CHAPTER 5: DOSE–RESPONSE ASSESSMENT

5.1 Introduction
Risk assessment approaches generally take one of two forms: analyses that provide a qualitative (or sometimes just qualitative) estimation of risk, and analyses that establish health-based guidance values, such as an ADI or TDI, which are levels of human exposure considered to be “without appreciable health risk” (IPCS, 1987, 1990). 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 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 effect data 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 JECFA and JMPR in three main ways (see chapter 7):

1) derivation of a health-based guidance value, such as an ADI, TDI or ARfD;
2) estimation of the margin of exposure (MOE) between a defined point on the dose–response relationship and the level of human exposure; and
3) quantification of the magnitude of the risk at a prespecified level or 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 prespecified level of risk, such as a 1 in 10^{-6} 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 relationship. 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 effect(s) and the highest experimental dose that does not produce a measurable effect, the 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(s) 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 prespecified 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 a draft EHC on Principles for Modelling Dose–Response for the Risk Assessment of Chemicals that has been developed as part of the IPCS Harmonization Project on Approaches to the Assessment of Risk from Exposure to Chemicals. The Dose–Response Modelling Planning Group met in October 2002 in Geneva to develop an outline, and a first draft working paper was discussed at the Task Group Meeting in September 2004 and subsequently revised and edited into the draft EHC (IPCS, 2008).

The draft EHC (IPCS, 2008) covers toxicants with threshold effects and those for which there may be no practical threshold, such as substances that are genotoxic and carcinogenic. The draft EHC 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. The draft EHC 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.
5.2 Dose–response assessment: basic concepts

5.2.1 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 a consequence of absorption, distribution, 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 section 4.3). Biomarkers of internal dose, 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 parent compound alone or the parent compound plus first-pass metabolites (see section 4.3). An additional consideration for tissue dose is whether the dose metric is the peak concentration or a time-weighted average, such as the AUC.

Two parameters are important determinants of dose: the dose frequency and 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), total body burden (e.g. ng/kg body weight), body burden averaged over a given period of time or tissue concentration.

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 arose (e.g. with the dioxin database) was that measurements of the biomarker were made many years after what was believed to be the period of highest exposure (WHO, 2002).

Sometimes the doses used in an animal study are transformed to the equivalent human exposures prior to dose–response modelling (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 animals and humans of the events controlling absorption, tissue distribution, metabolism, excretion and the other molecular and biochemical processes that ultimately lead to particular responses. Interspecies extrapolation of such a dose metric may be possible by the use of a PBTK model. While 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 5.3.2).
5.2.2 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 and/or tissue specific and have different degrees of variation across individuals. DRM can address each response, provide insight into their quantitative similarity across species and tissues and link responses in a mechanistically reasonable manner.

Response is generally considered to vary across experimental units (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 average value) and their effective range (usually based on the standard deviation).

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

- **Quantal 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).

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

- **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).

- **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 and/or time.

5.2.3 Models

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.

The simplest dose–response model is a linear model to describe a continuous response for which the key components are:

- assumptions: the mean added response is proportional to dose;
- functional form: \( R(D) = \alpha + \beta \cdot D \) where \( R(D) \) is the mean response as a function of dose \( D \); and
- parameters: \( \alpha \) is a parameter describing the mean response in the control (unexposed) group and \( \beta \) is a parameter describing the mean change in response per unit dose.

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.

### 5.2.4 Dose–response assessment and dose–response modelling

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, 2008). Dose–response analysis provides the linkage of a model to dose–response data for the purposes of predicting response to a 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.

<table>
<thead>
<tr>
<th>Step</th>
<th>Description</th>
<th>Options</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Data selection</td>
<td>Determine the response to be modelled and select appropriate data</td>
<td>End-point, quality, sample size, utility, availability</td>
</tr>
<tr>
<td>2. Model selection</td>
<td>Choose the type of model to be applied to the data</td>
<td>End-point, data availability, purpose</td>
</tr>
<tr>
<td>3. Statistical linkage</td>
<td>Assume what statistical distributions describe the response</td>
<td>End-point, data type, model choice, software availability</td>
</tr>
<tr>
<td>4. Parameter estimation</td>
<td>Combine the first three steps in an appropriate computer program to obtain estimates of the model parameters</td>
<td>Linkage function, software availability, variance</td>
</tr>
<tr>
<td>5. Implementation</td>
<td>Use the estimated model parameters and the model formula to predict response/dose as needed</td>
<td>Outputs, target selection, model predictions, BMD, direct extrapolation</td>
</tr>
</tbody>
</table>
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, while 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, 2008). 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 generally are 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 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. However, a considerable amount of DRM has been done by simpler linkage functions, such as drawing a straight line through the data points. 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(s) to the data. Since 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 linkage 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 often do not meet the criterion of transparency.

Step 5 is to make the inferences necessary to address the risk assessment questions developed at the problem formulation stage. Usually, when making a prediction, the emphasis is on the change in response seen in treated animals compared with response (baseline data) in the control animals. 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 impact 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.

DRM may be used to develop risk assessment advice in a variety of ways:

<table>
<thead>
<tr>
<th>Step</th>
<th>Description</th>
<th>Options</th>
</tr>
</thead>
<tbody>
<tr>
<td>6. Evaluation</td>
<td>Examine the sensitivity of the resulting predictions to the assumptions used in the analysis.</td>
<td>Model comparison, uncertainty</td>
</tr>
</tbody>
</table>
1) DRM may be used to define levels of exposure where the response measurement is assumed to be virtually unchanged relative to the control measurement. In this approach, a simple functional model such as a straight line can be used. An alternative and simpler approach is the use of pairwise comparisons of the data for different dose levels to define the NOAEL, which is used as a POD for the observed dose–response data.

2) The dose–response model(s) may be used to identify a dose with a known level of response at or slightly below the observable range. The prespecified level of response is known as the benchmark response (BMR), and the dose associated with that response the BMD. The lower one-sided confidence limit on 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(s) 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.

Approach 1 is used currently 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 (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 (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 detoxication processes, DNA 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.
5.2.5 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: those aimed at providing estimates of risk for exposures outside of the range of the data used in the dose–response analysis, and 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 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: extrapolating from the test species to humans and allowing for possible human differences in response. The methods employed for these extrapolation issues are varied, ranging from the use of uncertainty factors (see sections 5.3.2 and 5.3.3) to more complicated modelling schemes based upon differences in toxicokinetics and toxicodynamics between humans and animals and between different human individuals.

Extrapolation issues are most commonly dealt with by using uncertainty factors that have been developed over years of experience. Alternatively, an MOE may be calculated as the numerical ratio of the point of departure on the dose–response data divided by the estimated human exposure; however, such a ratio does not allow for issues of extrapolation, and these will need to be discussed as part of the MOE analysis.

5.3 Dose–response modelling

5.3.1 General perspectives

The NOAEL approach identifies the POD as a single dose that is assumed to be without appreciable effect, whereas the BMD approach provides a POD based on data from the entire observed dose–response range for the critical effect. The BMD approach thus reflects the characteristics of the dose–response curve, particularly in providing estimates of the slope. In the case of a regression framework, it provides the standard error and confidence interval for the model parameters. The NOAEL approach relies on the concept that the NOAEL is sufficiently distant from the lowest dose associated with an effect within the observable range that it will be without appreciable risk under the conditions of the experiment giving the dose–response data, but the magnitude of any possible effect at the NOAEL is not defined. In contrast, the BMD is a dose for which the size of the effect has been predefined, and thus it is under the control of the risk assessor.

Selection of the data used to calculate a health-based guidance value is similar for both the NOAEL and BMD approaches, although the different methods mean that optimum data
sets will have different characteristics. 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. An important criterion of the NOAEL approach is that the group sizes should be sufficient to ensure that any undetected effect at the NOAEL would be toxicologically insignificant. In the BMD approach, a NOAEL does not have to be established, but doses with graded responses are needed to provide optimum model output.

Calculations of a health-based guidance value based on the NOAEL or BMD approach are summarized in Table 5.2.

**Table 5.2. Comparisons of methods 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, 2008)**

<table>
<thead>
<tr>
<th>Step</th>
<th>NOAEL-derived ADI</th>
<th>BMD-derived ADI</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Data selection</td>
<td>Sufficient sample sizes, at least one dose with “no” effect and one dose with effect. Relevant end-points in a relevant species are important for any approach.</td>
<td>Sufficient number of doses with different response levels and a sufficient number of total subjects.</td>
</tr>
<tr>
<td>2. Model selection</td>
<td>Statistical method</td>
<td>Fit dose–response model (e.g. Weibull model).</td>
</tr>
<tr>
<td></td>
<td>[ R(D) = \begin{cases} 1 &amp; \text{if response at dose } D \ 0 &amp; \text{if response at dose } D \ 0 &amp; \text{not significantly different from control response} \ 1 &amp; \text{is significantly different from control response} \end{cases} ]</td>
<td>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).</td>
</tr>
<tr>
<td>3. Statistical linkage</td>
<td>Pairwise statistical tests between dose groups and control group.</td>
<td></td>
</tr>
<tr>
<td>4. Parameter estimation</td>
<td>Assessment of point of departure [ NOAEL = D_{NOAEL} ]</td>
<td>Choose an appropriate response, p, in the range of experimental response.</td>
</tr>
<tr>
<td></td>
<td>where ( R(D) = 0 ) for all ( D \leq D_{NOAEL} ) and ( R(D) = 1 ) for all ( D &gt; D_{NOAEL} )</td>
<td>Estimate BMDL(_p), the 95% lower confidence bound on the BMD(_p), where ( \frac{R(BMD_p) - R(0)}{1 - R(0)} = p )</td>
</tr>
<tr>
<td></td>
<td>This procedure presupposes that all doses below the NOAEL are non-significant and all doses above the LOAEL are significant. This is often not the case.</td>
<td></td>
</tr>
<tr>
<td>5. Implementation</td>
<td>[ ADI = \frac{NOAEL}{UFS} ]</td>
<td>[ ADI = \frac{BMD_p L}{UFS} ]</td>
</tr>
<tr>
<td>where UF is uncertainty factor.</td>
<td>Sensitivity of BMD to model choice can be checked by fitting various models.</td>
<td></td>
</tr>
</tbody>
</table>
The table shows calculation of an ADI, but the methods are applicable to any health-based guidance value.

5.3.2 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 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 as “The highest dose of a substance at which no toxic effects are observed” (IPCS, 1990). In reality, both terms have similar meaning, and the NOAEL/NOEL has been used similarly to set health-based guidance values by both Committees. At its sixty-eighth meeting (WHO, 2007), JECFA decided to differentiate between the terms NOEL and NOAEL in future risk assessments with the following definitions:

- **No-observed-effect-level (NOEL):** Greatest concentration or amount of a substance, found by experiment or observation, that causes no alteration of morphology, functional capacity, growth, development, or lifespan of the target organism distinguishable from those observed in normal (control) organisms of the same species and strain under the same defined conditions of exposure.

- **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 based on demonstrating the absence of an effect, and the result 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 considered that this level of intake is without 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 where 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 can only be one of the doses actually applied in the study. If the true threshold were 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 were 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, e.g. ADI, TDI) from NOAEL-based DRM (step 5 above) is given by the equation:

\[
HBGV = \frac{NOAEL}{UFs}
\]

Uncertainty factors (UFs) 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 an animal study into a health-based guidance value (IPCS, 1987). Additional uncertainty factors may be used to allow for database deficiencies such as the absence of a chronic study or when effects are detected at all experimental dose levels and a NOAEL therefore cannot be defined. In such cases, a LOAEL may 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 interspecies differences between the experimental animal species used and humans and the variability in responses in humans (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 as well as from a study in animals.

The concept of CSAFs (IPCS, 1994, 2005) has been introduced to allow appropriate data on species differences and/or human variability in either toxicokinetics (fate of the chemical in the body) or toxicodynamics (actions of the chemical on the body) to modify the relevant default 10-fold uncertainty factor. The strategy used by IPCS involves replacing the original 100-fold uncertainty factor with CSAFs (IPCS, 1994, 2005). For example, JECFA used comparative body burden data rather than external dose data in its calculation of a provisional tolerable monthly intake (PTMI) for dioxin-like substances, allowing the usual 100-fold uncertainty factor to be subdivided and replaced with lower values, since there was no need for an uncertainty factor for toxicodynamics differences between species (WHO, 2002).

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.

### 5.3.3 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 preselection of a level of response, the BMR. In general, the level chosen is such that it is close to the LOD 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 5.2. 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 USEPA BMD software program, a number of routinely used models are cited (http://cfpub2.epa.gov/ncea/cfm/recordisplay.cfm?deid=20167). Caution should be used in interpreting the results of DRM if widely varying BMDL estimates are given when multiple models are applied to the same data, because this could indicate inadequate data for modelling.

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 (WHO, 2006). Annex 3 of the report of that meeting (WHO, 2006) 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% response, that is close to the LOD 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 (2008).

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

$$HBGV = \frac{BMDL}{UF_s}$$

In this calculation, the values of the uncertainty factors could be the same as used for the NOAEL or adjusted to account for a slightly different interpretation for 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 by 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, since 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 considered
and 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”.

5.3.4 Uncertainty
A disadvantage with the NOAEL as a POD (“starting point”) for formulating advice to risk
managers is that it is not possible to quantify the degree of variability and uncertainty that
may be present. The NOAEL is assumed to be a dose without biologically significant effects.
This assumption is more likely to be valid in toxicological studies with larger sample sizes.
While uncertainty factors are amenable to uncertainty analysis (Slob & Pieters, 1998), the
NOAEL approach is not readily amenable to quantitative estimation of uncertainty or to a
sensitivity analysis.

A modelling approach facilitates both sensitivity and uncertainty analyses. Uncertainty
can be expressed numerically when the doses and responses are linked by a model (see Table
5.2). Such numerical analyses can also be subject to sensitivity analyses to test the
contribution of different aspects of the database or of model characteristics to the overall
uncertainty. The uncertainty that arises from aspects of study design, such as dose spacing,
sample size and biological variability, on the risk estimates can be assessed using DRM.

5.3.5 Study design
A design optimal for the NOAEL approach could limit the use of DRM, and vice versa. While
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 on 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.

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, although
unlike the NOAEL approach, the inherent uncertainty is more explicit.
5.3.6 Biological information

The NOAEL approach incorporates biological information through the application of subjective “expert” judgement. More complicated forms of DRM have the potential for a more “science-rich” analysis through the more formal quantitative inclusion of factors/covariates into the models, in the case of both human epidemiological and animal data. Such an approach can lead to more certain estimates centred on a toxicologically based concept of estimating the dose–response relationship on the basis of all available biological knowledge using empirical data and applying statistical inference.

5.4 Principles of dose–response modelling

5.4.1 Data

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 obviously larger NOAELs compared with the other studies. In addition, end-points not showing a clear dose–response on visual inspection 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 NOAEL approach, there should be an adverse effect detected at one or more dose levels and no response detected at one or more dose levels. 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.

There are various types of response data (see section 5.2.2), and the main distinction is between quantal and continuous data. Although the type of data is important for statistical reasons, the distinction between quantal and continuous data also has a crucial impact on interpretation of results and their use in risk assessment. In the case of quantal dose–response data, information on the change of incidence with dose is available at one particular degree of effect. For example, the incidence of inflammation as judged from histopathology may increase as dose increases, but under the quantal categories “no inflammation” or “inflammation”, there is no indication of the degree of effect. In ordinal and continuous data, in contrast, information on both the degree of effect and the incidence is available as a function of dose. So, for example, inflammation might be categorized into an ordinal variable using the categories “none”, “mild”, “moderate” and “severe”, or might be quantified in a continuous variable, for example, from measurements on biochemical markers of inflammation.

5.4.2 Mathematical models

A number of mathematical models have been or can be used to describe dose–response data. These are outlined below, but their application and interpretation require specialized expertise. Further details of the main models are provided in the report of the sixty-fourth meeting of JECFA (WHO, 2006) and in IPCS (2008).

5.4.2.1 Dose–response models for continuous data

The models listed in Table 5.3 are some of the forms that may be used to describe the relationship between dose and the magnitude of 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, like 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, 2008).

### Table 5.3. Dose–response models for continuous data

<table>
<thead>
<tr>
<th>Name(s)</th>
<th>Notes</th>
<th>Equation for response</th>
<th>Parameter explanations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Michaelis-Menten law of mass action</td>
<td>A theoretical account of enzyme- or receptor-based activity where the rate of action is a function of the rate of association (ka) and the rate of dissociation (kd).</td>
<td>$R_{\text{Max}} \frac{[S]}{K_M + [S]}$</td>
<td>$R_{\text{Max}}$ is the maximum rate of the reaction, $[S]$ is the substrate concentration, and $K_M$ is the Michaelis-Menten constant, which is equal to $k_a/k_d$.</td>
</tr>
<tr>
<td>Hill equation log-logistic</td>
<td>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.</td>
<td>$R_{\text{Max}} \frac{D^n}{K_D^n + D^n}$</td>
<td>$R_{\text{Max}}$ 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.</td>
</tr>
<tr>
<td>First-order exponential</td>
<td>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.</td>
<td>$R_{\text{Max}} \left(1 - e^{-rD}\right)$</td>
<td>$R_{\text{Max}}$ is the maximum response, $D$ is the dose, and $r$ is the exponential rate constant.</td>
</tr>
<tr>
<td>Power</td>
<td>Simple exponential model.</td>
<td>$= \beta D^\alpha$</td>
<td>$D$ is the dose, $\alpha$ is the shape parameter, and $\beta$ is the scale parameter.</td>
</tr>
<tr>
<td>Linear</td>
<td>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.</td>
<td>$= mD$</td>
<td>$D$ is the dose and $m$ is the slope.</td>
</tr>
</tbody>
</table>

5.4.2.2 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 5.4). For a group of homogenous 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.

<table>
<thead>
<tr>
<th>Name(s)</th>
<th>Theoretical basis</th>
<th>Equation for frequency (F)</th>
<th>Parameter explanations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Step function</td>
<td>No variability.</td>
<td>If $D &lt; T$, $F = 0$ If $D \geq T$, $F = 1$</td>
<td>$D$ is the dose and $T$ is the threshold parameter.</td>
</tr>
<tr>
<td>One-hit (single-hit)</td>
<td>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).</td>
<td>$1 - e^{-(\alpha + \beta D)}$</td>
<td>$D$ is the dose, $e$ is Euler’s constant, $\alpha$ is a location parameter, and $\beta$ is the slope parameter.</td>
</tr>
<tr>
<td>Gamma multi-hit</td>
<td>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.</td>
<td>$\Gamma(\text{gamma} \times D, k)$</td>
<td>$\Gamma()$ 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.</td>
</tr>
<tr>
<td>Probit normal</td>
<td>A descriptive model based on a normal or Gaussian distribution.</td>
<td>$\Phi(\alpha + D \times \beta)$</td>
<td>$\Phi()$ is the normal CDF, $D$ is the dose, $\alpha$ is a location parameter, and $\beta$ is the slope parameter.</td>
</tr>
<tr>
<td>Logistic</td>
<td>The statistical logistic model is also a descriptive tool with no theoretical basis.</td>
<td>$\frac{1}{1 + e^{-D \times \beta}}$</td>
<td>$D$ is the dose, $\alpha$ is a location parameter, and $\beta$ is the slope parameter.</td>
</tr>
<tr>
<td>Weibull</td>
<td>A flexible descriptive model originally developed to describe survival data in demography.</td>
<td>$e^{-(\alpha + (\beta \times D)^\gamma)}$</td>
<td>$D$ is the dose, $\alpha$ is the background parameter, $\beta$ is the slope parameter, and $\gamma$ is an exponent.</td>
</tr>
</tbody>
</table>

Background response rates should be accounted for by incorporating an additional parameter in the dose–response model (see IPCS, 2008).

5.4.2.3 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 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.

5.4.2.4 Biologically based dose–response models

While 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. Another class of model, the biologically based dose–response models, 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 both in resources and time to construct and thus would be expected to be developed fully only for exposures and toxicities of the highest concern.

5.4.3 Model fitting and estimation of parameters

Details of mathematical methods used to optimize the fitting of the model to the data and methods for comparisons of the outputs of different models are given in the report of the sixty-fourth meeting of JECFA (WHO, 2006) and in IPCS (2008). IPCS (2008) also provides details on incorporation of sampling errors, study errors and model errors into uncertainty analysis, together with detailed information on selection of the BMR and interpretation of the BMD and BMDL.

5.5 References


