be achieved only by an iterative "trial and error" approach (see below). In many applications, the logarithm of the likelihood function is used as the criterion. The likelihood derives directly from the distribution assumed for the scatter in the data. For quantal data, the binomial likelihood is typically used. For continuous data, the normal likelihood is often used, be it for the observed responses themselves or for the log-transformed responses. Note that maximizing the normal likelihood function is in fact equivalent to minimizing the sum of squares.

Computer software uses algorithms to find parameter values that optimize the fit of the model to the data, and the user does not need to worry about the exact nature of the calculations. However, some basic understanding of the search process is required in order to interpret the outcomes. An iterative search algorithm tries to find "better" parameter values in a process by evaluating whether the fit can be improved by changing the parameter values through a trial and error process. More advanced algorithms operate by evaluating the slope at which the fit is improved for one or more parameter value changes. The algorithm

can start searching only when the parameters have values to start with. Although the software often gives a reasonable first guess for the starting values, the user may have to change these. It is not unusual (in particular when the information in the data is hardly sufficient to estimate the intended parameters) that the end result depends on the starting values chosen, and the user should be aware of that.

### 5.1.3 Modelling with covariates

In some circumstances, it is desirable to include variables in addition to just an exposure variable in dose–response models. For example, in epidemiological studies, it is common to model disease risk in terms of not only exposure, but also age, sex, socioeconomic status, smoking status and other measurements that may be relevant to the disease state. These other factors may not themselves be directly affected by the exposure, but they may be correlated with exposure status because of the way in which the sample was taken. Unless the proper covariates are included in a model for the relationship between exposure and the health end-point, the effect of exposure will be incorrectly estimated.

In principle, this sort of confounding cannot occur in bioassay studies in which animals are randomized to treatment groups, but it may be useful to include a covariate such as sex or body weight to account for some of the variability in a related measure.

### 5.1.4 Biologically based dose–response models

Although biological considerations may motivate the choice of one or several empirical models, the level of biological detail in such models is minimal. Thus, their credibility for interpolating and extrapolating a data set derives mainly from their fit to the data, as evaluated statistically. The biologically based dose–response models, another class of model, are much more complicated and are explicitly designed to model the biological details that lead from initial exposure to a toxicant to the ultimate pathological outcome. Typically, such a model includes a PBTK model to describe the distribution and metabolism of the parent compound and toxic metabolites, as well as other mechanistic or toxicodynamic models that link target tissue concentration to the ultimate response. The toxicodynamic part of the model may be relatively simple or may be as complicated as a fully elaborated stochastic model for carcinogenesis.

Such a model is really a quantitative expression of a set of biological hypotheses and, when rigorously tested against critical experiments, becomes a credible tool for extrapolating from experimental results into exposure realms that are difficult or expensive to reproduce in controlled experiments. Such models are quite expensive to construct in terms of both resources and time and thus would be expected to be developed fully only for exposures and toxicities of the highest concern.

### 5.1.5 Uncertainty

Any parameters or predictions estimated from a given model are only point estimates and, to a larger or smaller extent, uncertain. This uncertainty arises from at least three sources: 1) the sampling error arising from inferences about a larger population from a single experiment; 2) the reality that dose–response estimates often differ among experiments with different experimental design, protocol or uncontrolled circumstances; and 3) the fact that the "true" model is not known, which results in additional uncertainty when interpolating between doses, but even more so when extrapolating outside the dose range containing observations. These uncertainties may all be represented in a dose–response assessment through the use of probability distributions or probability trees. The latter technique involves using multiple alternative plausible assumptions about what data sets or models are to be used to produce an estimate, which results in a range of plausible estimates.

### 5.1.6 Issues of extrapolation

Extrapolation is a necessary part of all risk assessments, except in those rare cases where DRM uses data from studies in sufficient numbers of humans who are representative of the potential exposed population and who have had a level of exposure similar to that which is of concern.

Most of the methods used to implement the results of a dose–response analysis (step 5) address these extrapolation issues. The strategies used for extrapolation basically fall into two categories: 1) those aimed at providing estimates of risk for exposures outside of the range of the data used in the dose–response analysis and 2) those

aimed at establishing health-based guidance values, such as the ADI, without quantification of risk. The methods that have been used for extrapolation are diverse and sometimes contentious, with different countries, and even different agencies within a given country, using different approaches.

Even when human data are available and suitable for dose–response analysis, they are generally from selected populations or study groups, such as workers in occupational settings, whose exposures differ from those of the general population. Thus, dose–response analyses normally need to be extrapolated from the observed conditions where scientific support is available to conditions where scientific support is weaker or non-existent. For dose–response analyses based on human studies, extrapolation is generally a downward extrapolation to different levels of exposure, but can also be to different life stages (e.g. fetus, child) or different populations with different environmental factors that might affect exposure (e.g. dietary differences).

In most cases considered by JECFA and JMPR, the data used for DRM come from experiments in laboratory animals administered doses significantly exceeding the potential human exposure. For such dose–response analyses, there are two issues of extrapolation: 1) extrapolating from the test species to humans and 2) allowing for possible human differences in response. The methods employed for these extrapolation issues are varied, ranging from the use of uncertainty factors (see section 5.2.3) to more complicated modelling schemes based upon differences in toxicokinetics and toxicodynamics between humans and experimental animals and variability between different human individuals.

## 5.2 Setting health-based guidance values

### *5.2.1 Introduction*

The setting of health-based guidance values provides quantitative information from risk assessment for risk managers, enabling them to make decisions concerning the protection of human health. Health-based guidance values developed by JECFA and JMPR for substances found in food and also drinking-water are the quantitative expression of the range of oral exposure (either acute or chronic) that would be expected to be without appreciable health risk.

For substances intentionally added to food, such as food additives, and for residues of pesticides and veterinary drugs in food, the health-based guidance value is termed the ADI. JECFA and JMPR determine ADIs based on all the known facts at the time of the evaluation.

Substances that have long half-lives and accumulate in the body are not suitable for use as food additives (FAO/WHO, 1962a). Data packages should include metabolism and excretion studies designed to provide information on the cumulative properties of food additives.

At the time of its first meeting, JECFA recognized that the amount of an additive used in food should be established with due attention to "an adequate margin of safety to reduce to a minimum any hazard to health in all groups of consumers" (FAO/WHO, 1957). The second JECFA meeting (FAO/WHO, 1958), in outlining procedures for the testing of intentional food additives to establish their safety for use, concluded that the results of laboratory animal studies can be extrapolated to humans, and that

> some margin of safety is desirable to allow for any species difference in susceptibility, the numerical differences between the test animals and the human population exposed to the hazard, the greater variety of complicating disease processes in the human population, the difficulty of estimating the human intake, and the possibility of synergistic action among food additives.

This conclusion formed the basis for establishing the ADI, which is defined as an estimate of the amount of a food additive, expressed on a body weight basis, that can be ingested daily over a lifetime without appreciable health risk.

JECFA generally sets the ADI on the basis of the lowest relevant NOAEL in the most sensitive species.

The ADI is expressed in amount (e.g. mg) per kilogram of body weight, usually as a range from 0 to an upper limit. ADIs are normally expressed numerically using only one significant figure. The use of more than one significant figure might be taken to imply a greater degree of accuracy than that which can be achieved when assessing the hazard from the wide range of factors that influence toxicity.

When appropriate, JMPR and JECFA develop ARfDs (see section 5.2.9). The ARfD is defined as (FAO/WHO, 2002a):

> an estimate of the amount of a substance in food and/or drinking-water, normally expressed on a body-weight basis, that can be ingested in a period of 24 h or less, without appreciable health risk to the consumer, on the basis of all the known facts at the time of the evaluation.

For food contaminants that are generally unavoidable, JECFA has used the term "tolerable" for health-based guidance values. This term was considered more appropriate than "acceptable", as it signifies permissibility for the intake of contaminants associated with the consumption of otherwise wholesome and nutritious food. These have included TDI, provisional maximum tolerable daily intake (PMTDI), provisional tolerable weekly intake (PTWI) and provisional tolerable monthly intake (PTMI). The use of the term "provisional" expresses the tentative nature of the evaluation, in view of the paucity of reliable data on the consequences of human exposure at levels approaching those with which JECFA is concerned.

Health-based guidance values may be derived from either NOAELs or BMDs (BMDLs), often called the POD or reference point. The NOAEL approach has been used for over 50 years, and testing guidelines (chapter 4) have been developed to ensure that toxicological data are suitable to identify the adverse effect of concern and also to define a NOAEL. In the BMD approach, a NOAEL does not have to be identified, but doses with graded responses are needed to provide optimum model output.

Calculation of the health-based guidance value (HBGV) can be described as follows:

$$\text{HBGV} = \frac{\text{POD}}{\text{UFs}}$$

where UF is the uncertainty factor, a term often used synonymously with safety factor.

When relevant, JMPR and JECFA use an overall NOAEL as a basis for the ADI, considering the most relevant studies together. JMPR

made the following comment with regard to an overall NOAEL (FAO/WHO, 2004b):

> During the toxicological evaluation of a compound, the Meeting often has available more than one study in which the same end-points have been addressed. In such situations, the dose spacing may be different, resulting in different NOAELs and lowest-observed-adverse-effect levels (LOAELs). The Meeting agreed that in such circumstances it might be appropriate to consider the studies together. When they are comparable, including consideration of study design, end-points addressed, and strain of animal, the "overall NOAEL" should be the highest value identified in the available studies that provides a reasonable margin ($\geq 2$) over the lowest LOAEL, provided that due consideration is given to the shape of the dose–response curve.

JECFA subsequently applied this approach in the evaluation of phytosterols, phytostanols and their esters (FAO/WHO, 2009b).

Calculations of a health-based guidance value based on the NOAEL or BMD approach for the example of quantal response data are summarized in Table 5.4.

The table shows calculation of an ADI, but the methods are applicable to any health-based guidance value.

### 5.2.2 Data

In selecting an experimental animal study for use in risk assessment, due consideration needs to be given to matching, as far as is possible, the pattern of potential human exposure—i.e. the route and duration of exposure (as a fraction of lifetime) and the pattern of exposure (e.g. intermittent bolus dosing or dietary administration).

When considering which data to use from a set of available toxicity studies on a particular compound, it is not necessary to undertake DRM for each observed end-point in each study. Whether the NOAEL or BMD approach is used for risk assessment, the aim is to define the adverse effect that is produced at the lowest levels of exposure. Therefore, a first step would be to exclude studies that have NOAELs that are obviously larger than those from the other studies. In addition, end-points clearly not showing a dose–response on visual and

**Table 5.4. Comparison of methods used to derive health-based guidance values based on NOAEL and BMD approaches (using the Weibull model for illustrative purposes) for the case of quantal data (adapted from IPCS, 2009)**

| Step | NOAEL-derived ADI | BMD-derived ADI |
|---|---|---|
| 1. Data selection | Sufficient sample sizes, at least one dose with no statistically significant effect. Relevant end-points in a relevant species are important for any approach. | Sufficient number of doses with different response levels and a sufficient total number of subjects. |
| 2. Model selection | None. | Fit dose–response model (e.g. Weibull model). |
| 3. Statistical linkage | Pairwise statistical tests between dose groups and control group. | Predicted fractions are linked to observed fractions, and their "distance" is minimized by optimizing some fit criteria function (e.g. likelihood function based on assumed distribution). |
| 4. Parameter estimation | No parameter; the NOAEL corresponds to one of the dose levels in the study. | Choose an appropriate response, $p$, in the range of experimental response. Estimate $BMDL_p$, the 95% lower confidence bound on the $BMD_p$, where $\frac{R(BMD_p) - R(0)}{1 - R(0)} = p$ |
| 5. Implementation | ADI = NOAEL / UFs<br>where UF is uncertainty factor.[a] | ADI = $BMDL_p$ / UFs |
| 6. Evaluation | Statistical power analysis can be performed to check if the test was sensitive enough to detect relevant effects. | Sensitivity of BMD to model choice can be checked by fitting various models. |

[a] The term "uncertainty factor" (UF) is used synonymously with the term "safety factor".

statistical inspection of the data can be omitted. Then, based on the toxicological impact together with the apparent magnitude of the response, a selection of end-points can be made as candidates for DRM. After selecting the potentially relevant end-points, the suitability of each dose–response data set for dose–response analysis is considered. For the BMD approach, it is generally desirable to have at least three or four different doses (including controls) and different levels of effect associated with different doses.

A design optimal for the NOAEL approach could limit the use of DRM, and vice versa. Whereas the NOAEL approach requires sufficient sample sizes within dose groups (to provide statistical power), the BMD approach requires a sufficient number of dose groups (to provide a description of the whole dose–response).

The BMD approach can be used to analyse data from studies carried out in the past and based on the traditional designs (with three dose groups and a control). Although these may not be optimal for model fitting, the BMD approach retains the advantages outlined above. The BMD approach can also be used for combined analysis of multiple similar studies.

Both the BMD and NOAEL approaches may prove inadequate when the number of animals per dose group is too small. For example, when the critical effect is seen in an experimental animal such as the dog, with few animals per dose group, the NOAEL may be high because of the insensitivity of the test. Although the BMD approach is better for evaluating sparse dose–response data, it may also provide very uncertain estimates; unlike the NOAEL approach, however, the inherent uncertainty is more explicit.

### 5.2.3 Safety/uncertainty factors

The terms “safety factor” and “uncertainty factor” are often used interchangeably, “safety factor” having been used historically, but the preference now is to use “uncertainty factor”. Comparable terms used by other bodies are “adjustment factor” and “assessment factor”. Application of the factors is intended to provide an adequate margin of safety for the consumer, considering sensitive human population subgroups.

Uncertainty factors are default factors used to account for both uncertainty and variability. Historically, an uncertainty factor of 100 has been used to convert the NOAEL from a study in experimental animals into a health-based guidance value for human exposure (IPCS, 1987). Additional uncertainty factors may be used to allow for important database deficiencies, such as the absence of a chronic study or when effects are detected at all experimental dose levels and a NOAEL has not been defined. In such cases, a LOAEL might be used for establishing a health-based guidance value (IPCS, 1994).

The default 100-fold uncertainty factor may be seen to represent the product of two separate 10-fold factors that allow for 1) differences between the average responses in the experimental animals used in the study identified to derive the POD and those in average humans and 2) the variability in responses between average humans and those who are highly sensitive (IPCS, 1987). The recognition that the original 100-fold uncertainty factor could be considered to represent two 10-fold factors allowed some flexibility, because different factors could be applied to the NOAEL from a study in humans and the NOAEL from a study in experimental animals.

Although uncertainty factors were to some degree determined arbitrarily and validated subsequently by scientific data and practical experience, they are dependent on the nature of the compound, the amount, nature and quality of the toxicological data available, and the nature of the toxic effects of the compound. When considering appropriate uncertainty factors, a number of aspects have to be taken into account, such as species differences, individual variations and incompleteness of available data.

A number of basic principles have been developed for considering appropriate uncertainty factors (adapted from EHC 104; IPCS, 1990), as described below.

When determining health-based guidance values, the 100-fold default factor is used as the starting point for extrapolating animal data to humans and may be modified in the light of the data that are available and the various concerns that arise when considering these data. Some of these are given below:

1) When relevant human data are available, the 10-fold factor for interspecies variability may not be necessary or may be reduced, depending on whether the available data represent the most susceptible part of the population as well as representing a sufficiently large group of individuals. Recommendations on numbers required can be found in IPCS (2005).

2) The quality of the data supporting the NOAELs or BMDLs determined in the animal experiments (and also in human experiments) influences the choice of the uncertainty factor. An increased factor may be appropriate to account for deficiencies in the studies.

3) The quality of the total database may affect the choice of uncertainty factor. Significant data deficiencies may warrant an increased factor due to increased uncertainty. A clear explanation needs to be given as to the exact nature of the deficiency.

4) The type and significance of the initial toxic response may alter the uncertainty factor. Thus, a response that is marginal and reversible may result in a reduced safety factor, if the effect is still considered relevant for human health.

5) The shape of the dose–response curve (in those cases where data are adequate to permit derivation of such a curve) may also be considered in assessing uncertainty factors. An increased factor can be considered when the dose–response curve is very steep, particularly when the NOAEL is close to the LOAEL.

6) Metabolic considerations may influence the choice of uncertainty factor. Thus, saturation of metabolic pathways resulting in toxic manifestations, biphasic metabolic patterns and data on comparative metabolism may all affect the magnitude of the uncertainty factor. Suitable toxicokinetic data may be used to derive chemical-specific adjustment factors (CSAFs) (see below).

7) Knowledge of the comparative mechanism or mode of toxic action in experimental animals and humans may influence the choice of uncertainty factor. More broadly, information on the dose–response relationships for one or more key events comprising a mode of action, in experimental animals or humans, can be

invaluable in informing the choice of uncertainty factors. Suitable toxicodynamic data may be used to derive CSAFs (see below).

Some experimental support for the default uncertainty factors was published by Dourson & Stara (1983). This paper also proposed additional factors for extrapolating from subchronic data (10-fold) and for converting LOAELs to NOAELs (1- to 10-fold, depending upon the severity and concern raised by the observed effect). Reviews (Renwick & Lazarus, 1998; Dorne & Renwick, 2005; Dorne et al., 2005) of clinical data on human variation in the major pathways of foreign compound metabolism and pharmacological sensitivity have shown that the 10-fold factor is a reasonable default value. In addition, clinical and/or epidemiological research in humans may provide information on variation in response within the human population to a chemical and hence allow a more accurate determination of uncertainty factors.

The concept of CSAFs (IPCS, 1994, 2005) has been introduced to allow appropriate data on species differences or human variability in either toxicokinetics (fate of the chemical in the body) or toxicodynamics (actions of the chemical within the body) to modify the relevant default 10-fold uncertainty factor (Table 5.5). The strategy proposed by IPCS involves replacing the original 100-fold uncertainty factor with CSAFs.

The CSAFs enable the incorporation in risk assessment of specific quantitative data on species differences or human variability in either toxicokinetics or toxicodynamics to replace part of the default uncertainty factor. Although such information is often not available, information on pathways of elimination or mode of action may be available. For example, JECFA used comparative body burden data rather than external dose data in its calculation of a PTMI for dioxin-like substances, allowing the usual 100-fold uncertainty factor to be subdivided and replaced with chemical-specific lower values, as there was no need for interspecies differences in toxicokinetics or for toxicodynamic differences between species (FAO/WHO, 2002b). Detailed guidance on the application of CSAFs in risk assessment has been published (Meek et al., 2002; IPCS, 2005).

As information is available on the extent to which some of these pathways or processes vary between experimental animals and humans

**Table 5.5. Values for default uncertainty subfactors that can be replaced by CSAFs to derive composite uncertainty factors (from IPCS, 2005)**

| | Default subfactor | | |
|---|---|---|---|
| Source of uncertainty | Toxicokinetic | Toxicodynamic | Combined |
| Interspecies variation | 4.0 | 2.5 | 10 |
| Human interindividual variation | 3.16 | 3.16 | 10 |

or within humans, an approach has been proposed to enable this information to be used to inform the choice of uncertainty factors. This approach therefore lies somewhere between the normal default situation (100-fold uncertainty factor) and the derivation of CSAFs on the basis of quantitative chemical-specific data. Such factors have been termed "categorical factors" (Walton et al., 2001) or pathway-related factors (Dorne et al., 2005). This concept is applied by JMPR for pesticides where the effect (mostly acute) is dependent on the peak plasma concentrations ($C_{max}$) rather than the plasma concentration integrated over time (area under the curve, or AUC) in order to derive a combined uncertainty factor based on categorical and default factors. This would lead, for example, to a factor of 25 instead of the default of 100 for use with carbamates (for details, refer to section 2.5 of FAO/WHO, 2009a).

Several of the factors cited above may apply in the consideration of any one compound. Certain factors may serve to increase and others to decrease the choice of the final uncertainty factor. Therefore, it must be stressed that the total weight of evidence has to be considered in determining the appropriate uncertainty factor to be used and that the determination of uncertainty factors must be considered on a case-by-case basis.

### 5.2.4 The NOAEL approach to deriving health-based guidance values

The critical steps in this approach are selection of the appropriate data and determination of the NOAEL. Historically, JECFA has used the term no-observed-effect level (NOEL), which was defined in EHC 70 (IPCS, 1987) as "The greatest concentration or amount of an agent, found by study or observation, that causes no detectable, usually

adverse, alteration of morphology, functional capacity, growth, development, or lifespan of the target". In contrast, JMPR has used the term NOAEL, which was defined in EHC 104 (IPCS, 1990) as "The highest dose of a substance at which no toxic effects are observed". In reality, both terms have similar meaning, and the NOAEL has been used similarly to set health-based guidance values by both JECFA and JMPR.

For the purpose of this monograph, NOAEL will be used, defined as follows:

- *No-observed-adverse-effect level (NOAEL)*: Greatest concentration or amount of a substance, found by experiment or observation, which causes no detectable adverse alteration of morphology, functional capacity, growth, development, or lifespan of the target organism under defined conditions of exposure.

The main difficulty with this approach is that it is critically dependent on the sensitivity of the test method. The statistical linkage (step 3) determines whether or not there is a statistically significant effect (e.g. at the 5% level) compared with background (e.g. the control group) for each dose level separately. When the response is not statistically significant, it is generally considered that this level of intake is without biologically significant adverse health effects, but the power of the study to detect an adverse effect is not analysed. Given the typical animal studies used in toxicology, the effect size that can be detected by a statistical test may be larger than 10% (additional risk). Therefore, the NOAEL may be expected to be a dose at which the effect is in reality somewhere between 0% and 10% or more. The selection of the NOAEL (step 4) identifies the highest dose level that does not produce a statistically significant effect. The NOAEL approach tends to give lower health-based guidance values for studies with a higher power to detect adverse effects, which in effect "penalizes" better-designed studies. This emphasizes the importance of adherence to testing guidelines in order to ensure that the data are suitable for risk assessment purposes.

The value of the NOAEL depends strongly on the following characteristics of the study design:

- *Group size*. The power to detect a NOAEL at some dose level is directly dependent on the sample sizes chosen at those dose

levels. The larger the group size, the smaller the possible undetected effect size at the NOAEL.

- *Dose selection*. The NOAEL must be one of the doses actually applied in the study. If the true threshold is higher than the NOAEL, the distance between the two can be expected to be limited (related to the dose spacing used), whereas if the true threshold is lower than the NOAEL, the distance between the two is potentially unlimited.

- *Experimental variation*. Experimental variation comprises biological (e.g. genetic) variation between subjects, variation in experimental conditions (e.g. time of feeding, location in experimental room, time of section or interim measurements) and measurement errors. Larger experimental variation between subjects will result in lower statistical power, and hence higher NOAELs.

Calculation of the health-based guidance value (HBGV) from NOAEL-based DRM (step 5 above) is given by the equation:

$$\text{HBGV} = \frac{\text{NOAEL}}{\text{UFs}}$$

Step 6 could be undertaken to analyse the power of the dose group identified as representing the NOAEL to detect the adverse effect found at higher dose levels. For example, DRM could be used to determine, with 95% confidence intervals, the magnitude of effect that would be predicted to occur in the NOAEL group. In addition, step 6 could be used for both the NOAEL and BMD approaches to evaluate the sensitivity of the calculated health-based guidance value to the values of the uncertainty factors chosen.

### 5.2.5 Benchmark dose approach to deriving health-based guidance values

As an alternative to the NOAEL approach, the BMD concept has been introduced (Crump, 1984; Kimmel & Gaylor, 1988). In contrast to the NOAEL approach, this method defines a level of exposure producing a non-zero effect size or level of response as the POD for risk assessment. The BMD method has a number of advantages,

including the use of the full dose–response data in the statistical analysis, which allows the quantification of the uncertainty in the data. Higher uncertainty in the data—for example, due to small group sizes or high variation within a group—would be reflected in lower health-based guidance values.

In choosing the data (step 1) for BMD modelling, similar basic considerations apply as for the NOAEL method. Group sizes are less critical, because the POD is not based on identifying a level of exposure at which the adverse effect was not detected. Studies showing a graded monotonic response with a significant dose-related trend provide the best experimental data for modelling.

The main difficulty with this approach is that it requires the selection of a level of response, the BMR. In general, the level chosen is such that it is close to the limit of detection of the study, or a level that would generally be considered as representing a negligible health effect. A generic form of the BMD and BMDL is presented in Table 5.4 for the example of quantal data. A variety of response levels, such as 1%, 5% and 10%, may be selected as the BMR; differences in selection of the BMR could lead to discrepancies in health-based guidance values between different regulatory bodies.

Choosing a model (step 2) for the BMD method is dependent upon the types of data available and the characteristics of the response being modelled. Complicated models will require a larger number of dose groups than simpler models, and several models have been proposed for each type of data. In the United States Environmental Protection Agency BMD software program (http://www.epa.gov/NCEA/bmds/), a number of routinely used models are cited. If widely varying estimates are given when multiple models are applied to the same data, it may be necessary to select a particular model to calculate the BMDL. Strategies that have been suggested include using a criterion for goodness of fit (e.g. the Akaike Information Criterion), model averaging or the model that yields the lowest BMDL (IPCS, 2009).

At its sixty-fourth meeting, JECFA calculated the MOEs for a number of genotoxic and carcinogenic food contaminants using BMDL values derived by fitting a range of models to the available experimental dose–response data (FAO/WHO, 2006). Annex 3 of the

report of that meeting provides useful background information on the use of the BMD approach for risk assessment purposes.

The statistical linkage (step 3) between the data and the model can assume a number of different forms. For quantal data, it is appropriate to assume that the data are binomially distributed for each dose group.

Selection of the POD (step 4) for the BMD method is in reality selection of the BMR, because the model outputs simply report the BMD and BMDL values for the selected BMR. It is often not clear what level of response (BMR) can be considered as representing a negligible health effect. Selection of the BMR requires discussion among toxicologists and clinicians. Although an explicit statement on the BMR is an improvement compared with the generally unknown response level that may be associated with a NOAEL, choices of a BMR need consensus building and will remain a subjective "expert" judgement in what is essentially a mathematical approach. An alternative approach to selection of the BMR is to choose an excess response, often 10%, that is close to the limit of detection of the study, below which there was insufficient support from the experimental data; however, this simply leaves open the issue of the possible health consequences of the resulting level of response at that BMR. Further information on the selection of the BMR is given in IPCS (2009).

The health-based guidance value (HBGV) can be calculated as follows:

$$\text{HBGV} = \frac{\text{BMDL}}{\text{UFs}}$$

In this calculation, the values of the uncertainty factors could be the same as those used for the NOAEL or adjusted to account for a slightly different interpretation of the BMDL relative to the NOAEL. Unlike with the NOAEL approach, an extra uncertainty factor would not be necessary if all dose levels resulted in significant levels of adverse effect (indeed, such data would be more suitable for modelling). Empirical investigations showed for a large and representative set of compounds that the BMDL may be regarded as an analogue to

a NOAEL, and substituting one for the other would result in similar health-based guidance values (Crump, 1984; Barnes et al., 1995).

Unlike the NOAEL approach, the BMD method includes the determination of the response at a given dose, the magnitude of the dose at a given response and their confidence limits. By extrapolation of the dose–response model below the biologically observable dose range, the response at specified (lower) dose levels can be estimated, as well as the dose corresponding to a specific response level. It should be noted, however, that extrapolation from a single model that fits the data in the observed range cannot be justified, as other models fitting the data equally well may result in substantially different estimates of low-dose risk. Linear extrapolation from a BMD for a 10% response ($BMD_{10}$) has been applied as a simple method for low-dose extrapolation, but the sixty-fourth meeting of JECFA (FAO/WHO, 2006) concluded that "Linear extrapolation from a point of departure offers no advantages over an MOE and the results are open to misinterpretation because the numerical estimates may be regarded as quantification of the actual risk."

### *5.2.6 Acceptable daily intakes*

#### *5.2.6.1 Food additives*

In calculating the ADI, an uncertainty factor is applied to the NOAEL to provide a conservative margin of safety on account of the inherent uncertainties in extrapolating toxicity data from experimental animal studies to potential effects in humans as well as variation within the human species. When results from two or more animal studies are available, the ADI is based on the most sensitive animal species—i.e. the species that displayed the toxic effect at the lowest dose, unless metabolic or pharmacokinetic data are available establishing that the test in the other species is more appropriate for humans.

Generally, the ADI is established on the basis of toxicological information and provides a useful assessment of safety without the need for data on intended or actual use or dietary exposure. However, in setting ADIs, it may be necessary to know whether particular subpopulations are exposed, as the ADI applies to the whole population. Therefore, general information about exposure patterns should be known at the time of the safety assessment (see chapter 6). For example, if a food

additive is to be used in infant formulas, the safety assessment is not complete without looking carefully at safety studies involving exposure of very young animals.

There are occasions when JECFA considers the setting of an ADI in numerical terms not to be appropriate. This situation arises when the estimated consumption of the additive is expected to be well below any numerical value that would ordinarily be assigned to it. Under such circumstances, JECFA uses the term ADI "not specified". The Committee defines this term to mean that, on the basis of available data (chemical, biochemical, toxicological and other), the total daily intake of the substance arising from its use at the levels necessary to achieve the desired effect and from its acceptable background in food does not, in the opinion of the Committee, represent a hazard to health. For that reason, and for the reasons stated in the individual evaluations, the establishment of an ADI in numerical form is not deemed necessary (e.g. FAO/WHO, 1984, Annex II). An additive meeting this criterion must be used within the bounds of Good Manufacturing Practice (GMP)—i.e. it should be technologically efficacious and should be used at the lowest level necessary to achieve this effect, it should not conceal inferior food quality or adulteration, and it should not create a nutritional imbalance (FAO/WHO, 1974). That the background occurrence of the chemical must be taken into account in the evaluation of its safety was articulated by the WHO Scientific Group on Procedures for Investigating Intentional and Unintentional Food Additives (WHO, 1967).

JECFA has encountered several situations in which either the body of available data on a new additive had some limitations or the safety of a food additive for which the Committee had previously assigned an ADI was brought into question by new data. When the Committee feels confident that the use of the substance is safe over the relatively short period of time required to generate and evaluate further safety data, but is not confident that its use is safe over a lifetime, it often establishes a "temporary" ADI, pending the submission of appropriate data to resolve the safety issue on a timetable established by JECFA. When establishing a temporary (numerical) ADI, the Committee always uses a higher than usual safety factor, usually increasing it by a factor of 2. The additional biochemical or toxicological data required for the establishment of an ADI are clearly stated,

and a review of these new data is conducted before expiry of the provisional period. In many cases, long-term studies are requested, but timetables are not met, which means that JECFA has had to extend temporary ADIs for further periods of time. In instances where data have not been forthcoming, JECFA has withdrawn temporary ADIs as a safety precaution.

#### 5.2.6.2 *Pesticides*

The Joint FAO/WHO Meeting on Principles Governing Consumer Safety in Relation to Pesticide Residues indicated that the assessment of the amount of pesticide to which humans can be exposed daily for a lifetime, without injury, was the primary aim of toxicological investigations. The meeting indicated that "when the (toxicological) investigations are completed, it is possible, by the use of scientific judgement, to name the acceptable daily intake" (FAO/WHO, 1962b).

JMPR stated that the following information should be available in order to arrive at an ADI (FAO/WHO, 1964):

- The chemical nature of the residue. Pesticides may undergo chemical changes and are frequently metabolized by the tissues of plants and animals that have been treated with them. Even when a single chemical has been applied, the residue may consist of a number of derivatives with distinct properties, the exact nature of which may differ in animals and plants and in different crops and products.

- The toxicities of the chemicals forming the residue from acute, short-term and long-term studies in animals. In addition, knowledge is required of the metabolism, mechanism of action and possible carcinogenicity of residue chemicals when consumed.

- A sufficient knowledge of the effects of these chemicals in humans.

The principles discussed above were adopted by subsequent Meetings but, as would be expected, have been further developed with time. Thus, the 1968 JMPR (FAO/WHO, 1968) indicated that metabolites would, under certain conditions, be considered to be included in the ADI. Generally, if the metabolites in food commodities are qualitatively and quantitatively the same as those observed in

laboratory test species, the ADI would apply to the parent compound as well as to metabolites. If the metabolites are not identical or not present at the same order of magnitude, separate studies on the metabolites may be necessary. When one or several pesticides are degradation products of another pesticide, a single ADI may be appropriate for the pesticide and its metabolites (e.g. oxydemeton-methyl, demeton-*S*-methyl sulfone and demeton-*S*-methyl) (FAO/WHO, 1989).

The use of the temporary ADI, first proposed by the WHO Scientific Group on Procedures for Investigating Intentional and Unintentional Food Additives (WHO, 1967), was adopted by JMPR. Criteria were set that had to be met prior to the establishment of the temporary ADI. These included the consideration of each chemical on its own merits, the establishment of the temporary ADI for a fixed period (usually 3–5 years) and the subsequent review of original and new data prior to expiry of the provisional period.

The establishment of a temporary ADI has always been accompanied by a requirement for further work by a specified date and, for numerical ADIs, by the application of an increased safety factor. The 1972 JMPR (FAO/WHO, 1973) considered the course of action to be taken if requested data were not forthcoming and indicated that, under these circumstances, the temporary ADI would be withdrawn. It emphasized, however, that such an action

> did not necessarily indicate a potential health hazard, but only that insufficient information is available at the time of review to permit the Meeting to state with reasonable certainty that there is no likelihood of adverse effects on health resulting from ingestion over a prolonged period.

In 1986, JMPR (FAO/WHO, 1986) indicated that the previously utilized terms "Further work or information required" or "Further work or information desirable" were being replaced, the former by the statement "Studies without which the determination of an ADI is impracticable" and the latter by the statement "Studies which will provide information valuable to the continued evaluation of the compound." Not only do these new statements reflect the actual work performed by JMPR much more clearly than the previous terms "required" and "desirable", but they also reflect the Meeting's increasing reluctance to allocate temporary ADIs as well as the desire to continue the evaluation of a compound even after an ADI has been allocated.

In 1988, JMPR (FAO/WHO, 1988a) recommended that temporary ADIs should not be allocated for new compounds and that an ADI should not be allocated in the absence of an adequate database. The Meeting intended that monographs would be published for all chemicals that are reviewed, regardless of whether an ADI is allocated, and that data requirements would be clearly specified for those chemicals with an inadequate database.

#### 5.2.6.3 *Veterinary drug residues*

Recognizing the principles applied in evaluating a substance for the purposes of establishing an ADI in the Principles for the Safety Assessment of Food Additives and Contaminants in Food (IPCS, 1987), the thirty-second JECFA meeting elaborated many of these principles as a framework for the specific assessment of residues of veterinary drugs in food (FAO/WHO, 1988b). Most importantly, where possible and appropriate, they affirmed that an ADI based on determination of a NOAEL from experimental animal or human toxicological data should be used as the end-point of the safety evaluation with use of an appropriate safety factor. The thirty-second meeting of the Committee recognized that in some instances it might be inappropriate to establish an ADI. When it has been determined that establishing an ADI is unnecessary because of a large margin of safety, the recommendation of a maximum residue limit (MRL) is also unnecessary. For example, at the fortieth meeting, an ADI "not specified" was established for the bovine somatotropins (FAO/WHO, 1993). The Committee noted the lack of activity of the recombinant somatotropins and insulin-like growth factor-1 after oral dosing as well as the low amounts and non-toxic nature of the residues of these compounds even at exaggerated doses. The Committee concluded that these results provide an extremely large margin of safety for humans consuming dairy products from animals treated with the recombinant somatotropins and therefore warranted the establishment of an ADI "not specified".

The Committee has noted that an ADI for a drug is usually based on the toxicity of the parent drug rather than on its metabolite or metabolites. However, it may sometimes be necessary to calculate an ADI for individual metabolites. Although most compounds have been evaluated as individual substances, there are instances where an ADI has been established as a group ADI (e.g. streptomycin/

dihydrostreptomycin, enrofloxacin/ciprofloxacin; see section 5.2.8) and where an ADI has been established on a microbiological end-point rather than a toxicological end-point (e.g. spiramycin and spectinomycin). The thirty-eighth meeting of the Committee (FAO/WHO, 1991) noted that if the pharmacological effects are more relevant and sensitive than the toxicological effects, the ADI should be established on the basis of pharmacology.

There have been a limited number of situations where an ADI numerical value or range was not identified. For allergenic considerations, the Committee did not establish an ADI for benzylpenicillin, as there were insufficient data with which to establish a NOEL (FAO/WHO, 1990). The Committee recommended that the daily intake from food should be kept as low as possible (below 0.03 mg/person per day).

The thirty-eighth meeting of the Committee also addressed the issue of establishing ADIs and MRLs for those substances that are rapidly converted to their metabolites when they are administered to the target animal or host (FAO/WHO, 1991). The Committee recognized that there may be occasions when drug metabolites present as residues are responsible for the specific activity of concern possessed by the parent drug. In these situations, the activity of the parent drug would be discounted in establishing the ADI on which to base the MRL; the ADI would instead be based on a toxicological property of the metabolites with an appropriate safety factor applied. In the case of febantel, an ADI was established for febantel per se, based on a study in animals administered the parent compound, but the MRL was established for metabolites, measured as oxfendazole sulfone, using an ADI established for oxfendazole.

The fortieth meeting of the Committee noted that certain conditions apply regarding the identity and quality of veterinary drugs subject to Committee review (FAO/WHO, 1993). The Committee evaluations depend on studies performed with a chemical substance of defined identity, purity and physical form. In particular, the ADI is valid only for substances that do not differ significantly in identity and quality from the material used to generate the data on which the Committee's evaluation is based (see chapter 3).

The thirty-eighth meeting of the Committee (FAO/WHO, 1991) affirmed that in calculating the ADI, the Committee has usually

followed the procedures described in Principles for the Safety Assessment of Food Additives and Contaminants in Food (IPCS, 1987), applying a safety factor to the NOAEL derived from the most relevant and appropriate toxicological, microbiological or pharmacological end-point study. The safety factor usually chosen is 100 in the situation where a NOAEL is derived from a long-term animal study, on the assumptions that humans are 10 times as sensitive as the test animals used in such studies and that a 10-fold range of sensitivity within the human population may exist. When no adverse health effects are seen in long-term studies, an uncertainty factor of 100 may be applied to the NOAEL derived from short-term studies where higher dose levels have been used and an effect has been noted. Typically, acceptable short-term studies need to be at least 3-month studies. The Committee noted, however, that, depending on the quantity, quality and nature of the available data, a safety factor of 100 might be insufficient. This may occur when the required data are incomplete, when the study from which the NOAEL is established is inadequate (e.g. insufficient numbers of animals per test group or when no individual animal data are reported) or when irreversible effects such as teratogenicity or non-genotoxic carcinogenicity are noted. The Committee may employ, and on limited occasions has employed, higher safety factors (e.g. 200, 500 and 1000), depending on the quality and quantity of relevant data. The Committee noted that safety factors are usually not appropriate for genotoxic carcinogens. When the only noteworthy toxicological effects are observed in human studies, a lower safety factor (e.g. 10) may be applied. The Committee stressed that the safety factor applied with each drug would be assessed on its own merits, considering all the above factors.

A different approach is used for the establishment of an ADI based on an effect on the human gut microflora. A decision tree approach that complies with Guideline 36 of the International Cooperation on Harmonisation of Technical Requirements for Registration of Veterinary Medicinal Products (VICH, 2004) has been developed by JECFA. It is used to determine the need to establish a microbiological ADI for the compound under review. The decision tree approach initially seeks to determine if there may be microbiologically active residues entering the human colon. This is done in three steps, in which the questions are:

1) Step 1: Are residues of the drug, and (or) its metabolites, microbiologically active against representatives of the human intestinal flora?

2) Step 2: Do residues enter the human colon?

3) Step 3: Do the residues entering the human colon remain microbiologically active?

If the answer is "no" to any of the first three steps, then no microbiological ADI is necessary. However, should such residues be present, then two end-points of public health concern are to be considered: 1) disruption of the colonization barrier and 2) increase of the population(s) of resistant bacteria.

At Step 4 of the decision tree process, it is possible to provide scientific justification to eliminate testing (i.e. the need for a microbiological ADI) for either one or both end-points.

Step 5 is where a microbiological ADI is determined. Should a microbiological ADI not be necessary, then the toxicological or pharmacological ADI would be used.

The decision tree approach makes use of observations in humans if such data are available. If this is the case, it is reflected in the uncertainty factor used by the Committee. However, the typical situation is that the ADI is based on in vitro minimum inhibitory concentration (MIC) data. The following formula is used to derive a microbiological ADI from in vitro MIC data:

$$ADI = \frac{MIC_{calc} \times \text{Mass of colon content}}{\text{Fraction of oral dose available to microorganisms} \times UF \times 60\ kg}$$

where:

- The $MIC_{calc}$ represents the lower 90% confidence limit for the mean $MIC_{50}$ (the minimum inhibitory concentration for 50% of strains of the most sensitive relevant organism) of the relevant genera for which the drug is active.

- The mass of colon content is assumed to be 220 g/day.
- The fraction of an oral dose available to microorganisms is ideally based on in vivo measurements for the drug administered orally. Alternatively, if sufficient data are available, the fraction of the dose available for colonic microorganisms can be calculated as 1 minus the fraction of an oral dose excreted in urine. Human data are encouraged; in their absence, non-ruminant animal data are recommended. In the absence of data to the contrary, it should be assumed that metabolites have antimicrobial activity equal to that of the parent compound. The fraction may be lowered if the applicant provides quantitative in vitro or in vivo data to show that the drug is inactivated during transit through the intestine.
- UF is the uncertainty factor.
- 60 kg is the standard human body weight used by JECFA.

In these cases, the uncertainty factor is used in an entirely different way than when applied to an ADI based on toxicological data. When establishing a microbiologically based ADI, the safety factor is used to account for uncertainty about the amount and relevance of the MIC data available for review. For example, where microbiological effects were studied directly in humans or in a sufficient number of microorganisms representative of the potentially susceptible fraction of the human gut microflora, an uncertainty factor of 1 may be used. Generally, uncertainty factors considered appropriate for microbiological end-points are 1–10, depending on the quantity and quality of the data.

Several meetings of the Committee on residues of veterinary drugs in food have had substances with limited toxicological data available upon which to establish an ADI. The thirty-sixth meeting of the Committee (FAO/WHO, 1990) noted that when the Committee, in its scientific judgement, is confident that the consumption of residues of the veterinary drug is without toxicological hazard over a limited amount of time (e.g. the amount of time required to generate and evaluate further data for toxicological assessment), but not sufficiently confident that consumption of these residues over a lifetime may not pose a health concern, it may establish a temporary ADI. In applying this approach, the Committee considers whether those data might be made available to the Committee within a relatively short period of time. As is noted below, temporary MRLs may be recommended for similar

or additional reasons, such as the availability of reliable analytical methods or additional information on the nature of the quantification of residues.

Where the Committee has established temporary ADIs, it specifies what information is required to resolve the data needs and sets a date by which the data are requested for re-evaluation by the Committee. The same approach is applied with MRLs. At the reassessment, if one is done, the Committee has the option to 1) establish a full ADI, 2) extend the temporary ADI or 3) not extend the temporary ADI (i.e. the ADI is withdrawn). The same options are available with temporary MRLs. The thirty-sixth meeting of the Committee established a temporary ADI and temporary MRLs for levamisole and requested additional toxicological and residue data for re-evaluation by the Committee (FAO/WHO, 1990). Based on the additional data provided, the forty-second meeting of the Committee established an ADI; however, it withdrew the temporary MRL for levamisole in milk, as no additional data were made available. Similarly, the Committee withdrew the MRL in eggs because of high amounts of residues (FAO/WHO, 1995).

### 5.2.7 Tolerable intakes

JECFA has considered the presence of food contaminants on many occasions since 1972, when mercury, lead and cadmium were first assessed (FAO/WHO, 1972). These food contaminants have included, in addition to heavy metals, environmental contaminants such as dioxins, mycotoxins, impurities arising in food additives, solvents used in food processing, packaging material migrants and residues arising from the use of animal feed additives or the non-active components of veterinary drug formulations. Each of these classes of food contaminants possesses its own unique characteristics and evaluation requirements. Thus, JECFA has recognized through the years that evaluation principles should pertain to classes or groups of contaminants rather than to food contaminants in toto. Guidelines for the evaluation of classes of contaminants are provided in various sections of this report.

When JECFA considered mercury, cadmium and lead in 1972, it established the concept of a PTWI, which was a departure from the traditional ADI concept (FAO/WHO, 1972). JECFA has continued to use this concept, with some modifications, ever since for

contaminants with cumulative properties. The use of the term "provisional" expresses the tentative nature of the evaluation, in view of the paucity of reliable data on the consequences of human exposure at levels approaching those with which JECFA is concerned.

PMTDIs are established for food contaminants that are known not to accumulate in the body. The value assigned to a PMTDI represents permissible human exposure as a result of the background occurrence of the substance in food and also in drinking-water.

For contaminants that may accumulate within the body over a period of time, JECFA has used the PTWI and PTMI. On any particular day, consumption of food containing above-average levels of the contaminant may exceed the proportionate share of its weekly or monthly tolerable intake (TI). JECFA's assessment takes into account such daily variations, its real concern being prolonged exposure to the contaminant, because of its ability to accumulate within the body over a period of time.

The principles for establishing tolerable intakes are the same as for acceptable intakes as described above. For contaminants, there are often epidemiological studies available that can form the basis for derivation of tolerable intakes. If sufficient information is available to perform a dose–response assessment, the POD can be defined from epidemiological studies, and uncertainty factors can then be applied according to the principles outlined above. JECFA often applies the concept of CSAFs when deriving tolerable intakes for contaminants.

### *5.2.8 Group ADIs/TIs*

If several substances that produce similar toxic effects are to be considered for use as food additives, pesticides or veterinary drugs or occur as contaminants (e.g. dioxins), it may be appropriate in establishing an ADI or TI to consider the group of substances in order to limit their overall intake. For this procedure to be feasible, the substances should have a similar mode of action and a similar range of toxic potency. Flexibility should be used in determining which NOAEL is to be used in calculating the ADI or TI. In some cases, the average NOAEL for all the substances in the group may be used for calculating the group ADI. A more conservative approach is to base

the group ADI or TI on the substance with the lowest NOAEL. The relative quality and length of studies on the various substances should be considered when setting the group ADI or TI. When the NOAEL for one of the substances is out of line with the others in the group, it should be treated separately.

When considering a substance that is a member of a series of substances that are very closely related chemically (e.g. fatty acids), but for which toxicological information is limited, it may be possible to base its evaluation on the group ADI or TI established for the series of substances. This procedure can be followed only if a great deal of toxicological information is available on at least one member of the series and if the known toxic properties of the various substances fall along a well-defined continuum. Interpolation, but not extrapolation, can be performed. The use of this procedure by JECFA represents one of the few situations in which the Committee has used structure–activity relationships in its safety assessments.

In some instances, group ADIs can be established primarily on the basis of metabolic information. For example, the safety of esters used as food flavours could be assessed on the basis of toxicological information on their constituent acids and alcohols, provided that it is shown that they are quantitatively hydrolysed in the gut.

The calculation of a group ADI is also appropriate for substances that cause additive physiological or toxic effects, even if they are not closely related chemically. For example, it may be appropriate to establish a group ADI for additives such as bulk sweeteners that are poorly absorbed and cause a laxative effect.

### 5.2.9 Setting of acute reference doses (ARfDs)

#### 5.2.9.1 General considerations

JMPR routinely evaluates the acute and chronic effects of pesticide residues in food and has developed guidance on the setting of ARfDs for pesticides (FAO/WHO, 1999, 2001a,b, 2002c, 2004a; Solecki et al., 2005). The guidance provided in these documents for agricultural pesticides should be of value in general considerations of the necessity of establishing an ARfD, as well as in the specific end-point considerations in the derivation of an ARfD. The text that

follows relates mainly to pesticide residues, but JECFA may apply similar principles to other types of compounds when the establishment of an ARfD is needed.

The ARfD of a chemical refers to the amount of a substance that can be ingested in a period of 24 h or less (see section 5.2.1). Because the ARfD is compared with exposure data for a 24 h period, this will provide a conservative risk assessment for rapidly reversible effects (e.g. cholinesterase inhibition by carbamates) where the ARfD would be applicable to a shorter time period.

The decision as to whether the setting of an ARfD is necessary should be based on the hazard profile of a pesticide, as well as on specific end-points that may be particularly relevant to acute effects. Most of the scientific concepts applying to the setting of ADIs apply equally to the setting of ARfDs (e.g. consideration of the scientific quality of studies, selection of the critical effect). When assessing the need for an ARfD, the entire database should be reviewed using a weight of evidence approach to determine whether adverse effects seen in repeated-dose toxicity studies might be relevant to single exposures. Usually a single ARfD is set, but two values may be required (e.g. one for the general population and one for a subgroup of the population) in exceptional cases. In some cases, it may also be necessary to set an additional ARfD for main metabolites if they occur on crops and are therefore included in the residue definition (e.g. if these metabolites are likely to show an acute toxicity profile that is different from that of the parent compound) or when metabolites formed in humans are not observed in experimental animal metabolism studies.

#### 5.2.9.2 *Practical cut-off value for ARfDs*

Bearing in mind practical considerations, such as the maximum quantity of a particular food likely to be consumed in a single sitting, a value above which the formal setting of an ARfD is unnecessary can be proposed. This practical cut-off value (upper limit) for an ARfD should be considered with reference to the potential range of dietary exposures to an acutely toxic pesticide. For example, the acute exposure to a pesticide used on fruit, for which there is an MRL set according to Good Agricultural Practice (GAP), may be calculated as follows:

- A 50 kg person consumes 500 g of fruit in a single sitting. The fruit consists of a single large item (e.g. a melon) and has been treated with a pesticide having an MRL of, for example, 20 mg/kg. Trial data show that a variability factor[1] of 5 is applicable.

- The estimated maximum exposure could be [20 mg/kg (MRL) × 5 (variability factor) × 0.5 kg (mass)] / 50 kg body weight = 1 mg/kg body weight.

However, further issues need consideration when deciding on a practical cut-off value for ARfDs:

- A small number of pesticide/commodity combinations have MRLs in excess of 20 mg/kg, although they might not have a toxicity profile indicating acute toxicity concern.

- Infants and small children might have a higher rate of consumption relative to body weight.

- For certain commodities, a variability factor greater than 5 might be applicable.

This estimate indicates that any general cut-off for ARfDs should be at a value greater than 1 mg/kg body weight. A value of 5 mg/kg body weight is proposed as a conservative value to cover all eventualities for agricultural pesticides, based on practical considerations on consumption and maximum residue levels in foods. An ARfD cut-off at 5 mg/kg body weight would equate to a NOAEL of 500 mg/kg body weight per day in an animal study, when default uncertainty factors are applied. Thus, if acute toxicity were seen only at doses greater than 500 mg/kg body weight, then there would be no necessity to set an ARfD.

By analogy, relevant upper limits might be considered for other chemicals (e.g. for non-agricultural pesticides).

[1] The variability factor is defined as the ratio of the 97.5th percentile of the distribution of pesticide residue per unit to the mean residue for the lot (ν = 97.5th percentile divided by the mean) (FAO, 2002).

If, during the derivation of an ARfD, it becomes apparent that a previously derived ADI is higher than the ARfD, the ADI should be reconsidered. Such a situation can occur for a number of reasons (e.g. the availability of additional studies, or compounds producing more severe effects when given by gavage than when administered in the diet) (FAO/WHO, 2001b). Even when there is no obvious basis to revise the ADI, it is recommended that the lower of the ARfD and the ADI be used as the ADI.

#### *5.2.9.3 Biological and toxicological considerations*

The following are key points for consideration when evaluating the database regarding the potential for acute toxicity:

- In the absence of data to the contrary, all indications of acute toxicity observed in repeated-dose studies should be considered as potentially relevant to setting an ARfD.

- Particular weight should be given to observations and investigations at the beginning of repeated-dose studies.

- The NOAEL from the most sensitive species should be used unless there is evidence to demonstrate that it is not appropriate for a human risk assessment.

- Isolated findings showing no specificity or clear pattern are not necessarily indications of acute toxicity.

In determining the appropriateness of using doses and end-points from subchronic or chronic toxicity studies to establish an ARfD, a weight of evidence evaluation should be conducted that considers all relevant data. This includes what is known about the toxic mode of action and the pertinent biology of the system that is affected. One of the main challenges is to evaluate whether those effects are also likely to occur at the same observed dose levels following an acute exposure.

Toxicological information from interim results or consideration of progression of a lesion in repeated-dose studies may provide insights into the relevance of end-points for setting ARfDs. For example, if interim data indicate that the response is minimal and

becomes pronounced or severe after increasing exposure duration, then repeated exposures are probably the determining factor in the response. Interpretation of the relevance of end-points should also consider toxicokinetic information that would raise concern for acute toxicity, such as slow elimination kinetics or toxicities dependent on the maximum plasma concentrations ($C_{max}$) achieved, as well as information on the acute toxicity of chemicals with a similar structure.

#### 5.2.9.4 *Stepwise process for setting ARfDs*

The following stepwise process for setting ARfDs for agricultural pesticides is recommended:

- Evaluate the total database for the pesticide, and establish a toxicological profile for the active substance.
- Consider the principles for not needing to set an ARfD:
  - No findings indicative of effects elicited by an acute exposure are observed at doses up to about 500 mg/kg body weight per day; and/or
  - No substance-related mortalities are observed at doses up to 1000 mg/kg body weight in single-dose oral studies; and/or
  - If mortality is the only trigger, the cause of death should be confirmed as being relevant to human exposures.

If a decision is taken at this stage not to set an ARfD, the reasons should be clearly explained.

If the above criteria do not exclude the setting of an ARfD, then one should be set as follows using the most appropriate end-point and safety factors:

- Select the toxicological end-points most relevant for a single (day) exposure (see section 5.2.9.5).
- Select the most relevant study in which these end-points have been adequately investigated.
- Identify the NOAELs for these end-points.
- Select the most relevant end-point providing the lowest NOAEL.

- Derive the ARfD using the most appropriate safety factors[1] (see section 5.2.9.6).

An end-point from a repeated-dose toxicity study should be used if the critical effect of the compound has not been adequately evaluated in a single-dose study. This is likely to be a more conservative approach and should be stated as such. This does not mean that a safety factor other than the default value should be applied. A refinement of such a NOAEL (e.g. in a special single-dose study) may be necessary if the acute intake estimation (see section 5.2.9.9) exceeds such a potentially conservatively established ARfD. This will be necessary only for a very limited number of substances, according to a retrospective analysis (Moeller et al., 2009). Under the Organisation for Economic Co-operation and Development test guidelines programme, a document is under development on "Guidance for a single-dose study" (OECD, 2009), based on the guidance developed by JMPR, to inform investigators should a specific study be necessary as a basis for derivation of the ARfD.

If at this stage, after consideration of all the end-points, an ARfD is not set, then the reasons should be clearly explained.

### 5.2.9.5 *Toxicological end-points relevant for ARfD derivation*

A number of effects could be caused by a single exposure. The relevance of these effects for ARfD derivation should be considered on a case-by-case basis. The route of substance administration should be considered carefully with regard to available toxicokinetic data, in order to minimize influences that are not relevant for the intake of residues (e.g. effects induced by gavage or by a specific vehicle or formulation used).

The following list of target effects is not an exhaustive list of all possible relevant end-points (FAO/WHO, 2004a), but these toxic mechanisms are regarded as alerts for acute toxicity, relevant for the consideration of the need to set an ARfD:

[1] The term "safety factor" is based on current JMPR terminology and applied as a synonym for the terms "uncertainty factor", "adjustment factor" and "assessment factor" used by other bodies.

- *Haematotoxicity*: The induction of methaemoglobinaemia is regarded as a critical effect in consideration of acute responses to chemical exposure. Haemolytic anaemia is considered to be less relevant for ARfD derivation, as the severity of such an effect generally appears to depend on prolonged exposure.

- *Immunotoxicity*: Immunotoxicity data derived from subchronic studies are not likely to be appropriate for setting an ARfD for acute adult exposure limits, because immune system cells are constantly replaced and because of inherent redundancy in the system.

- *Neurotoxicity*: Any neurotoxicity seen in repeated-dose studies could be the result of a single exposure that is not repairable; thus, any evidence of neurotoxicity should be considered relevant to the setting of an ARfD, unless it can be demonstrated that the effects are produced only after repeated exposures.

- *Kidney and liver effects*: If the effects on these organs cannot be discounted as being either adaptive or the result of prolonged exposure, an ARfD can be derived on the basis of such effects. Such an ARfD is likely to be conservative, and it may be possible to subsequently refine it using an appropriately designed single-dose study.

- *Endocrine effects*: In general, adverse effects on the endocrine system observed in routine toxicological testing for regulatory purposes—other than those affecting female reproduction and development of the offspring—are considered to be unlikely to arise as a consequence of acute exposure. However, exceptions may occur, and a case-by-case analysis is required.

- *Developmental effects*: Any treatment-related adverse effect on embryos, fetuses or offspring that has resulted from exposure during any phase of development should be considered as potentially appropriate to use in acute dietary risk assessment, despite the fact that the treatment period typically consists of repeated dosing, as it could be the result of a single exposure during a critical window of development.

Direct effects on the gastrointestinal tract or stomach should be assessed carefully to determine their relevance to human exposure.

Considerations would include whether they are due to irritation or a pharmacological action or whether they are related to the method of administration (e.g. occur with bolus dosing but not with incorporation into the diet). Similarly, diarrhoea and vomiting in dogs should be considered as not relevant for setting an ARfD if these effects are related to high concentrations following specific dosing methods (e.g. capsule administration or gavage) and local (irritant) effects.

Other findings relevant for setting an ARfD, such as clinical signs, changes in body weight/body weight gain, changes in food and/or water intake and mortalities observed after one or several doses in repeat oral exposure toxicity studies, may suggest the need to establish an ARfD.

#### 5.2.9.6 *Uncertainty factors for ARfDs*

The process for deriving an ARfD is essentially the same as that for deriving an ADI, involving the identification of the most appropriate NOAEL (or BMDL) and application of safety factors, usually 100-fold or 10-fold for data from studies in experimental animals or humans, respectively. Safety factors are used to extrapolate from animals to the average human and to allow for variation in sensitivity within the human population. The default factor of 10 for extrapolating from laboratory animals to humans can be subdivided into 2.5 for toxicodynamics and 4 for toxicokinetics, whereas the default human variability factor of 10 can be subdivided into identical factors of 3.2 for both toxicokinetics and toxicodynamics (IPCS, 2005), as described above under the concept of CSAFs (section 5.2.3).

A number of other situations may justify the use of safety factors higher or lower than the default values of 100 or 10 that are conventionally used on the basis of experimental animal or human data, respectively (FAO/WHO, 2001a). Such situations may arise when certain types of data are available. For example, data on the mode of toxic action are often available for chemicals such as veterinary medicines and pesticides that have a common mechanism against both the target species and non-target mammals. These data, together with information on the time course of effects, can provide an indication as to whether the action is reversible. Data on absorption, excretion and toxicokinetics, together with information on the mode of action, may

help to assess whether effects are likely to be related to $C_{max}$ or AUC. Human toxicity data are available for a small number of chemicals and can be used either directly to derive ARfDs or as part of the overall consideration of interspecies sensitivity.

When the effect under consideration is due to reversible interaction of the substance with a pharmacological target (e.g. a receptor or ion channel), then the concentration of the substance rather than total exposure should determine the magnitude of the effect (i.e. the $C_{max}$ is likely to be more relevant than the AUC). Similarly, if the effect of concern is due to direct irritation, then the concentration at the site of action is more relevant than the total exposure expressed on a body weight basis. In such cases, there will be less interspecies and interindividual variation in toxicokinetics; this would justify a 2-fold reduction in the respective safety factors, leading to an overall composite factor of 25 for extrapolation from animal studies (i.e. 5 × 5 instead of 10 × 10 for interspecies and intraspecies factors) and 5 (instead of 10) for human studies.

JMPR has used such categorical factors in the derivation of ARfDs for several carbamate insecticides that inhibit acetylcholinesterase (FAO/WHO, 2009a). These compounds do not require metabolic activation, they react reversibly with a pharmacological target (acetylcholinesterase), the magnitude of the pharmacological effect is proportional to the $C_{max}$ rather than the AUC and the excretion is rapid. In such circumstances, the determining factor is the $C_{max}$, which has been shown to have lower variability than clearance, as it depends mainly on the rate and extent of gastrointestinal absorption. This reduced variability in toxicokinetics is used by JMPR to derive a composite factor that is 50% of the default value.

If human data are available but are not used directly to derive the ARfD, they might be sufficient to demonstrate that the findings in experimental animals are qualitatively and quantitatively similar to those in humans, thereby supporting the use of a reduced, data-derived factor (e.g. data on the production and degradation of a toxic metabolite). Similarly, if data show that a wide range of species exhibit similar qualitative and quantitative effects, it could be possible to conclude that the variation between the most sensitive of these and humans would be less than 10.

A reduced safety factor might also be appropriate if the end-point used to derive an ARfD is of minimal adversity and the critical NOAEL is from a repeated-dose study (e.g. increased organ weight with minimal pathological change, or reduced food consumption and body weight gain observed in the first days of dosing). When considering whether body weight changes are relevant for setting of an ARfD, consideration should be given to potential problems of palatability of the feed.

When a NOAEL has not been identified for the most appropriate end-point, the LOAEL can be used in exceptional cases as the basis for the ARfD. In such a situation, the selection of an additional safety factor up to 10 will depend upon the magnitude of the effect and the steepness of the dose–response curve. If dose spacing results in a LOAEL that is markedly higher than the NOAEL, then the BMD approach, with the usual safety factor, would be a better alternative for defining the ARfD.

An extra uncertainty factor has sometimes been adopted for the severity of the effect. However, judging the degree of severity of an effect may be somewhat subjective, and it would not be feasible to grade all possible toxicological effects by their severity. Therefore, if a toxicological effect is judged to be irreversible or particularly severe, this should be a trigger to consider the finding in more detail before choosing an appropriate uncertainty factor. The following considerations may be helpful:

- Has the study shown an adequate margin between the NOAEL and the LOAEL?
- Is the finding supported by data from other studies or by knowledge of the mode of action of the compound?
- Is there a high level of uncertainty in the database?
- Have measurements been taken at appropriate times, and have they used appropriately sensitive methods?
- Has the study on which it is proposed to base the ARfD used adequate group sizes?

In determining the appropriate uncertainty factors for deriving an ARfD, a stepwise approach is proposed:

- Determine if the data are adequate to support the derivation of scientifically based assessment factors (i.e. CSAFs).

- If CSAFs cannot be derived, consider if there is any other information to indicate reduced or increased uncertainty. If not, the 10-fold or 100-fold default should be used.

- Whenever an uncertainty factor other than the default is used, a clear explanation of the derivation of the factor should be provided.

#### 5.2.9.7 *Different ARfDs for population subgroups*

It is preferable, especially for clarity of subsequent risk management and enforcement, to set a single ARfD to cover the whole population. It is important to ensure that any ARfDs established are adequate to protect the embryo or fetus from possible in utero effects. Although an ARfD based on developmental (embryo/fetal) effects would necessarily apply to women of childbearing age, it is recognized that such an ARfD may be conservative and not relevant to other population subgroups. This may also be the case for children 1–6 years of age for whom specific acute consumption data are available and thus can be separately modelled with respect to acute dietary intake of pesticide residues. The use of an ARfD for a sensitive end-point in pregnant women could lead to an unreasonably conservative short-term dietary risk assessment for the population as a whole. Thus, in those situations in which a developmental end-point drives an ARfD for a compound exhibiting no other toxicity at the developmental NOAEL, consideration could be given to setting a second value based on another (non-developmental) end-point for the rest of the population.

#### 5.2.9.8 *Use of human data in setting ARfDs*

Human data on a pesticide can be extremely valuable in setting the laboratory animal data into context and, when available, should always be evaluated, even if they are not used to derive an ARfD. Not only may a human study sometimes allow identification of end-points (NOAELs/LOAELs) for use in risk assessment, other important information may be gained, such as the nature of the adverse effect and its pattern of onset and duration.

Human data may be available from a number of sources, including epidemiological studies of acute effects in human populations exposed to the chemical, direct administration to volunteers, monitoring of those exposed following normal use of the chemical, exposures from accidental or deliberate poisonings, and exposures from use of the same substances as human pharmaceuticals. Such studies often involve single or short-term exposures that can be of relevance, directly or indirectly, to the derivation of ARfDs.

Further guidance from JMPR on the use of human data for setting ARfDs can be found in the review by Solecki et al. (2005).

#### *5.2.9.9 Intake considerations in relation to ARfDs*

For risk characterization purposes, the ARfD of a compound is compared with the estimated acute intake of a pesticide through various foods. This allows risk managers to identify for which crops and pesticide applications regulatory actions may be necessary for public health protection. The methodology for estimating acute dietary intakes for pesticides is described in detail in chapter 6.

#### *5.2.9.10 Specific guidance on the derivation of ARfDs*

JMPR has given more detailed consideration to the use of particular toxicological end-points (as outlined in section 5.2.9.5) that are relevant to the establishing of ARfDs. This guidance can be found in the review by Solecki et al. (2005). The guidance is not intended to cover all potentially relevant end-points comprehensively but focuses on the interpretation of those that have proved to be problematic in reaching a decision as to whether an effect is relevant to an acute exposure to residues of agricultural pesticides in foods.

## 5.3 References[1]

Barnes DG, Daston GP, Evans JS, Jarabek AM, Kavlock RJ, Kimmel CA, Park C & Spitzer HL (1995) Benchmark dose workshop: criteria for use of a benchmark dose to estimate a reference dose. Regul Toxicol Pharmacol, **21**: 296–306.

[1] Internet links provided in these references were active as of the date of final editing.

Crump KS (1984) A new method for determining allowable daily intakes. Fundam Appl Toxicol, **4**: 854–871.

Dorne JLCM & Renwick AG (2005) Refinement of uncertainty/safety factors in risk assessment by the incorporation of data on toxicokinetic variability in humans. Toxicol Sci, **86**: 20–26.

Dorne JLCM, Walton K & Renwick AG (2005) Human variability in xenobiotic metabolism and pathway-related uncertainty factors for chemical risk assessment: a review. Food Chem Toxicol, **43**: 203–216.

Dourson ML & Stara JF (1983) Regulatory history and experimental support of uncertainty (safety) factors. Regul Toxicol Pharmacol, **3**: 224–238.

FAO (2002) Submission and evaluation of pesticide residues data for the estimation of maximum residue levels in food and feed, 1st ed. Rome, Food and Agriculture Organization of the United Nations (FAO Plant Production and Protection Paper, No. 170; http://www.fao.org/ag/agp/agpp/pesticid/JMPR/Download/FAOM2002.pdf).

FAO/WHO (1957) General principles governing the use of food additives. First report of the Joint FAO/WHO Expert Committee on Food Additives. Rome, Food and Agriculture Organization of the United Nations; Geneva, World Health Organization (FAO Nutrition Meetings Report Series, No. 15; WHO Technical Report Series, No. 129; http://whqlibdoc.who.int/trs/WHO_TRS_129.pdf).

FAO/WHO (1958) Procedures for the testing of intentional food additives to establish their safety for use. Second report of the Joint FAO/WHO Expert Committee on Food Additives. Rome, Food and Agriculture Organization of the United Nations; Geneva, World Health Organization (FAO Nutrition Meetings Report Series, No. 17; WHO Technical Report Series, No. 144; http://whqlibdoc.who.int/trs/WHO_TRS_144.pdf).

FAO/WHO (1962a) Evaluation of the toxicity of a number of antimicrobials and antioxidants. Sixth report of the Joint FAO/WHO Expert Committee on Food Additives. Rome, Food and Agriculture Organization of the United Nations; Geneva, World Health Organization (FAO Nutrition Meetings Report Series, No. 31; WHO Technical Report Series, No. 228; http://whqlibdoc.who.int/trs/WHO_TRS_228.pdf).

FAO/WHO (1962b) Principles governing consumer safety in relation to pesticide residues. Report of a meeting of a WHO Expert Committee on Pesticide Residues held jointly with the FAO Panel of Experts on the Use of Pesticides in Agriculture. Rome, Food and Agriculture Organization of the United Nations; Geneva, World Health Organization (FAO Plant Production and Protection Division Report, No. PL/1961/11; WHO Technical Report Series, No. 240).

FAO/WHO (1964) Evaluation of the toxicity of pesticide residues in food. Report of a Joint Meeting of the FAO Committee on Pesticides in Agriculture and the WHO Expert Committee on Pesticide Residues. Rome, Food and Agriculture

Organization of the United Nations; Geneva, World Health Organization (FAO Meeting Report, No. PL/1963/13; WHO/Food Add./23).

FAO/WHO (1968) Pesticide residues in food—1968. Report of the Joint Meeting of the FAO Working Party of Experts on Pesticide Residues and the WHO Expert Committee on Pesticide Residues. Rome, Food and Agriculture Organization of the United Nations; Geneva, World Health Organization (FAO Agricultural Studies, No. 78; WHO Technical Report Series, No. 417; http://whqlibdoc.who.int/trs/WHO_TRS_417.pdf).

FAO/WHO (1972) Evaluation of certain food additives and the contaminants mercury, lead, and cadmium. Sixteenth report of the Joint FAO/WHO Expert Committee on Food Additives. Geneva, World Health Organization (WHO Technical Report Series, No. 505 and corrigendum; http://whqlibdoc.who.int/trs/WHO_TRS_505.pdf).

FAO/WHO (1973) Pesticide residues in food—1972. Report of the Joint Meeting of the FAO Working Party of Experts on Pesticide Residues and the WHO Expert Committee on Pesticide Residues. Rome, Food and Agriculture Organization of the United Nations; Geneva, World Health Organization (FAO Agricultural Studies, No. 90; WHO Technical Report Series, No. 525; http://whqlibdoc.who.int/trs/WHO_TRS_525.pdf).

FAO/WHO (1974) Toxicological evaluation of certain food additives with a review of general principles and of specifications. Seventeenth report of the Joint FAO/WHO Expert Committee on Food Additives. Rome, Food and Agriculture Organization of the United Nations; Geneva, World Health Organization (FAO Nutrition Meetings Report Series, No. 53; WHO Technical Report Series, No. 539 and corrigendum; http://whqlibdoc.who.int/trs/WHO_TRS_539.pdf).

FAO/WHO (1984) Evaluation of certain food additives and contaminants. Twenty-eighth report of the Joint FAO/WHO Expert Committee on Food Additives. Geneva, World Health Organization (WHO Technical Report Series, No. 710; http://whqlibdoc.who.int/trs/WHO_TRS_710.pdf).

FAO/WHO (1986) Pesticide residues in food—1986. Report of the Joint Meeting of the FAO Panel of Experts on Pesticide Residues in Food and the Environment and the WHO Expert Group on Pesticide Residues. Rome, Food and Agriculture Organization of the United Nations (FAO Plant Production and Protection Paper, No. 77).

FAO/WHO (1988a) Pesticide residues in food—1988. Report of the Joint Meeting of the FAO Panel of Experts on Pesticide Residues in Food and the Environment and the WHO Expert Group on Pesticide Residues. Rome, Food and Agriculture Organization of the United Nations (FAO Plant Production and Protection Paper, No. 92).

FAO/WHO (1988b) Evaluation of certain veterinary drug residues in food. Thirty-second report of the Joint FAO/WHO Expert Committee on Food Additives.

Geneva, World Health Organization (WHO Technical Report Series, No. 763; http://whqlibdoc.who.int/trs/WHO_TRS_763.pdf).

FAO/WHO (1989) Pesticide residues in food—1989. Report of the Joint Meeting of the FAO Panel of Experts on Pesticide Residues in Food and the Environment and the WHO Expert Group on Pesticide Residues. Rome, Food and Agriculture Organization of the United Nations (FAO Plant Production and Protection Paper, No. 99).

FAO/WHO (1990) Evaluation of certain veterinary drug residues in food. Thirty-sixth report of the Joint FAO/WHO Expert Committee on Food Additives. Geneva, World Health Organization (WHO Technical Report Series, No. 799; http://whqlibdoc.who.int/trs/WHO_TRS_799.pdf).

FAO/WHO (1991) Evaluation of certain veterinary drug residues in food. Thirty-eighth report of the Joint FAO/WHO Expert Committee on Food Additives. Geneva, World Health Organization (WHO Technical Report Series, No. 815; http://whqlibdoc.who.int/trs/WHO_TRS_815.pdf).

FAO/WHO (1993) Evaluation of certain veterinary drug residues in food. Fortieth report of the Joint FAO/WHO Expert Committee on Food Additives. Geneva, World Health Organization (WHO Technical Report Series, No. 832; http://whqlibdoc.who.int/trs/WHO_TRS_832.pdf).

FAO/WHO (1995) Evaluation of certain veterinary drug residues in food. Forty-second report of the Joint FAO/WHO Expert Committee on Food Additives. Geneva, World Health Organization (WHO Technical Report Series, No. 851; http://whqlibdoc.who.int/trs/WHO_TRS_851.pdf).

FAO/WHO) (1999) Procedures for estimating acute reference dose. In: Pesticide residues in food—1998. Report of the Joint Meeting of the FAO Panel of Experts on Pesticide Residues in Food and the Environment and the WHO Core Assessment Group on Pesticide Residues. Rome, Food and Agriculture Organization of the United Nations (FAO Plant Production and Protection Paper, No. 148; http://www.fao.org/ag/AGP/AGPP/Pesticid/JMPR/Download/98_rep/REPORT1998.pdf).

FAO/WHO (2001a) Proposed guidance for interpretation of data generated in studies with single oral doses. In: Pesticide residues in food—2000. Report of the Joint Meeting of the FAO Panel of Experts on Pesticide Residues in Food and the Environment and the WHO Core Assessment Group on Pesticide Residues. Rome, Food and Agriculture Organization of the United Nations (FAO Plant Production and Protection Paper, No. 163; http://www.fao.org/ag/AGP/AGPP/Pesticid/JMPR/Download/2000_rep/cont.pdf).

FAO/WHO (2001b) Further issues related to establishing the acute reference dose. In: Pesticide residues in food—2001. Report of the Joint Meeting of the FAO Panel of Experts on Pesticide Residues in Food and the Environment and the WHO Core Assessment Group on Pesticide Residues. Rome, Food

and Agriculture Organization of the United Nations (FAO Plant Production and Protection Paper, No. 167; http://www.fao.org/ag/AGP/AGPP/Pesticid/JMPR/Download/2001_rep/REPORT2001.pdf).

FAO/WHO (2002a) Pesticide residues in food—2002. Report of the Joint Meeting of the FAO Panel of Experts on Pesticide Residues in Food and the Environment and the WHO Core Assessment Group on Pesticide Residues. Rome, Food and Agriculture Organization of the United Nations (FAO Plant Production and Protection Paper, No. 172; http://www.fao.org/ag/AGP/AGPP/Pesticid/JMPR/Download/2002_rep/2002JMPRReport2.pdf).

FAO/WHO (2002b) Evaluation of certain food additives and contaminants. Fifty-seventh report of the Joint FAO/WHO Expert Committee on Food Additives. Geneva, World Health Organization (WHO Technical Report Series, No. 909; http://whqlibdoc.who.int/trs/WHO_TRS_909.pdf).

FAO/WHO (2002c) Further guidance on derivation of the ARfD. In: Pesticide residues in food—2002. Report of the Joint Meeting of the FAO Panel of Experts on Pesticide Residues in Food and the Environment and the WHO Core Assessment Group on Pesticide Residues. Rome, Food and Agriculture Organization of the United Nations (FAO Plant Production and Protection Paper, No. 172; http://www.fao.org/ag/AGP/AGPP/Pesticid/JMPR/Download/2002_rep/2002JMPRReport2.pdf).

FAO/WHO (2004a) Guidance on the establishment of acute reference doses. In: Pesticide residues in food—2004. Report of the Joint Meeting of the FAO Panel of Experts on Pesticide Residues in Food and the Environment and the WHO Core Assessment Group on Pesticide Residues. Rome, Food and Agriculture Organization of the United Nations (FAO Plant Production and Protection Paper, No. 178; http://www.fao.org/ag/AGP/AGPP/Pesticid/JMPR/DOWNLOAD/2004_rep/report2004jmpr.pdf).

FAO/WHO (2004b) Pesticide residues in food—2004. Report of the Joint Meeting of the FAO Panel of Experts on Pesticide Residues in Food and the Environment and the WHO Core Assessment Group on Pesticide Residues. Rome, Food and Agriculture Organization of the United Nations (FAO Plant Production and Protection Paper, No. 178; http://www.fao.org/ag/AGP/AGPP/Pesticid/JMPR/DOWNLOAD/2004_rep/report2004jmpr.pdf).

FAO/WHO (2006) Evaluation of certain food contaminants. Sixty-fourth report of the Joint FAO/WHO Expert Committee on Food Additives. Geneva, World Health Organization (WHO Technical Report Series, No. 930; http://whqlibdoc.who.int/trs/WHO_TRS_930_eng.pdf).

FAO/WHO (2009a) Pesticide residues in food—2008. Report of the Joint Meeting of the FAO Panel of Experts on Pesticide Residues in Food and the Environment and the WHO Core Assessment Group on Pesticide Residues. Rome, Food and Agriculture Organization of the United Nations (FAO Plant Production and Protection Paper, No. 193; http://www.fao.org/fileadmin/templates/agphome/documents/Pests_Pesticides/JMPR/JMPRReport08.pdf).

FAO/WHO (2009b) Evaluation of certain food additives and contaminants. Sixty-ninth report of the Joint FAO/WHO Expert Committee on Food Additives. Geneva, World Health Organization (WHO Technical Report Series, No. 952; http://whqlibdoc.who.int/trs/WHO_TRS_952_eng.pdf).

Gelman A, Carlin JB, Stern HS & Rubin DB (2004) Bayesian data analysis, 2nd ed. Boca Raton, FL, Chapman and Hall/CRC.

Hasselblad V & Jarabek AM (1995) Dose–response analysis of toxic chemicals. In: Berry DA & Stangl DK eds. Bayesian biostatistics. New York, NY, Marcel Dekker.

IPCS (1987) Principles for the safety assessment of food additives and contaminants in food. Geneva, World Health Organization, International Programme on Chemical Safety (Environmental Health Criteria, No. 70; http://www.inchem.org/documents/ehc/ehc/ehc70.htm).

IPCS (1990) Principles for the toxicological assessment of pesticide residues in food. Geneva, World Health Organization, International Programme on Chemical Safety (Environmental Health Criteria, No. 104; http://www.inchem.org/documents/ehc/ehc/ehc104.htm).

IPCS (1994) Assessing human health risks of chemicals: derivation of guidance values for health-based exposure limits. Geneva, World Health Organization, International Programme on Chemical Safety (Environmental Health Criteria, No. 170; http://www.inchem.org/documents/ehc/ehc/ehc170.htm).

IPCS (2004) IPCS risk assessment terminology. Geneva, World Health Organization, International Programme on Chemical Safety (Harmonization Project Document, No. 1; http://www.who.int/ipcs/methods/harmonization/areas/ipcsterminologyparts1and2.pdf).

IPCS (2005) Chemical-specific adjustment factors for interspecies differences and human variability: guidance document for use of data in dose/concentration–response assessment. Geneva, World Health Organization, International Programme on Chemical Safety (Harmonization Project Document, No. 2; http://whqlibdoc.who.int/publications/2005/9241546786_eng.pdf).

IPCS (2009) Principles for modelling dose–response for the risk assessment of chemicals. Geneva, World Health Organization, International Programme on Chemical Safety (Environmental Health Criteria, No. 239; http://whqlibdoc.who.int/publications/2009/9789241572392_eng.pdf).

Kimmel CA & Gaylor DW (1988) Issues in qualitative and quantitative risk analysis for developmental toxicology. Risk Anal, **8**: 15–20.

Meek ME, Renwick A, Ohanian E, Dourson M, Lake B, Naumann BD & Vu V (2002) Guidelines for application of chemical-specific adjustment factors in dose/concentration–response assessment. Toxicology, **181–182**: 115–120.

Moeller T, Stein B & Solecki R (2009) Retrospective evaluation of acute reference doses (ARfD) for pesticides in the European Union. Toxicol Lett, **189**(Suppl 1): S209–S210.

OECD (2009) Project 4.27: GD for a single dose study for the derivation of an acute reference dose (ARfD). In: Work plan for the Test Guidelines Programme (TGP). Paris, Organisation for Economic Co-operation and Development, Environment Health and Safety Division, Environment Directorate (http://www.oecd.org/dataoecd/50/63/43723559.pdf).

Renwick AG & Lazarus NR (1998) Human variability and noncancer risk assessment—an analysis of the default uncertainty factor. Regul Toxicol Pharmacol, **27**: 3–20.

Solecki R, Davies L, Dellarco V, Dewhurst I, Van Raaij M & Tritscher A (2005) Guidance on setting of acute reference dose (ARfD) for pesticides. Food Chem Toxicol, **43**: 1569–1593.

VICH (2004) Studies to evaluate the safety of residues of veterinary drugs in human food: general approach to establish a microbiological ADI. Brussels, International Cooperation on Harmonisation of Technical Requirements for Registration of Veterinary Medicinal Products (VICH GL36; http://www.vichsec.org/pdf/05_2004/Gl36_st7_F_rev.pdf).

Walton K, Dorne JL & Renwick AG (2001) Categorical default factors for interspecies differences in the major routes of xenobiotic elimination. Hum Ecol Risk Assess, **7**: 181–201.

WHO (1967) Procedures for investigating intentional and unintentional food additives. Report of a WHO Scientific Group. Geneva, World Health Organization (WHO Technical Report Series, No. 348).