From THE INSTITUTE OF ENVIRONMENTAL MEDICINE
Karolinska Institutet, Stockholm, Sweden

Quantitative Approaches for Health Risk Assessment of Environmental Pollutants

Estimation of Differences in Sensitivity, Relative Potencies, and Margins of Exposure

Fereshteh Kalantari

Stockholm 2012

Published by Karolinska Institutet. Printed by University Service, US-AB
Nanna Svartz väg 4 – SE 171 77 Solna

ISBN 978-91-7457-868-3

To my family

ABSTRACT

Historically, quantitative health risk assessment of chemical substances is based on deterministic approaches. For a more realistic and informative health risk assessment, however, the variability and uncertainty inherent in measurements should be taken into consideration. The variability and uncertainty may be assessed by applying probabilistic methods when performing dose-response assessment, exposure assessment and risk characterization.

The benchmark dose (BMD) method has been suggested as an alternative to the no observed (adverse) effect level (NO(A)EL) approach in dose-response assessment of non-cancer health effects. In contrast to the NO(A)EL that is limited to being one of the experimental dose levels, the BMD is estimated as the dose corresponding to a predetermined change in response, according to a model fitted to the dose-response data.

In the present thesis, quantitative differences in sensitivity between dioxin sensitive Long-Evans (L-E) and dioxin resistant Han/Wistar (H/W) rats following long-term exposure to 2,3,7,8-tetrachlorodibenzo-*p*-dioxin (TCDD) was investigated. Sensitivity differences were analyzed by comparing BMDs for the two strains, considering a number of conventional toxicological endpoints, endpoints relevant for the endocrine system, and a group of bone parameters. Differences between the strains were most pronounced for hepatic foci; L-E rats were approximately 20-40 times more sensitive than H/W rats. For body and organ weight parameters, L-E rats were 10-20 times more sensitive than H/W rats. For retinoid parameters and hepatic CYP1A1 induction, estimated differences between the strains were generally about 5-fold. For bone effects, significant strain differences were observed with the L-E rat being the most sensitive strain. This difference was most pronounced (about 49-fold) for cross-sectional area of proximal tibia. It was also concluded that the BMD approach is a more suitable method for evaluation of bone parameters compare to the NOAEL approach. In another application, relative potency values (REPs) were established for a group of dioxin-like (DL) and non-dioxin-like (NDL) polychlorinated biphenyl (PCB) congeners as the ratio between BMDs, median effective doses (ED_{50}s), or NOELs. This analysis was based on increased liver weight, decreased hepatic vitamin A levels, and increased hepatic EROD activity. The findings indicated that the BMD approach results in more reliable REP values compared to the ED_{50} and NOEL approaches. The BMD approach also provides more information about the precision of the estimated REP values by the calculation of a two-sided 90% confidence interval; a confidence interval may also be established for a ED_{50} ratio but not for a NO(A)EL ratio. Overall findings in this analysis supported further development and use of endpoint specific systems for assessment of human exposure to mixtures of chemicals with similar as well as different mode-of-actions.

Finally, the potential health impact of a group of PCBs was characterized by estimating the cumulative margins of exposure (MOEs) for the adult Swedish population. A cumulative MOE distribution was quantified by simultaneous integration of a reference dose (RfD) distribution and a distribution for the human dietary exposure. Both a relative potency factor (RPF) based approach and an RPF-free approach were used for estimating the cumulative MOE. Results indicated that the cumulative MOE could be up to four times lower for women compared to men. The cumulative MOE reflected the MOE for PCB 126; other PCB congeners had little contribution. Compared to conventional MOE approaches, the newer approaches considered herein provide an improved tool under which potential health concerns can be assessed by accounting for both variability and various uncertainties, quantitatively, contributing to improving cumulative health risk assessments for the human population.

LIST OF PUBLICATIONS

I. Sand S, Fletcher N, von Rosen D, **Kalantari F**, Viluksela M, Tuomisto JT, Tuomisto J, Falk-Filipsson A, Hakansson H. 2010. Quantitative and statistical analysis of differences in sensitivity between Long-Evans and Han/Wistar rats following long-term exposure to 2,3,7,8-tetrachlorodibenzo-p-dioxin. Regul Toxicol Pharmacol 57(2-3): 136-145.

II. Herlin M, **Kalantari F**, Stern N, Sand S, Larsson S, Viluksela M, Tuomisto JT, Tuomisto J, Tuukkanen J, Jamsa T, Lind PM, Hakansson H. 2010. Quantitative characterization of changes in bone geometry, mineral density and biomechanical properties in two rat strains with different Ah-receptor structures after long-term exposure to 2,3,7,8-tetrachlorodibenzo-p-dioxin. Toxicology 273(1-3): 1-11.

III. **Kalantari F**, Westerholm E, Fattore E, Öberg M, Sand S, Håkansson H. 2012. Estimation of relative potency values for polychlorinated biphenyl (PCB) congeners based on hepatic endpoints of toxicity. ***Manuscript.***

IV. **Kalantari F**, Bergkvist C, Berglund M, Fattore E, Glynn A, Håkansson H, Sand S. 2012. Establishment of the cumulative margin of exposure for a group of polychlorinated biphenyl (PCB) congeners. ***Submitted.***

CONTENTS

LIST OF ABBREVIATIONS

ADI	Acceptable daily intake
AhR	Aryl hydrocarbon receptor
BMD	Benchmark dose
BMR	Benchmark response
CYP	Cytochrome P450
DL	Dioxin like European Food Safety Authority
GM	Geometric mean
GSD	Geometric standard deviation
EFSA	European Food Safety Authority
EROD	Ethoxyresorufin-*O*-Deethylase
H/W	Han/Wistar
L-E	Long-Evans
LO(A)EL	Lowest-observed-(adverse)-effect-level
MOE	Margin of exposure
NO(A)EL	No-observed-(adverse)-effect-level
NFA	National Food Agency
NDL	Non-dioxin-like
PCB	Polychlorinated biphenyl
PoD	Point of departure
RfD	Reference dose
REP	Relative potency
RPF	Relative potency factor
TCDD	2,3,7,8-tetrachloro-dibenzo-*p*-dioxin
TDI	Tolerable daily intake
TEF	Toxic equivalency factor
TEQ	Toxic equivalence (TCDD toxic equivalence)
U.S. EPA	United States Environmental Protection Agency
WHO	World Health Organization

1 BACKGROUND

Humans are continuously exposed to a variety of compounds present in the environment and in food. Human health risk assessments are performed to assess the risk of adverse health effects of exposure to chemical substances. According to the World Health Organization (WHO) this process consists of four different steps; hazard identification, hazard characterization, exposure assessment and risk characterization (WHO/IPCS 2004).

The first step is to identify the type and nature of the adverse health effects that an agent has the capability to cause in an organism, system or population. The aim of hazard identification is to identify the potential critical endpoints that can be assumed to be relevant for humans. The second step, hazard characterization (dose-response assessment) involves investigating the relationship between the administered or absorbed dose of a chemical and the adverse effects it may produce in more detail. This assessment mostly relies on data from experimental animal studies, while consideration of data from human epidemiological studies is made from time to time. The overall aim of the hazard characterization is to determine the dose below which the compound of interest does not cause adverse health effects in humans. This is the estimated maximum exposure level of an agent, normally expressed on a body mass basis, to which individuals in a (sub) population may be exposed over a lifetime without appreciable health risk. Various terms are used to for such health-based guidance values; for example acceptable daily intake (ADI), tolerable daily intake (TDI) (the term "tolerable" is used for agents that are not deliberately added, such as contaminants in food), reference dose (RfD), and reference concentration (RfC). ADI/TDI are terms used e.g. by the WHO and the European Food Safety Authority (EFSA), while RfD/RfC are used by the U.S. Environmental Protection Agency (EPA). In this thesis the term RfD will be used. The RfD is estimated by applying uncertainty factors, such as inter- and intra-species factors, to a point of departure (PoD) derived from the dose-response assessment.

Exposure assessment involves estimating the exposure, via one or multiple routs, to a particular agent in a population of interest. The average exposure alone is a poor descriptor of the exposure in a population. Thus, assessment of the potential variation in

exposures is an important part of this step. More recently, there is also a focus towards assessing the uncertainty, besides the variability.

Risk characterization involves the integration of the first three steps in the risk assessment process. Ideally, it should produce a quantitative estimate of the risk in the exposed population, and/or estimates of the potential risk under different plausible exposure scenarios. Such estimates may be derived by combing the exposure information with information on the relationship between dose/exposure and the adverse health effect of interest. The potential health concern for a particular agent is often assessed by estimating the margin of exposure (MOE) or margin of safety (MOS), which involves comparison of the exposure situation with PoDs or reference points (RPs, i.e. equivalent to PoDs) derived in the dose-response assessment (Renwick et al. 2003; WHO 1999).

The current thesis mainly relates to hazard characterization, but it also deals with exposure assessment and risk characterization to some extent. Even though human data sometimes is used for hazard characterization the case of using animal data as a basis is exclusively considered. The main objective within this work is the further development and application of new methods for quantitative health risk assessment of chemicals.

1.1 HAZARD CHARACTERIZATION

1.1.1 Threshold vs. non-threshold assumption

Traditionally, different assumptions have been made regarding the shape of the dose-response relationship in the low dose region depending on whether a non-genotoxic or a genotoxic effect is considered. For non-genotoxic effects an exposure threshold is assumed; i.e. there exists a dose level below which no biologically significant effects are observed. Since only differences in response that are higher than the experiment detection limit can be identified, it is hard to prove or disprove the presence of a threshold from the experimental data (Slob 1999). Therefore the existence of a threshold should be supported by expert judgment of the underlying biology of the effect. For genotoxic carcinogens, however, it has theoretically been regarded that even a single molecule of genotoxicants could damage DNA, leading to the development of a tumor. Therefore, a non-threshold assumption is considered for genotoxic carcinogens.

The present thesis relates to risk assessment based on non-gentoxic effects, such as organ weight effects, alteration of tissue retinoid levels, changes in bone parameters and it does not specifically discuss the case of genotoxic carcinogens.

1.1.2 The NO(A)EL approach

A purpose of the dose-response assessment is to specify the highest dose without appreciable risk (effect). Traditionally, the no-observed-adverse-effect-level (NOAEL) is used as the PoD for the establishment of such health-based guidance values (e.g. RfDs) when the health risk assessment is based on non-genotoxic effects. The NOAEL is usually derived from animal data, and is defined as the highest experimental dose level for which the (mean) response is not significantly different compared with the (mean) response in the control group (threshold assumed). The effects that are considered for derivation of a NOAEL should also be toxicologically relevant effects, meaning that they should be relevant to human health risk. Since, it is very difficult if not impossible to judge the health relevance of an induced effect, some has supported the use of the term no-observed-effect-level (NOEL) (Berry 1988). If a NO(A)EL cannot be derived, the lowest-observed-(adverse)-effect-level (LO(A)EL) is used instead and is extrapolated to NO(A)EL by the application of an extrapolation factor. The final PoD is normally determined as the lowest relevant NO(A)EL, among the potentially critical effects, in the critical study/s. For establishment of health based-guidance values for human, uncertainty factors are applied to the PoD.

In this thesis the term NOAEL/LOAEL will be used when a studied endpoint can be considered as an adverse toxicological effect and the term NOEL/LOEL will be used when the investigated endpoint is not considered to be toxic *per se* but rather a marker of toxicity.

1.1.3 The Benchmark dose approach

Several limitations have been associated with the NO(A)EL approach (Crump 1984; Davis et al. 2011; USEPA 1995). Due to these shortcomings, the benchmark dose (BMD) approach was introduced by Crump (1984) as an alternative method. The BMD approach involves fitting of a mathematical model to the dose-response data and estimating the BMD as the dose corresponding to a predetermined change in the

response, referred to as the benchmark response (BMR). The BMD approach is illustrated in Figure 1. This concept was initially discussed for quantal dose-response data with a focus towards developmental toxicity (Allen et al. 1994a, 1994b; Faustman et al. 1994; Kavlock et al. 1995). However, this approach is now also frequently applied to (experimental and epidemiological) continuous data (Budtz-Jorgensen et al. 2001; EFSA 2009b; Sand et al. 2004; Slob 2002).

For determination of health-based guidance values using the BMD approach the extrapolation/uncertainty factors are applied to the lower confidence bound of BMD, i.e. the BMDL, instead of the NO(A)EL. The BMDL is usually defined as the one-sided lower 95% confidence limit on the BMD (which equals the lower bound of a two-sided 90% confidence interval); it can be interpreted as the dose corresponding to a response not likely to be larger than the specified BMR (with 95% confidence). Studies with larger sample sizes will produce tighter confidence intervals and therefore result in larger BMDLs (considering all else equal).

The fact that the BMD approach accounts for sample size, by assessing the statistical uncertainty associated with the BMD, has been regarded as a major advantage relative to the NO(A)EL approach. When increasing the sample size of the experiment the power of the test to detect an effect increases; if there are significant differences between groups it is easier to detect this as the sample size increases (i.e. the NO(A)EL decreases). The opposite characteristics have been discussed to be more appropriate from regulatory viewpoint arguing that larger experiments should provide a greater evidence of safety and therefore result in a higher PoD, which is theoretically the case for the BMD approach. Because the BMD concept uses the entire data instead of comparing each dose group against the control, it should also be less sensitive to study design and experimental error compared to the NO(A)EL approach. The shape of the dose-response relationship is also taken into account to a higher extent relative to the case when using a NO(A)EL.

Another advantage of using the BMDL is that the PoD for risk assessment then corresponds to an explicit response level which introduces consistency, but at the same time the specifications of suitable levels of response/risk has also shown to be one of the main challenges with BMD concept. The NO(A)EL may sometimes be interpreted as a dose threshold without adverse health effect. However, as pointed out previously,

based on the observed dose-response data the existence of threshold cannot be proven; an infinite number of measurements, both in terms of doses and effects, would be needed, and such an experiment cannot be conducted from a practical point of view (Edler et al. 2002; Slob 1999, 2007). As will be discussed below (section 1.3.1.5) when expressing response/risk as defined within the BMD framework (in terms of a BMR), a response/risk may also be present at the NO(A)EL in spite of it being a "no-effect level". An extensive discussion and comparison of the BMD and the NOAEL is presented in (EFSA 2009a) and in (Sand et al. 2011).

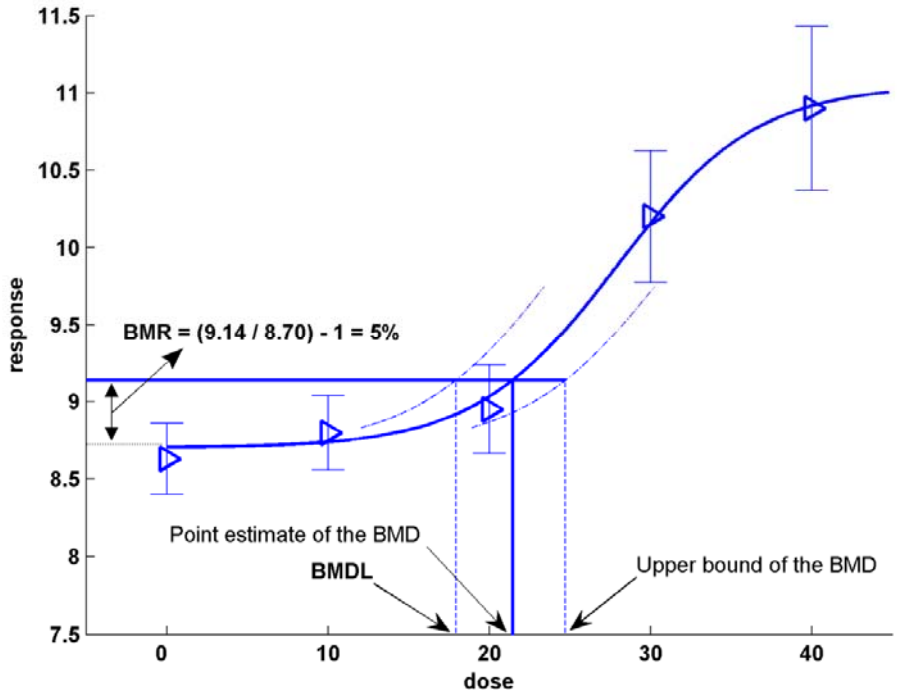

Figure 1. A dose-response model fitted to data. The triangles represent the observed mean responses at each dose level. The BMD which is a dose corresponding to a predefined change in response is determined from the fitted model (solid curve). The dashed lines demonstrate the upper and lower 95% confidence bounds for doses corresponding to specified response levels (BMRs). Their intersections with the horizontal line result in the lower and upper bounds of the BMD, i.e. BMDL and BMDU respectively. In this figure the BMD corresponding to 5% change in response relative to background is illustrated. The estimated background level from the fitted model is 8.7 and therefore 5% increase of that equals 9.14 (8.7+0.05*8.7= 9.14). The BMD of 21.50 is obtained from the intersection of horizontal line at the response level of 9.14 with the fitted curve.

1.1.3.1 Quantal data and dose-response models

Quantal data implies data where experimental animals are classified as responders or non-responders, e.g. the number of animals with (one or several) tumors in a particular dose group. Dose-response data may also be presented in ordinal format. Such data result when several sub-classes describing the severity of response (minimal, mild, moderate, etc) are defined. The simplest case of ordinal data is equivalent to quantal data and, theoretically, as the number of sub-classes in ordinal data approaches infinity, continuous data results. It has been discussed that quantal (or ordinal) data may be "connected" to an "underlying continuous response", i.e. when subjects are categorized (e.g. by the experimenter) as responders they have exceeded a certain level of the "underlying continuous response" (Sand et al. 2008; Slob and Pieters 1998). This type of reasoning may however not be applicable to any type of quantal endpoint.

A number of quantal dose-response models have been applied in health risk assessment of chemicals (Bailer et al. 2005; Sand et al. 2002). Some of these models are standard probability distribution function, such as the (log) logistic, the (log) probit, and the Weibull models, and some of them, like the gamma and the multistage models, are stochastic models that are based on the assumption that a positive response in an animal is the result of random occurrence of one or more biological event (Krewski and Van Ryzin 1981). The latter models were previously suggested for assessment of genotoxic carcinogens. To date, however, this distinction is not made; models that adequately describe the data are further considered in the risk assessment process regardless of whether a non-genotoxic or a genotoxic effect is considered. The minimum and maximum responses of quantal models are restricted to be in the range of 0 and 100%. For quantal data of general type it is assumed that the effect observed for an individual is independent of the effect observed for another individual. However, this is not the case for developmental toxicity data where fetuses from the same litter may respond more similar than fetuses from different litters. For such data intra-litter correlation should be regarded. Therefore, specific models and methods have been developed for evaluation of developmental toxicity data (Kodell et al. 1991; Kupper et al. 1986; Rai and Van Ryzin 1985).

1.1.3.2 Continuous data and dose-response models

For continuous data, the degree of response is observed in the individual subject. Enzyme activities, organ weights, and hormone levels are examples of continuous responses. A wide range of dose-response models may be used to describe the relationship between the dose of chemical and the mean response of a continuous endpoint, for example, the polynomial models (including the linear model), the power model (Allen et al. 1994a; Allen et al. 1996; Crump 2002; Crump 1984; Crump 1995; EFSA 2009a; Kavlock et al. 1995), a family of nested exponential models (Slob 2002), and the Hill model family (Barton et al. 1998; Davis et al. 2011; EFSA 2009a; Gaylor and Aylward 2004; Kim et al. 2002; Murrell et al. 1998; Sand et al. 2004; Toyoshiba et al. 2004; Zhou et al. 2001). The Hill and exponential families contain a parameter that describes the dynamic range of response (the absolute or relative difference between the maximum and minimum response levels) that allows the model to plateau at high doses. Therefore, these models can produce S-shaped dose-response relationships.

Continuous response data is typically assumed to be normally or log-normally distributed. For simplicity, the variance is generally assumed to be constant among dose groups. However, this assumption is not always appropriate. For some datasets, the variance increases/decreases as the mean response increases/decreases, In such situations the variance can be modeled by a statistical function such as power function or exponential model (Davis et al. 2011; Sand et al. 2006). In the present thesis, the constant model (**Papers II** and **III**) and dose-dependent models for variance were considered (**Papers I** and **II**).

1.1.3.3 Data requirement for BMD modeling

There are some minimal data requirements for BMD modeling. For quantal data, the sample size and the incidence of response is needed for each dose group. For continuous data, the observed response data for each experimental animal or the summary data (the mean response, the standard deviation, and the number of animals at each dose level) are needed. It should be noted that outliers cannot be recognized when continuous summary data is used.

In the process of dose-response modeling and BMD analysis it is investigated whether or not there is a statistically significant dose-response trend; a BMD is generally not

recommended to be derived if no such relationship can be demonstrated. It has been debated if a minimal number of dose groups are required for BMD analysis. According to U.S. EPA, a dataset should at least contain three dose groups in addition to the control group (Davis et al. 2011). No specific recommendation is given by EFSA. Datasets with few dose groups may, depending on dose placement, provide limited information on the shape of the dose-response curve. Resulting BMDs and BMDLs may then differ substantially depending on the dose-response model used, which indicates that a PoD derived from such data will be uncertain. Importantly, applying a NO(A)EL approach in such a case is not a better approach. Rather it may be discussed if such data are adequate for use in quantitative risk assessment.

1.1.3.4 Selection of dose-response models and overall BMDL

There are numerous models that may be used to describe the dose-response relationships and subsequently drive the BMD. This results in BMD values that are model dependent i.e. by fitting different models different BMDs are obtained. To take model selection uncertainty into account, it has been suggested that a family of nested models such as Hill family or exponential family is fitted to the same dataset (EFSA 2009a). From each family of models, the model with the appropriate number of parameters is then selected, e.g. based on a likelihood ratio test statistic. Considered models (e.g. the most appropriate member in each model family) also needs to fulfill standard statistical requirements for goodness-of-fit. Visual inspection of the plotted dose-response curve can give additional indication of how well the model fits the data. There may be cases where the goodness of fit test is violated whereas the model appears to fit the data well. There could also be cases where the goodness of fit test supports the model while an unrealistic curve fit is observed.

The accepted models are then further considered. Thus, a range of BMDLs may result which becomes an indicator of the degree of model uncertainty for the particular data set considered (EFSA 2009a). Model averaging approaches have been discussed for deriving a single BMDL value from a given dataset (Wheeler and Bailer 2007, 2009), but at this point the most conservative (lowest) value is used in practice as the BMDL for that particular dataset. The overall PoD is then typically determined, similar as in the NOAEL approach, as the lowest BMDL, among the potentially critical effects/datasets, in the critical study/s (EFSA 2009a).

It should be noted that the proposed models are used descriptively based on some general considerations, and are not regarded as biologically based at any detailed level. The reliability of the BMD is more related to the quality of the data than the selected model. If the data contains sufficient information different models will result in similar BMD values (Slob 2002) ("Quality of data" and "sufficient information" here mostly refers to the number of dose groups applied).

1.1.3.5 Definitions and specifications of the BMD and BMR

Depending on the type of response data (quantal or continuous), the BMR associated with the BMD is defined differently. For quantal data, the BMR is commonly defined in terms of additional or extra risk:

Additional risk: $$BMR = p(BMD) - p(0) \quad . \qquad (1)$$

Extra risk: $$BMR = \frac{p(BMD) - p(0)}{1 - p(0)}, \qquad (2)$$

Where $p(BMD)$ denotes the probability of response at the BMD and $p(0)$ is the background probability of response.

For continuous data, several suggestions have been made for how to define the BMD (and BMR). The BMD for continuous data has for example been defined as the dose corresponding to a percentage change in response relative to background (Crump 1984; Slob and Pieters 1998):

$$BMR = \frac{|\mu(BMD) - \mu(0)|}{\mu(0)} \qquad (3)$$

where $\mu(0)$ denotes the mean background response, and $\mu(BMD)$ denotes the mean response at the BMD. The BMR for continuous data has also been presented in terms of a change in response relative to the standard deviation of the control group, $\sigma(0)$ (Crump 1984; Crump 1995; Kavlock et al. 1995);

$$BMR = \frac{|\mu(BMD) - \mu(0)|}{\sigma(0)} \qquad (4)$$

In addition to the definitions presented above the BMD has alternatively been expressed as the dose corresponding to a change in response relative to the estimated dynamic range of response (the difference of the maximum and minimum response levels) (Murrell et al. 1998):

$$BMR = \frac{|\mu(BMD) - \mu(0)|}{\max(response) - \min(response)} \qquad (5)$$

This definition is only applicable in situations where the model describing the dose-response data levels off to some limiting value, i.e. consist of a parameter that describes the dynamic range of response.

Observe that for quantal data the BMD corresponds to a change in the probability of response, while no probability based interpretation can (directly) be made for the BMD/BMR definitions discussed above for continuous data. However, an indirect probability based method has also been presented for continuous data (Crump 2002; Crump 1995; Gaylor and Slikker 1990), which has mostly been applied to epidemiological data in practice (Budtz-Jorgensen et al. 2000; Clewell et al. 2003; Jacobson et al. 2002). In this approach, a probability model that describes the proportion of the distribution (describing the variability in the population) below (or above) a specific cut-off value as a function of dose is estimated. The BMD is the dose where the probability of exceeding (or falling below) the cut-off level has increased by a particular percentage according to the additional or extra risk definition (equations 1 and 2). BMDs calculated for continuous endpoints according to this approach have probability based interpretation in similarity to BMDs derived from quantal data.

A critical issue concerns the specification of the BMR value in BMD analysis. The BMR should ideally be based on toxicological considerations (e.g. represent a non-adverse effect size). However, it should not be set at levels outside the range of observed responses. Extrapolation far beyond the range of the data can result in BMDs which become dramatically different depending on the model used (EFSA 2009a).

Default values for the BMR for quantal and continuous data have been suggested; however the selection of other BMRs can be justified with biological and statistical considerations.

For quantal data, U.S. EPA and EFSA have suggested a BMR of 10% as default, defined in terms of extra risk (Davis et al. 2011; EFSA 2009a). Some studies have compared NOAEL and BMDL values across several datasets to address the issue of suitable BMR levels. Results from those studies suggest that the risk at NOAEL corresponds to a BMR level in the range of 5-10% or more than 10% (Allen et al. 1994a, 1994b; Fowles et al. 1999). More recently, dose-response analysis of 786 dataset from the U.S. National Toxicology Program Carcinogenesis Bioassay Program indicated that the upper bound on extra risk at the NOAEL was close to 10% at median (Sand et al. 2011). Hence, while the NOAEL does not correspond to a statistically significant change in risk (or effect), a risk at the NOAEL is apparent in terms of BMR and depends on the dataset considered.

In Sand et al (2011) the issue of BMR for quantal data was further addressed by the definition of a signal-to-noise crossover dose (SNCD). The SNCD is the dose at which the additional risk is equal to some fraction of background noise. The SNCD was discussed as an objective estimate of the lowest dose applicable as a PoD for a particular dataset, without having the signal (i.e. the point estimate of additional risk at a certain dose, d) overwhelmed by noise (i.e. the difference between the upper and lower bound of a two-sided 90% confidence interval on absolute risk at the same dose, d).

For continuous data, EFSA recommends a BMR of 5% as default; defined as a 5% change in the mean response relative to the background mean response. EFSA states that a 5% response is often within the range of observation, and would provide BMDL estimates that are not critically dependent on the dose-response model. A re-analysis of a large number of U.S. National Toxicology Program studies showed that the $BMDL_{05}$ was, on average, close to the NOAEL derived for the same dataset, while in most individual datasets they differed within one order of a magnitude (EFSA 2009a).

The data evaluated in the present thesis is continuous, and the BMD was defined as the dose corresponding to 5%, 10% and 100% (depending on the endpoint considered)

change in response according to Equation 3 (**Papers I, II, III** and **IV**) and Equation 5 (**Paper I**).

1.1.3.6 Human dose-response data

The definitions of BMD for quantal and continuous data discussed above are dependent on the response level in unexposed subjects. However, many epidemiological studies may not include an unexposed control group and therefore application of the BMD method may be problematic. In this situation a model can still be fitted to the data since fitting a dose-response model does not necessarily require observations at the control group. The response in the unexposed group can be determined by low-dose extrapolation, but it is important to note that a BMD derived in such a case may become highly model-dependent (Budtz-Jorgensen et al. 2001; EFSA 2009a). Furthermore, epidemiological studies involve different sources of uncertainty that may lead to a biased estimate of the BMD. In order to obtain less biased estimates, factors such as imprecision in human exposure, confounders and effect modifiers need to be handled by suitable statistical methods or by including them in the model as covariates (Budtz-Jorgensen et al. 2004; EFSA 2009b).

1.1.3.7 Comparison of dose-response relationships

An important issue in risk assessment is to assess the relative potency between different chemicals or relative differences in sensitivity between subpopulations (e.g. animal strains, sexes or species). Such a comparison can be performed by introducing a covariate/s in the model. It is then possible to investigate whether the dose-response relationships of two sub-populations are similar or not, i.e. are the sub-populations equally sensitive or not. Comparison of dose-response curves in the context of BMD analysis and risk assessment has generally been discussed by (Slob 2002). In **Papers I** and **II**, a framework was developed for assessing quantitative and statistical differences in sensitivity following chemical exposure, and in **Paper III**, relative potency (REP) values for a group of polychlorinated biphenyls (PCBs) were estimated.

1.1.3.8 Study design

The question whether or not the typical study designs used in animal experiments are adequate for BMD analysis, and/or what the optimal design is for estimating the BMD has been discussed in a few studies. Slob et al. (2005) studied this for continuous data

considering the BMD defined as the dose corresponding to 5% change in response relative to background. It was concluded that the performance of a design is determined by the total number of animals, and increasing the number of dose groups while reducing the number of animals per dose group, does not result in poorer performance of the design. In addition to the sample size, the dose placement showed to be an important factor and to minimize the risk of inadequate dose placement, studies with multiple dose levels are favorable (Slob et al. 2005). Investigations have also been made for continuous data by Kuljus et al. (2006) considering the Hill model where the BMD was defined as a change in response relative to the dynamic range of response (equation 5). The main purpose of this study was to investigate the effect of increasing the number of dose groups and at the same time decreasing the number of animals per dose group. The results suggested that to avoid unfavorable dose placement, studies should be designed with more than four dose groups. To further improve the study design any prior information about the dose-response relationship, e.g. information from similar previous studies should also be taken into account (Kuljus et al. 2006).

1.1.4 Extrapolation factors

1.1.4.1 Interspecies and Intraspecies extrapolation

The PoD derived from experimental or epidemiological studies cannot directly be used to set an RfD for the human population. To establish the RfD, the PoD is commonly divided by one or more extrapolation factors (EFs) depending on the characteristics of the study (e.g. animal study or human study) used for derivation of the PoD.

In the mid-1950s, an assessment factor of 100 was introduced in the United States by the Food and Drug Administration for deriving safe exposure levels for food additives. This approach was later modified slightly and applied by U.S. EPA in setting RfDs for environmental pollutants (Falk-Filipsson et al. 2007). Although it is hard to find a firm scientific basis for a 100-fold assessment factor, it is generally considered to consist of a 10-fold EF for describing interspecies differences and a 10-fold EF for describing intraspecies differences. The default interspecies EF of 10 is intended to account for the difference between the average experimental animal and the average human, and the default intraspecies EF allows extrapolating from the average to the sensitive human.

In an attempt to replace default EFs with data-derived EFs by incorporation of mechanistic information, Renwick (1993) subdivided each of 10-fold EFs into a toxicokinetic factor of 4 and a toxicodynamic factor of 2.5. A more flexible framework that develops a series of EFs related to different processes of metabolism and excretion when chemical-specific toxicokinetic or toxicodynamic data are not available has also been proposed (Dorne et al. 2001, 2004; Renwick et al. 2000; Walton et al. 2004).

Several authors have suggested data-based empirical distributions for inter- and intraspecies factors (Edler et al. 2002; Price et al. 1997; Slob and Pieters 1998; Swartout et al. 1998; van der Voet and Slob 2007; Vermeire et al. 1999). Most of these distributions are based on ratios of NOAEL values from different studies (for the same compounds). As an alternative, Bokkers and Slob (2007) have proposed an empirical data-based interspecies EF distribution based on the BMD ratio of mouse and rat studies. van der Voet and Slob (2007) suggested that the EF interspecies factor can be described by a log-normal distribution with geometric mean (GM) equal to 4 and a geometric standard deviation (GSD) equal to 1.48. Under these settings the 99th percentile of the log-normal distribution is equal to the default interspecies factor of 10. To describe the intraspecies EF distribution van der Voet and Slob (2007) combined both variability and uncertainty. They assumed that the variability among individuals in the population is log-normally distributed with GM of 1 and a GSD which has a chi-square distribution. The log-normal distribution reflects variability while distribution for the GSD reflects the uncertainty of this variability. The degrees of freedom for the chi-square distribution was specified such that the default intraspecies factor of 10 equals the 99th percentile of the resulting intraspecies (log-normal) distribution (van der Voet and Slob 2007). In the present thesis, the inter- and intraspecies EF distributions proposed by van der Voet and Slob (2007) were considered (**Paper IV**).

1.1.4.2 Sub-chronic to chronic extrapolation

The most relevant experimental exposure duration to be used as basis for establishment of the RfD is chronic exposure, since these health-based guidance values should protect the human population from adverse health effects of chemical substances after life-long exposure. However such data may not always be available. By default, a factor of 10 is applied to for sub-chronic to chronic extrapolation. A data-based EF has been suggested, by considering a certain percentile of the sub-chronic to chronic NOAEL ratio distribution (Pieters et al. 1998). Instead of using ratios of NOAELs, Bookers and

Slob (2005), derived a data-based EF distribution by estimating BMD ratios for sub-chronic and chronic studies (body weight and liver weight data on mice and rats). They proposed a lognormal distribution with a GM of 1.7 and GSD of 2.9 to describe the sub-chronic to chronic EF (Bokkers and Slob 2005). In this thesis, the latter sub-chronic to chronic EF distribution was considered (**Paper IV**).

1.1.5 Establishment of health-based guidance value

Traditionally, the RfD is established by dividing a deterministic PoD by various point estimates of EFs. As an alternative, a probabilistic approach may be applied. In this case bootstrap procedures may be used to assess the uncertainty in the BMD (the complete BMD distribution is considered as the PoD). The uncertainty distribution of the BMD is then divided by the distributions for different EFs (as discussed above) to derive an uncertainty distribution for the RfD for the sensitive human (Slob and Pieters 1998; van der Voet and Slob 2007):

$$\boldsymbol{RfD_{sensitive}distrib.} = \frac{\boldsymbol{BMDdistribution}}{\boldsymbol{EF_1distrib.} \times \boldsymbol{EF_2distrib.} \times \ldots \times \boldsymbol{EF_idistrib.}} \quad \textbf{(6)}$$

To date a probabilistic approach for establishing human exposure guidelines has not been applied in risk assessments in practice.

It should be kept in mind that treating the RfD as a definite level, and considering exposures above the guideline as being associated with a health risk can be misleading. More informative indications of both severity and frequency of health effects that may occur in an exposed population should instead be provided (Clewell and Crump 2005).

1.2 EXPOSURE ASSESSMENT

1.2.1 Deterministic and probabilistic methods

Human exposure to chemicals occurs as a result of inhaling air, drinking water, eating food, or through dermal contact with products that contain the chemical. For assessment of dietary exposure, information of consumption habits and chemical occurrence in foods are needed. Food consumption data is obtained from consumption surveys in the population, and the concentration of chemicals in different food items are e.g. derived from food monitoring programs. If the consumption data concern foods as

eaten (i.e. pancakes) while the concentration data is available only for the raw foods (i.e. egg), information on recipes and food processing is ideally needed to match the consumption and concentration data.

Human exposure assessment can be performed using different methodologies. Routinely, a deterministic approach is applied by combining point estimates of food consumption with point estimates of concentrations of chemicals in corresponding food items. This approach regards only a single consumption scenario and assumes a constant concentration of chemicals present in the food. In order to protect the majority of the population worst case scenario approaches may be applied by using conservative estimates of the input parameters, i.e. the 95^{th}/99^{th} percentile or the maximum value may be considered (Kroes et al. 2002). This type of deterministic approach may be applicable as a primary step for evaluating the exposure situation. However, if a risk cannot be excluded under such an assessment a more realistic approach is required as a higher tier to judge whether or not the exposure may be associated with adverse health effects (Bosgra et al. 2009).

In contrast to the deterministic approach, the probabilistic analysis accounts for both variability (due to variation between individuals in food consumption and variability between concentrations of chemicals in the consumed foods) and uncertainty. Hence, this approach results in a more realistic illustration of health risk that can occur in the population. This approach employs all available information and allows us to avoid choosing worst case estimates for input parameters (Bosgra et al. 2005; Kroes et al. 2002; van der Voet and Slob 2007). Variability and uncertainty should be treated separately. Variability between individuals in exposure to chemicals is evaluated by Monte Carlo simulations. Bootstrap simulations can be used to assess the uncertainty; random values are drawn directly from the input data (non-parametric bootstrapping) or from probability distributions describing the input data (parametric bootstrapping). Uncertainty distributions with respect to different percentiles of the exposure can be obtained since the exposure is evaluated for several bootstrap samples. While probabilistic modeling confers many advantages, in order to create reliable input distributions sufficient data on input parameters is needed (Kroes et al. 2002).

1.2.2 Cumulative exposure

Health risks related to chemical exposures are mostly assessed separately for each compound whereas in reality humans are exposed to a complex mixture of chemicals (Kortenkamp 2007). Despite the recognition of this problem for decades, there are only a few models for prediction of cumulative exposure to chemical mixtures. The TEF-system proposed by WHO is the most common method to characterize the toxicity of human exposure to dioxins and dioxin-like (DL) compounds (Van den Berg et al. 2006). For pesticides, the relative potency factor (RPF) approach has been developed (Boobis et al. 2008; Muller et al. 2009; USEPA 2000), and this approach has also been applied for PAHs (Pufulete et al. 2004). Both the TEF and RPF approaches are based on selecting a reference compound and assessing the potency of other relevant compounds in relation to this reference. In the RPF approach, the cumulative exposure of the relevant compounds is then expressed in terms of reference compound equivalents, which are obtained by normalizing the exposure of each substance with its RPF. The RPF and TEF approaches are based on the assumption that there are no interactions between the considered chemicals; if the chemicals interact so that effects are modulated, information for the individual chemicals is not sufficient for predicting their combined effect.

TEFs are determined based on a scientific judgment of multiple REP values from different studies for multiple endpoints. A set of criteria was developed to determine the REP values that can be included in the WHO TEF system. In order to assign a TEF value to a compound it should have a chemical structure that is similar to that of polychlorinated dibenzo-p-dioxins (PCDD) or polychlorinated dibenzofurans (PCDFs), bind to aryl hydrocarbon receptor (AhR), and show AhR mediated biochemical and toxic responses. The compound should also be persistent and accumulate in the food chain (Ahlborg et al. 1994; Van den Berg et al. 2006). The RPF method can be used when a class of compounds appears to share a common mode of action; however, the exact mechanism is complex and may not be known in detail (USEPA 2000).

Several groups of compounds cannot be included in the TEF system as they do not fulfill the criteria of binding to AhR, although they have similar effect as the DL compounds on an endpoint basis. Therefore, when estimating the combined effect by only focusing on TEF assigned compounds, which have similar mechanism of action,

and not including the potency of other compounds with similar effects, the combined effect of mixtures might be underestimated. Hence, developing endpoint-specific RPF databases for endpoints that are observed after exposure to compounds with and without assigned TEFs seems to be appropriate (Ahlborg et al. 1994). When applying the RPF approach for estimating the cumulative exposure, one should bear in mind that the RPFs can be applied for the endpoints for which they were obtained and not for all endpoint (Muller et al. 2009).

1.3 RISK CHARACTERIZATION

The purpose of risk management is to protect the population from the adverse effect of chemicals. Advice to risk manager about the nature and magnitude of risk from exposure to chemicals can be given in a quantitative or qualitative form. Integrating exposure assessment and hazard characterization into risk characterization is an important step in risk assessment (Renwick et al. 2003).

Normally, a deterministic approach is used at the level of risk characterization. A margin of exposure (MOE) or margin of safety (MOS) may be established by comparing a PoD like the NOAEL or BMDL with the mean/median or some percentile of the estimated human exposure. The MOE and MOS may be considered equivalent terms (IPCS 2004; sometimes the MOS is used when the health-based guidance value is applied as reference instead of the NOAEL or BMDL). For genotoxic carcinogens quantitative estimates of cancer risk at given exposure (using linear extrapolation models) have traditionally been performed by the WHO and the U.S. EPA. EFSA proposed a MOE approach for risk assessment of compounds that are both genotoxic and carcinogenic; the MOE is here defined as the ratio between the BMDL, obtained from animal dose-response data for the critical effect, and a point estimate of the human exposure (EFSA 2005).

When appropriate data are available risk characterization may be carried out by probabilistic modeling. Such approaches have so far mostly been discussed at the level of the exposure assessment domain. The probabilistic approaches have not yet been used to aid the risk characterization or at the level of the hazard characterization domain in practice, e.g. for establishing health-based guidance values. A probabilistic MOE approach has been presented by van der Voet and Slob (2007). This is an

integration of the probabilistic approaches already described in this thesis at the level of hazard characterization and exposure assessment; it account for both variability and uncertainty in the estimation of the MOE. In **Paper IV** of current thesis, the model proposed by van der Voet and Slob (2007) was extended and refined to estimate the cumulative MOE for a group of PCBs using vitamin A as an example endpoint. Selected percentiles of the MOE distribution can be used to characterize the proportion of the population that can be considered to be at risk (Bosgra et al. 2009; van der Voet and Slob 2007; van der Voet et al. 2009).

2 PRESENT INVESTIGATION

2.1 AIM

The general objective of this thesis relates to development of methods applied for quantitative health risk assessment of chemical substances. The focus can be divided in two parts:

In **part one** the objective was to further develop and promote approaches for comparison of dose-response relationships for estimating 1) differences in sensitivity between sub-populations, and 2) differences in potencies between chemical substances. The specific project objectives were:

- To estimate the relative difference in sensitivity between L-E and H/W rat strains (with different AhR structure) considering a number of conventional toxicological endpoints as well as endocrine system relevant changes in tissue retinoid levels (**Paper I**).
- To estimate the relative difference in sensitivity between L-E and H/W rat strains for changes in bone geometry, mineral density and biomechanical properties (**Paper II**).
- To estimate REP values for a group of DL and non-dioxin-like (NDL) PCB congeners with the application of traditional approaches, i.e. using the ratio between NO(A)ELs/LO(A)ELs or median effective doses (ED_{50}s), and a new approach, i.e. using the ratio between BMDs (**Paper III**).

In **part two** the objective was to evaluate and further develop a generalized MOE approach for risk characterization of chemical mixtures by integrating dose-response modeling based on experimental effect data and exposure modeling based on human consumption data. The specific project objective was:

- To implement and evaluate an approach for estimating the cumulative MOE, with and without incorporation of RPFs, considering a group of DL and NDL PCBs (**Paper IV**).

2.2 MATERIALS AND METHODS

2.2.1 Data

2.2.1.1 Dose-response animal data

Some of the animal dose-response data used in this thesis were generated in our lab and some were obtained from scientific collaboration (**Papers I** and **II**) or derived from the literature (**Papers III** and **IV**). The dose-response data used was continuous in nature, describing the degree of severity of effect at the level of individual animals.

For **Papers I** and **II**, dose-response data on body and organ weights, altered hepatic foci, hepatic ethoxyresorufin-*O*-deethylase (EROD) activity, as well as retinoid and bone parameters, observed in female Long-Evans (L-E) and Hans/Wistar (H/W) rats following long-term exposure to 2,3,7,8-tetrachlorodibenzo-*p*-dioxin (TCDD) was used.

In **Papers III** and **IV** dose-response data on post-mortem observations, including gross changes such as fatty liver and spleen enlargement, organ weight, biochemical changes, tissue vitamin A levels, hematological changes such as anemia, etc, was derived from previously published studies (Chu et al. 2000; Chu et al. 1998; Chu et al. 1995; Chu et al. 1994; Chu et al. 1996a; Chu et al. 1996b; Lecavalier et al. 1997).

2.2.1.2 Human exposure data

For estimating the human dietary exposure, food consumption data and data on chemical concentrations in food was obtained by collaboration with the Swedish National Food Agency (NFA) (**Paper IV**). Data from the food consumption survey conducted by the NFA in collaboration with the Swedish Statistical Agency in 1997-1998, on adult men and women (Riksmaten 97-98) was used (Becker and Pearson 2002). The concentration data on PCBs in food originates from the NFA dioxin control program from 2003 and onwards, and from NFA food monitoring studies from 1999 and onwards. The concentration data for each PCB congener was extrapolated to the year 2009 with a 3.3 – 8.6 % decrease per year depending on the congener.

2.2.2 Dose-response assessment

2.2.2.1 Assumptions and model fitting

In **Papers I**, **II** and **III** data from two different sub-populations (**Papers I** and **II**) or two different compounds (**Paper III**) were analyzed simultaneously. Data was assumed to be normally distributed. The log-likelihood function can then be written as:

$$\ln L = -\frac{N}{2}\ln(2\pi) - \sum_{i=1}^{2}\sum_{j=1}^{g_i}\left[\frac{n_{ij}}{2}\ln\sigma^2\left(d_{ij}\right) + \frac{\left(n_{ij}-1\right)s_{ij}{}^2}{2\sigma^2\left(d_{ij}\right)} + \frac{n_{ij}\left(\bar{y}_{ij}-\mu(d_{ij})\right)^2}{2\sigma^2(d_{ij})}\right], \qquad (7)$$

where N is the total number of animals; g_i is the number of dose groups for dataset i; n_{ij} is the number of animals in the j'th dose group of dataset i; and $s_{ij}{}^2$ and $\bar{y}_{ij}$ are unbiased sample variance and the sample mean response in the j'th dose group for dataset i, respectively. The parameters defining the mean response $\mu(d_{ij})$ and the variance $\sigma^2\left(d_{ij}\right)$ are estimated by maximizing the log-likelihood.

In **Papers III** and **IV**, data was assumed to be log-normally distributed. However, equation 7 still applies since the response data was log-transformed before analysis. It should be noted, however, that in this case the estimated dose-response models describe the mean of the log-transformed response.

In **Papers I** and **II** equation 7 was used for fitting a model to dose-response data on L-E and H/W rats simultaneously (for a given endpoint), while in **Paper III** it was used for fitting a model to dose-response data on PCB 126 and another PCB congener simultaneously (for a given endpoint). In **Paper III**, the log-likelihood function was also further extended so that a model was fitted to both the male and female dose-response data for PCB 126 and another PCB simultaneously (for a given endpoint). In **Paper IV**, models were fitted to dose-response data on individual PCB congeners separately i.e. using equation 7, without summing over several datasets.

2.2.2.2 One way analysis of variance (ANOVA)

In **Papers II** and **III** in order to determine NO(A)ELs and LO(A)ELs, one-way ANOVA followed by Dunett's *post-hoc* test (**Paper II**) and pairwise t-test (**Paper III**)

was used to compare different dose groups against the corresponding control group at a significance level equal to 0.05.

2.2.2.3 Dose-response models

In all papers, the model used for analyzing the dose-response data was the Hill function, which in its general form can be written as:

$$\mu(d_i) = \alpha + \theta\left(\frac{d_i^{\eta}}{k^{\eta} + d_i^{\eta}}\right), \tag{8}$$

where $\boldsymbol{\alpha}$ is a parameter that describes the background response; $\boldsymbol{k}$ is the location parameter (which equals the ED_{50} dose); $\boldsymbol{\eta}$ is a parameter that describes the shape of the dose response curve; $\boldsymbol{\theta}$ is a parameter that describes the dynamic range of response (the difference between the estimated maximum and minimum response levels); and where d_i is the dose administered to the *i'th* treatment group. If η equals 1, the Hill function coincides with the Michael-Menten equation. These two models are nested and therefore it was tested if the data could just as well be described by the Michael-Menten equation which is a simpler model. The variance was assumed to be constant among dose groups, $\sigma^2 = c$. In Papers **I** and **II**, in addition to the constant model, a dose dependent exponential model, $\sigma^2 = e^{\lambda + \rho \ln(d+1)}$, was also considered for the variance.

2.2.2.4 Likelihood ratio test

The likelihood ratio test was used to assess individual model fit and for selecting the most appropriate model in a family of nested models. It can be shown that twice the difference of the log-likelihoods associated with the two models approximately follows a chi-square distribution with the degrees of freedom equal to the difference in the number of parameters between the two models. In addition, this test is the basis for the likelihood profile method, which was used to estimate the confidence interval for the BMD ratios in **Papers I**, **II** and **III**.

2.2.2.5 RfD distribution

In **Paper IV**, an approach for estimating an RfD under a probabilistic framework was applied. A "new" dose-response dataset was generated using a parametric bootstrap procedure. The BMD resulting from fitting a selected Hill model to the new dataset was divided by an interspecies extrapolation factor (EF_{inter}) that was drawn from log-normal distribution (with GM= 4 and GSD = 1.48) . The resulting value was divided by 20 000 random values drawn from an intraspecies extrapolation distribution (EF_{intra}) corresponding to a log-normal distribution (with GM = 1 and GSD = $\mathbf{1.98} \times \sqrt{5/\chi_5^2}$) where χ_5^2 is a random value draw from a Chi-square distribution (with five degrees of freedom) (van der Voet and Slob 2007). This results in an RfD distribution describing variability among individuals. The RfD distribution was also divided by an additional factor to account for subchronic to chronic extrapolation; this factor was randomly drawn from a lognormal distribution (with GM = 1.7 and GSD = 2.3) (Bokkers and Slob 2005). In order to evaluate the uncertainty in the RfD distribution the process described above was repeated for 500 times.

2.2.3 Exposure assessment

In **Paper IV**, the human dietary exposure to single PCB congeners, as well as the cumulative exposure for a whole group of PCBs was estimated.

A non-parametric bootstrap sample was drawn with replacement from the consumption data, with respect to individuals in the consumption survey (Riksmaten 97-98). Similarly, a non-parametric bootstrap sample was drawn with replacement from concentration data, with respect to each food item (correlation between compounds for each food item was accounted for in this process). A mean concentration was estimated with regard to each food category for each PCB congener. For all individuals in the consumption survey, the (average) daily body weight-adjusted dietary exposure over all food groups was estimated with regard to each PCB congener. To assess the uncertainty in the estimated exposures, the procedure described above was repeated 500 times.

For assessing the cumulative exposure, RPFs estimated for the PCBs in Paper III (using PCB 126 as the reference compound) were used. The RPFs were assumed to have a

log-normal distribution with GM equal to the RPF point estimate. The GSD of this distribution was estimated from the two-sided confidence interval for the RPF. In each uncertainty round, the estimated exposure for each PCB congener (described above) was multiplied by a random value drawn from the respective PCB-specific RPF distribution. The cumulative exposure was then estimated as the sum of the adjusted exposures (in terms of PCB 126 equivalents) of all PCBs.

2.2.4 Margin of exposure

In **Paper IV**, the cumulative MOE was estimated. This was performed with and without the incorporation of RPFs. In the RPF-based approach, the cumulative MOE was calculated as the ratio between the RfD distribution (based on 20 000 values) for PCB 126 and the cumulative dietary exposure distribution (20 000 values were randomly generated from a fitted log normal model) expressed in terms of PCB 126 equivalents. In the RPF-free approach, PCB-specific MOEs were estimated as the ratio between the PCB-specific RfD distribution and the PCB specific dietary exposure distribution. The inverse of the PCB-specific MOEs were summed, and the cumulative MOE was calculated as the inverse of this sum. Figures 2A and 2B describe the procedure to derive the RPF-based cumulative MOE.

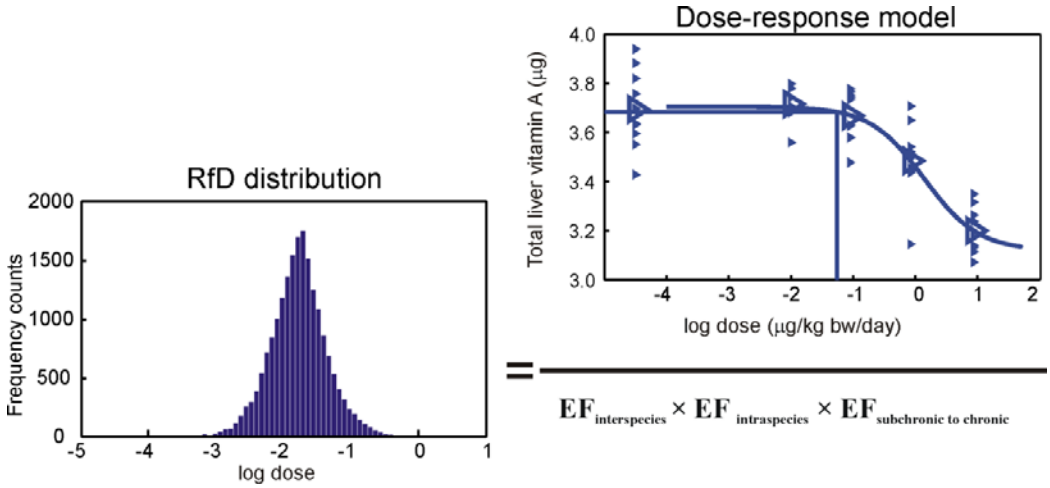

Figure 2A. Animal toxicity data for PCB126 (male rats) was analyzed using the BMD approach. The BMD was divided by extrapolation factors, EFinter, EFintra and EFsubchronic-chronic, by Monte Carlo simulations, resulting in a distribution for the RfD. To account for uncertainties, the process of estimating the RfD was repeated, n = 500 times.

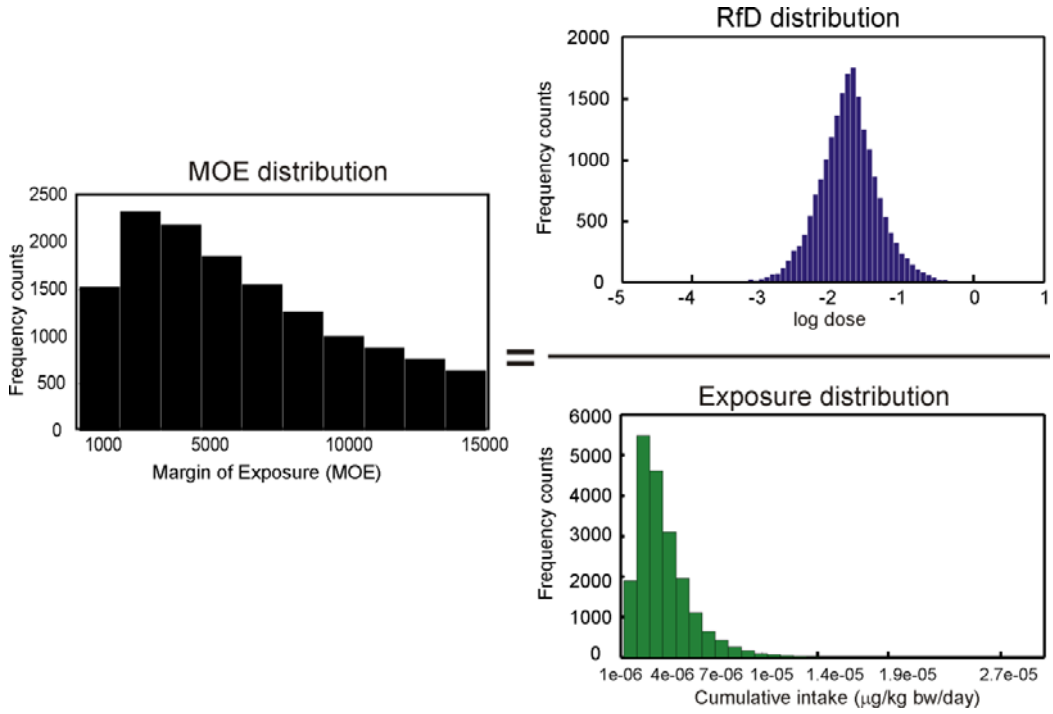

Figure 2B. The distribution for the cumulative exposure (expressed in terms of the PCB126 equivalents) is obtained by 1) combining consumption amounts and average PCB concentrations for each food group; 2) adjusting the total exposure for each PCB by applying the appropriate RPF, and 3) summing adjusted exposures over all PCB congeners. The distribution of the cumulative MOE is the ratio between the RfD distribution and the cumulative exposure distribution, obtained by Monte Carlo simulations. To account for uncertainties, the process of estimating the cumulative MOE was repeated, n = 500 times.

2.2.5 Software

The mathematical and statistical procedures applied in this thesis were developed in Matlab (version 7.0).

2.3 RESULTS AND DISCUSSION

2.3.1 Comparison of dose-response relationships

2.3.1.1 Strain differences in sensitivity

The differences in sensitivity between two strains, species or the difference in potency of two chemicals is traditionally assessed by the comparison of their corresponding NO(A)ELs (LO(A)ELs) or ED_{50}s. As has been discussed previously, a NO(A)EL depends heavily on the study setup and therefore it is regarded to contain substantial error. This error becomes even larger for the NO(A)ELs ratio. Whereas, the BMD is derived by using the whole data and it is expected to be more informative compared to the NO(A)EL.

In **Papers I** and **II**, differences in sensitivity between L–E (dioxin sensitive) and H/W (dioxin resistant) rats following long-term exposure to TCDD were quantitatively investigated. Although the effects of dioxin exposure in the two rat strains has been extensively studied before, this was the first time a detailed statistical analysis was employed to quantify the relative difference in sensitivity between the two strains following long-term TCDD exposure.

In **Paper I**, the difference in sensitivity between L-E and H/W rats was investigated for a variety of toxicological endpoints including data on body and organ weights, hepatic foci, hepatic CYP1A1 induction, as well as tissue retinoid levels. For a given endpoint, the BMDs for L-E and H/W rats were estimated by the Hill function that was fitted to the dose-response data. In **Paper I**, the BMD derived for every endpoint was defined as the dose corresponding to a percentage change in response relative to background response as well as a percentage change in response relative to the range between maximum and background response (equations 3 and 5). The sensitivity difference was then estimated in terms of a BMD ratio, i.e. $BMD_{H/W}/BMD_{L-E}$. The BMD ratio becomes a constant value that is independent of the BMR level used in the BMD calculation when the dose-response models are parallel, i.e. when the two curves have the same dose-response shapes, but the location of the curves on the dose scale may still differ. The BMD ratio becomes dependent on the selected BMR level when the dose-response curves are not parallel. The BMD ratio confidence interval reflecting the uncertainty in the sensitivity difference was also estimated. If this confidence interval

contained the value 1, the two strains were not regarded to differ statistically in their sensitivity to dioxin exposure.

The statistical analysis demonstrated that the assumption of parallel dose-response curves was accepted for most of the parameters investigated. The parallel curve assumption was rejected for the data on volume fraction of hepatic foci, and hepatic retinyl palmitate, and hence for these endpoints the BMD ratio is not constant. A range of BMRs were therefore employed for calculating the BMD ratio for these endpoints.

It was concluded that L-E and H/W rats differed statistically in their response to TCDD treatment for most of the parameters investigated in this study, i.e. the confidence interval corresponding to the $BMD_{H/W}/BMD_{L-E}$ ratio did generally not include a BMD ratio = 1. Differences in response between the strains were most pronounced for hepatic foci; L-E rats were approximately 20-40 times more sensitive than H/W rats. For body and organ weight parameters, L-E rats were approximately 10-20 times more sensitive than H/W rats. For retinoid parameters and hepatic CYP1A1 induction, estimated differences between the strains were generally about 5-fold, and associated with a low uncertainty (Paper I, Table 2). The dose-response model and profile likelihood curve generated for thymus weight data are shown in Figures 3A and 3B.

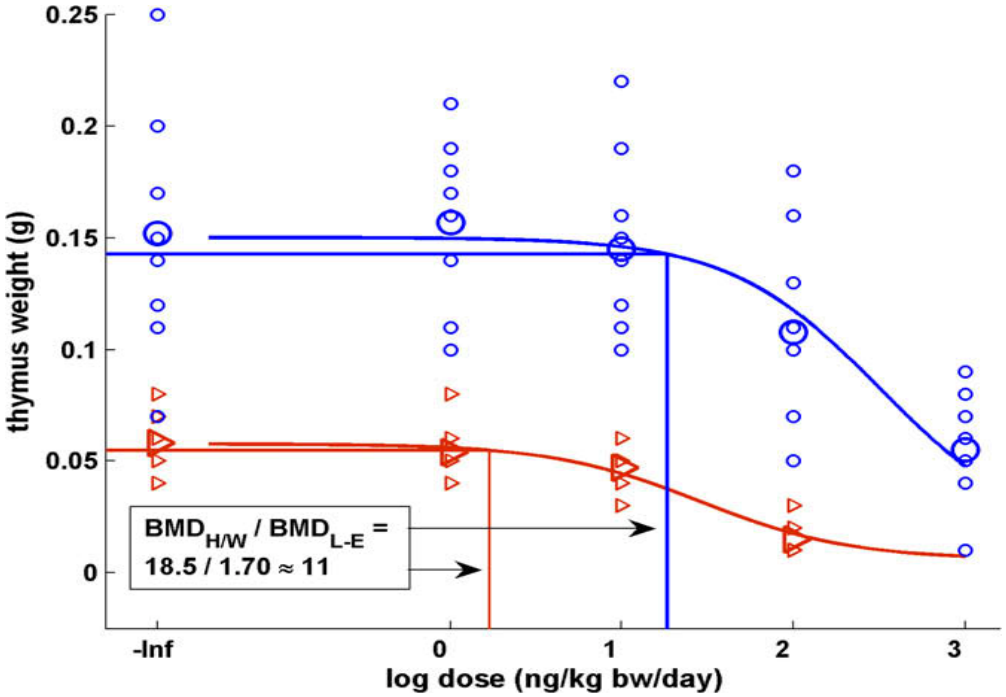

Figure 3A. The Hill function fitted to data on thymus weight observed in L-E (triangles) and H/W (circles) rats following long-term TCDD exposure. The point estimate of the BMD ratio presenting the strain difference in sensitivity to TCDD exposure is 11. The BMD is defined as a dose corresponding to 5% change in response relative to background. However, as the dose-response curves could be assumed to be parallel the BMD ratio is independent of the BMR level selected.

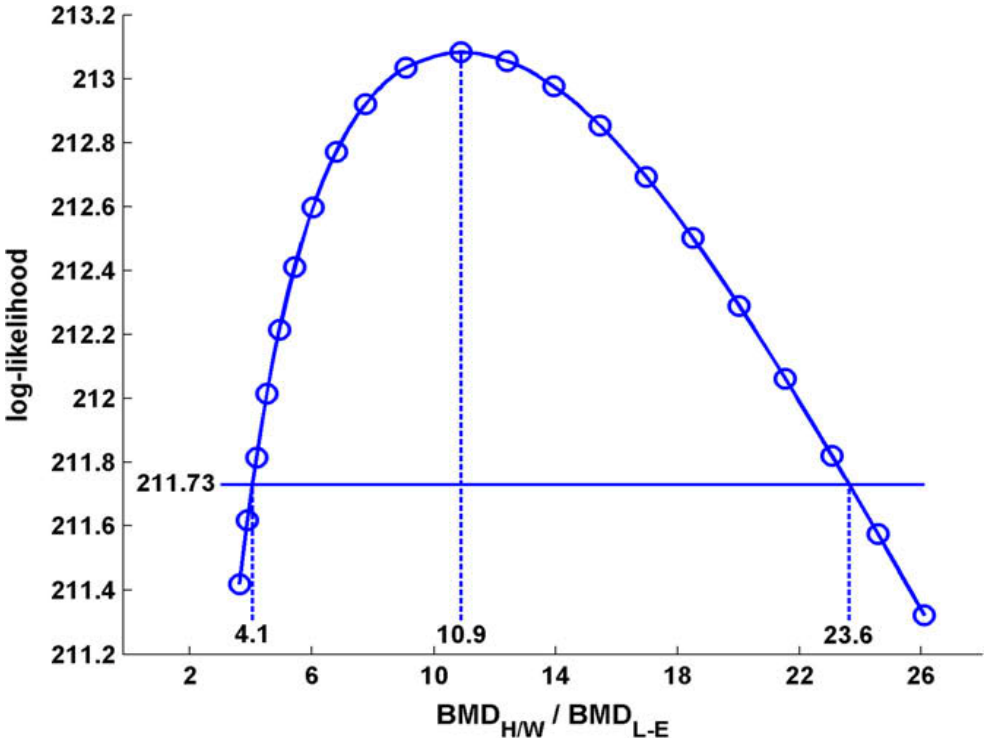

Figure 3B. Establishment of a confidence interval for the BMD ratio using the profile likelihood method. The 90% confidence interval for the BMD ratio was estimated to 4.1 - 23.6. Since the confidence interval does not include a BMR ratio = 1, the two strain are considered to differ statistically in their sensitivity to TCDD exposure.

In **Paper II**, data on bone geometry, mineral density and biomechanical properties that were derived from the same toxicity study as presented in **Paper I** were considered. In this investigation, for every endpoint, the BMD and BMDL were estimated and compared to the associated NOAEL value. The relative difference in sensitivity between L-E and H/W rat strains, in terms of the ratio between their corresponding BMDs, was also investigated for bone parameters where both strains showed statistically significant effects. The BMD was defined as corresponding to the default BMR level of 5% change in response relative to background, according to equation 3. In order to assess the influence of BMR level on the results, additional calculations were also performed considering a BMR level equal to a 10% change in response relative to background.

The results in **Paper II** indicated that the BMD approach is a more suitable method for evaluation of bone parameters compared to the NOAEL method. For a few endpoints, NOAEL values could not be established whereas the corresponding BMDs could be identified. For the cases where a NOAEL value could be derived, it was in

the same range as the BMDL corresponding to a 5% response. This indicates that despite the assumption that NOAEL is considered as a risk/effect-free dose, it does in fact correspond to an unknown effect size. The effect size associated with BMD, on the other hand, is explicitly presented which may be considered to result in more transparent risk assessments. BMDLs corresponding to a response level of 10% were for most parameters higher than the corresponding NOAELs and 2-10 times higher than BMDLs corresponding to a response level of 5%. In overall, a response level of 5% seems to be relevant for the estimation of BMDs and BMDLs corresponding to the different bone parameters that were investigated in **Paper II**.

For those bone parameters where both strains showed a significant dose-response relationship, strain sensitivity differences were estimated by comparing their corresponding dose-response relationship as illustrated for data on cross sectional area of proximal tibia in Figure 4. For these nine parameters, the dose-response curves of L-E and H/W rats could be assumed to be parallel according to the likelihood ratio test; there was a 10-fold strain difference for energy absorption of proximal tibia and length of tibia, and a 49-fold difference for cross-sectional area of proximal tibia (Figure 4), while there was no statistically significant strain difference with respect to the other six parameters (Paper II, Table 6). It should be born in mind that for some parameters there was no observed effect for H/W rats and therefore taken all endpoints together the data suggest a significant difference in sensitivity between two strains with the L-E rat being the most sensitive strain. This finding provides support for the distinct role of AhR for the effects of TCDD on adult bone. The results obtained in **Paper II** provide new quantitative information about TCDD-induced bone alterations at doses which are of relevance from a health risk assessment point of view and suggest that the BMD approach is an appropriate method for such evaluation.

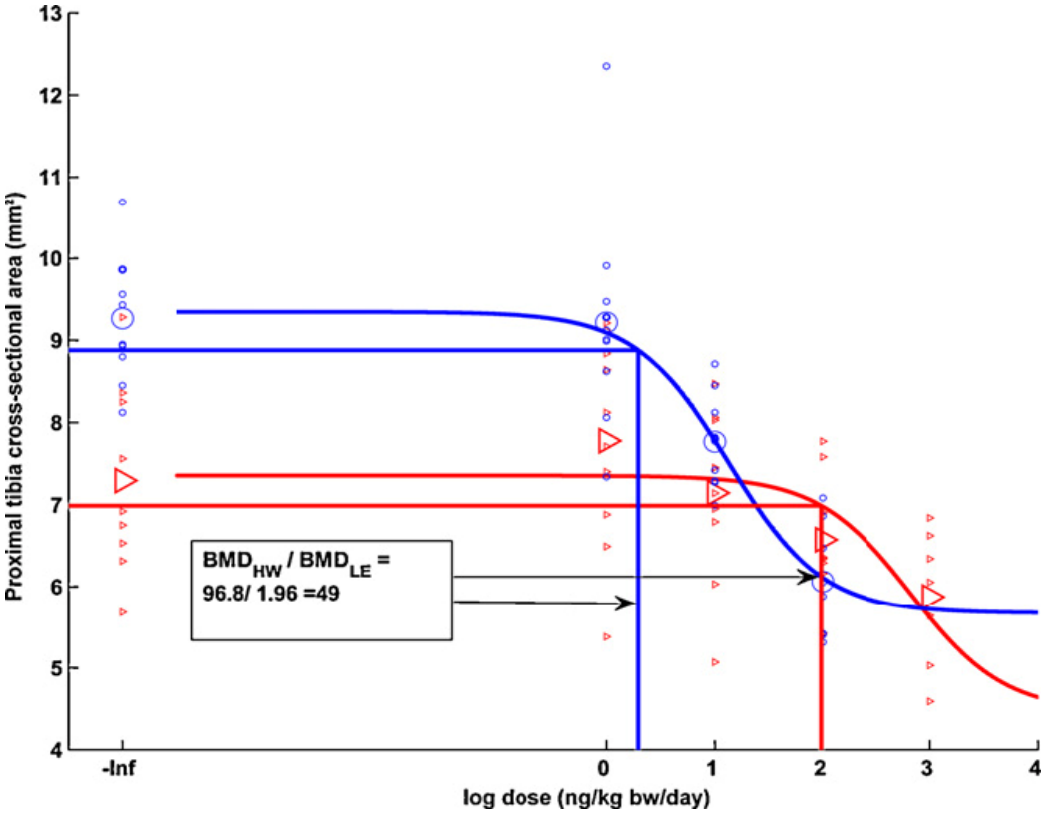

Figure 4. The Michael-Menten equation fitted to data on cross-sectional area of proximal tibia observed in L-E (circles) and H/W (triangles) rats following long-term TCDD exposure. The point estimate of the BMD ratio presenting the strain difference in sensitivity to TCDD exposure is 49. The BMD is defined as a dose corresponding to 5% change in response relative to background. However, as the dose-response curves could be assumed to be parallel the BMD ratio is independent of the BMR level selected.

2.3.1.2 Establishment of relative potency values

In **Paper III,** REP values for individual PCB congeners compared to PCB 126 (as the reference compound) were established using traditional approaches, i.e. as the ratio between NO(A)ELs/LO(A)ELs or median effective doses (ED_{50}s), and a more recent approach, i.e. as the ratio between BMDs. Data on male and female rats from a series of single compound experiments including DL PCBs 77, 105, 118 and 126 (included in WHO-TEF concept), as well as NDL PCBs 28, 128 and 153 (not included in the WHO-TEF concept) was evaluated. Increased liver weight, decreased hepatic vitamin A, and hepatic EROD induction were the endpoints considered for this evaluation. Increased liver weight is a well-established organohalogen endpoint. Although alterations in hepatic vitamin A and hepatic EROD activity may not be toxic *per se*, hepatic vitamin A reduction can be considered as a sensitive indicator of an altered retinoid homeostasis, and induction of EROD activity is indicative of AhR activation.

The BMR associated with the BMD was defined as a percent change in response relative to the background (equation 3). The BMR level was specified to a 5% change for hepatic vitamin A and relative liver weight data, and a 100% change for hepatic EROD activity data. In **Paper III**, the data was analyzed on the log-response scale. The discussed response changes above concerns the normal response scale, and they translate to an absolute response change on the log-response scale which was used in practice (Paper III, Supplementary materials).

In **Paper III**, results demonstrated that for hepatic vitamin A and liver weight data the BMD-based REP values were for some cases higher and for some cases lower than the NOEL-based REP values. The NOEL-based REP values were 1-5 times higher than the BMD-based REP values for female liver weight data after exposure to PCBs 128, 153 and 156, female hepatic vitamin A data after exposure to PCBs 77, 153 and 156 and male hepatic vitamin A data after exposure to PCBs 77, 128 and 156. Whereas, for the remaining PCB congeners the BMD-based REP values were 1-3 times higher than the NOEL-based REP values. For the EROD activity data, the BMD-based REP values were in general 1-56 times higher than NOEL-based REP values (Table 1).

The assumption of parallel dose-response curves was supported for most congeners considering the data on hepatic vitamin A and relative liver weight. This assumption was, however, not supported for most congeners considering the EROD activity data. Moreover, further statistical analysis showed that male and female BMD-based REP values were equal for most congeners in the case of hepatic vitamin A and liver weight endpoints, but only for PCB 128 in the case of hepatic EROD activity data (Figure 5, Table 1).

Under the parallel curve assumption the BMD-based REP values become identical to the ED_{50}-based REP values since the dose ratios do not depend on the selected response level, or BMR, in this case. In the case of non-parallel dose-response relationships the ED_{50}-based REP values tended to be higher than BMD-based REP values (Table 1). Furthermore, in contrast to the BMD values that were derived in the low dose area, accurate estimation of ED_{50} values demands more reliable estimation of the maximum response level. Therefore in the cases where there is a lack of response information at high dose levels, ED_{50}-based REP values can be less precise than BMD-based REP values. Among the endpoints investigated in **Paper III** this situation was most

prominent for PCB 156 (based on male liver weight data) where the very low ED_{50} REP was due to estimation of ED_{50} for PCB 156 outside the range of observations.

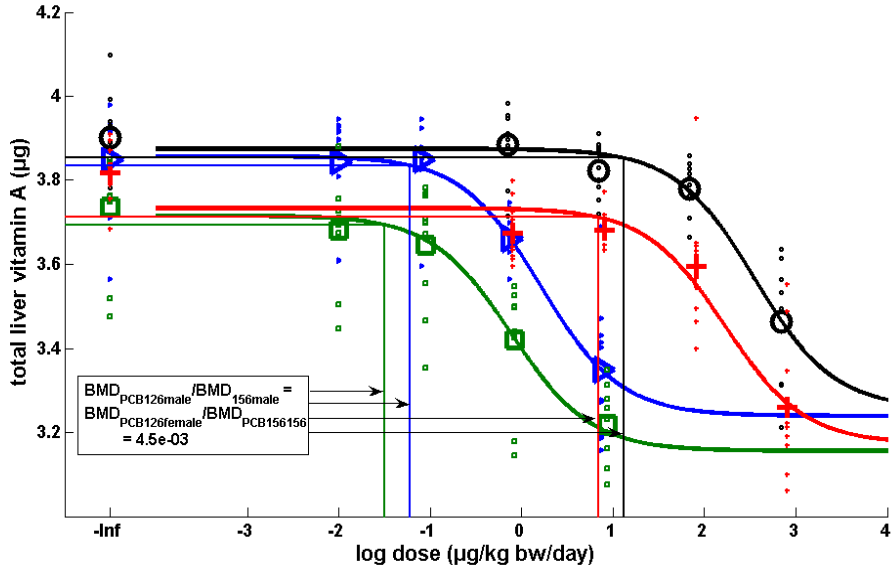

Figure 5. Michael-Menten equations fitted to hepatic vitamin A data (µg) observed in male and female Sprague-Dawley rats following subchronic exposure to PCB 156 (µg/kg bw/day) and PCB 126 (µg/kg bw/day). Squares and triangles indicate female and male data on PCB 126, respectively. Crosses and circles represent female and male data on PCB 156, respectively. According to the statistical analysis, the dose-response curves for PCB 126 and PCB 156 are parallel for the male data. The parallel curve assumption was also supported for the female data. The statistical analysis supports the assumption of equal REP values for the male and female data, i.e. $k_{malePCB\ 126} / k_{malePCB\ 156} = k_{femalePCB\ 126} / k_{femalePCB\ 156}$. Hence, the REP value (BMD ratio) is independent of the response level selected and can be summarized as a single value for male and female data.

In **Paper III**, the uncertainty associated with the BMD-based REP values was presented as two-sided 90% confidence intervals. A confidence interval may also be implemented for ED_{50}-based REP values, however such estimation is not possible for NO(A)EL-based REP values. The REP point estimates as well as the associated confidence intervals can be used as relative potency factors (RPFs) in cumulative risk assessments.

REP values could be established based on one or more of the endpoints analysed for congeners assigned a TEF (PCBs 77, 105, 118 and 156) as well as congeners not assigned a TEF (PCBs 28, 128 and 153). The highest REP values were obtained for

relative liver weight; REP values based on decreased hepatic vitamin A levels were slightly lower; and the lowest REP values were in general obtained for data on hepatic EROD activity (**Table 1**). The BMD-based REP values corresponding to DL PCB congeners were similar to their corresponding established WHO-TEF values. For assessment of the cumulative effects of persistent chemical mixtures, the findings in **Paper III** suggest that in addition to the mode-of-action approach, development of a system that includes chemicals with similar effects despite their different underlying mechanisms of action could be useful.

Table 1. Relative potency (REP) values estimated as BMD, ED_{50}, and NO(A)EL ratios based on relative liver weight (%), hepatic vitamin A (μg) and hepatic EROD activity (nmol/ mg protein/hr) observed in male and female rats after exposure to individual PCB congeners.

PCB congener	TEF[a]	Liver weight			Hepatic vitamin A			Hepatic EROD activity		
		BMD REP	ED_{50} REP	NOEL REP	BMD REP	ED_{50} REP	NOEL REP	BMD REP	ED_{50} REP	LOEL[b] REP
MALE										
28								$6.8e^{-07}$	$8.6e^{-05}$	$2.6e^{-06}$
77	$1.0e^{-03}$				$9.0e^{-04}$	$8.6e^{-03}$	$1.1e^{-03}$	$1.8e^{-04}$	$3.7e^{-04}$	$1.3e^{-05}$
105	$3.0e^{-04}$	$5.0e^{-04}$	$5.0e^{-04}$	$2.0e^{-04}$	$5.7e^{-04}$	$5.7e^{-04}$	$2.0e^{-04}$	$7.9e^{-05}$	$6.5e^{-05}$	$2.3e^{-06}$
118	$3.0e^{-04}$							$1.2e^{-05}$	$5.9e^{-05}$	$1.5e^{-05}$
128					$4.0e^{-05}$	$4.0e^{-05}$	$1.9e^{-04}$	$4.9e^{-06}$	$4.9e^{-06}$	$2.4e^{-06}$
153		$5.4e^{-04}$	$5.4e^{-04}$	$2.3e^{-04}$	$2.7e^{-04}$	$2.7e^{-04}$	$2.3e^{-04}$	$2.1e^{-04}$	$1.3e^{-03}$	$2.9e^{-04}$
156	$3.0e^{-04}$	$1.5e^{-03}$	$5.0e^{-16}$	$1.2e^{-03}$	$4.5e^{-03}$	$4.6e^{-03}$	$1.2e^{-02}$	$3.0e^{-04}$	$2.3e^{-04}$	$1.4e^{-05}$
FEMALE										
28								$2.4e^{-06}$	$2.3e^{-06}$	$2.5e^{-06}$
77	$1.0e^{-03}$				$6.0e^{-04}$	$6.0e^{-04}$	$1.1e^{-03}$	$2.9e^{-04}$	$7.2e^{-04}$	$1.1e^{-04}$
105	$3.0e^{-04}$	$3.0e^{-03}$	$3.0e^{-03}$	$1.9e^{-03}$	$3.2e^{-04}$	$3.2e^{-04}$	$2.3e^{-04}$	$1.4e^{-04}$	$1.1e^{-04}$	$2.5e^{-06}$
118	$3.0e^{-04}$							$1.1e^{-04}$	$1.1e^{-04}$	$5.9e^{-05}$
128		$4.1e^{-04}$	$4.1e^{-04}$	$2.0e^{-03}$	$3.9e^{-05}$	$3.9e^{-05}$	$2.3e^{-05}$	$5.4e^{-06}$	$6.5e^{-06}$	$2.3e^{-06}$
153		$7.0e^{-04}$	$7.0e^{-04}$	$2.0e^{-03}$	$4.5e^{-05}$	$4.5e^{-05}$	$2.3e^{-05}$	$7.5e^{-03}$	$3.6e^{-02}$	$2.4e^{-04}$
156	$3.0e^{-04}$	$9.8e^{-03}$	$9.8e^{-03}$	$1.0e^{-02}$	$3.7e^{-03}$	$3.7e^{-03}$	$1.1e^{-02c}$	$6.2e^{-04}$	$5.3e^{-04}$	$1.2e^{-05}$

[a]WHO TEF values (Van den Berg et al. 2006) have been adjusted to account for the use of PCB 126 as a reference compounds, i.e. by dividing with the TEF for PCB 126 (TEF = 0.1).
[b]For hepatic EROD activity data, a NOEL value could not be established for PCB 126 and the calculations have been based on LOEL values for all congeners.

2.3.2 Establishment of the cumulative margin of exposure

In **Paper IV**, a probabilistic approach for integrated evaluation of dose-response and exposure data was further developed and applied to estimate the cumulative MOE for

the combined exposure to DL PCBs 77, 105, 118, 156 and NDL PCB 153. There are a number of studies that have used a probabilistic risk assessment framework for MOE estimation, but these studies have mainly focused on pesticides (Bosgra et al. 2009; Muller et al. 2009). In contrast to the conventional MOE approaches, which are calculated as the ratio between a BMDL (or NO(A)EL) and a point estimate of the human exposure, the variability and uncertainty in both components of the MOE are taken into account under a probabilistic framework. Along these lines, a distribution (describing variability) for the MOE was estimated as the ratio between a distribution for a reference dose and a distribution for the human dietary exposure in **Paper IV**. The uncertainty in the respective distribution was also assessed in this process. Data on hepatic vitamin A from the same series of toxicity studies on individual PCB congeners that was used as basis in **Paper III** was used as an example endpoint for this analysis.

In **Paper IV**, an RPF-based approach as well as an RPF-free approach was considered for estimating the cumulative MOE. In the RPF-based approach, the cumulative PCB exposure was estimated by the use of RPFs. Individual REPs derived for PCB 28, 105, 118, 153 and 156 (using PCB 126 as reference) in **Paper III** were used in **Paper IV** as RPFs to express the concentration of individual compounds in equivalents of PCB 126 (Table 2). The analysis in **Paper III** suggested that these REPs were similar for male and female data except for PCB 153. In **Paper IV**, however, a common/single RPF was also used for male and female data on PCB 153 since they were close enough to be considered as similar in a practical context. The RPFs were assumed to be log-normally distributed, characterized by a geometric mean equal to the point estimates of the RPFs (Table 2) and a geometric standard deviation that was estimated from a two-sided 90% confidence interval established for each RPF (Table 2).

Under the RPF-based approach, the cumulative MOE was estimated by comparing the RfD for PCB 126 with the cumulative exposure expressed in terms of PCB 126 equivalents. Under the RPF-free approach, the cumulative MOE is estimated by the use of compound specific MOEs. Detailed description of these two approaches can be found in the materials and methods section in **Paper IV**.

Table 2. Relative potency factors (RPFs) for individual PCB congeners (relative to PCB 126), based on hepatic Vitamin A data for male and female Sprague-Dawley rats.

PCB Congener	**RPF (Lower bound (P5), Upper bound (P95))**
PCB 28	NA[a]
PCB 77	9.0e-04 (5.6e-04, 1.3e-03)
PCB 105	4.5e-04 (2.6e-04, 7.8e-04)
PCB 118	NA
PCB 153	1.5e-04 (1.0e-04, 2.0e-04)
PCB 156	4.5e-03 (2.6e-03, 6.9e-03)

[a] Not applicable: There was no statistically significant difference between the saturated model and the no-response model, i.e. no dose-response relationship was observed.

The 0.1st, 1st, 5th, and 50th percentiles of the cumulative MOE were estimated under the two approaches considered (Paper IV, Table 5). The lowest value for the cumulative MOE was 20; it suggests that 1 out of 1000 women have a MOE less than 20 (and with high confidence this value is not less than 5, if also accounting for the uncertainty). The corresponding value for men was about 70. Depending on the percentile considered, the cumulative MOE could be a factor of 2 - 4 lower for women compared to men. This difference was mainly due to differences in the estimated RfD values for men and women, while the estimated exposure levels for men and women were similar (Paper IV, Tables 3 and 4). The results indicated that the cumulative MOE, more or less, reflected the MOE for PCB 126; the other PCB congeners had little contribution to the cumulative exposure, and thus the cumulative MOE (Figure 6).

It has been discussed that the RPF-based and RPF-free approaches used in **Paper IV** are equivalent (van der Voet et al. 2009). However in practice, the implementation of the two approaches is different. The RPF-based approach requires the direct use of RPFs; such values are usually established as single (average/median) values for each compound and the uncertainty in the RPFs can also be taken into account, as discussed in **Paper IV**. However, in van der Voet (2009), the RPFs have been defined as the ratio between individual reference doses that accounts for both variability and uncertainty, i.e. not conventional BMD ratios based on experimental data only. Hence, the RPF-free approach can be considered as indirectly accounting for both variability

and uncertainty in relative potency. Thus, the main difference between the implementation of the two approaches discussed in **Paper IV** is that uncertainty in relative potency is only accounted for in the RPF-based approach, while both variability and uncertainty in relative potency is indirectly accounted for in the RPF-free approach. This may explain the minor differences observed between the two approaches (Paper IV, Table 5).

While the RPF-free approach accounts for variability (beside uncertainty) at the level of relative potency, it may be regarded to be more data intensive, compared to the RPF-based approach, since it requires detailed dose-response data for all compounds included. An advantage of the RPF-based approach is that the implementation of this approach only requires detailed dose-response information for the reference compound (PCB 126 in this case), and RPFs that have been derived in previous studies can be used.

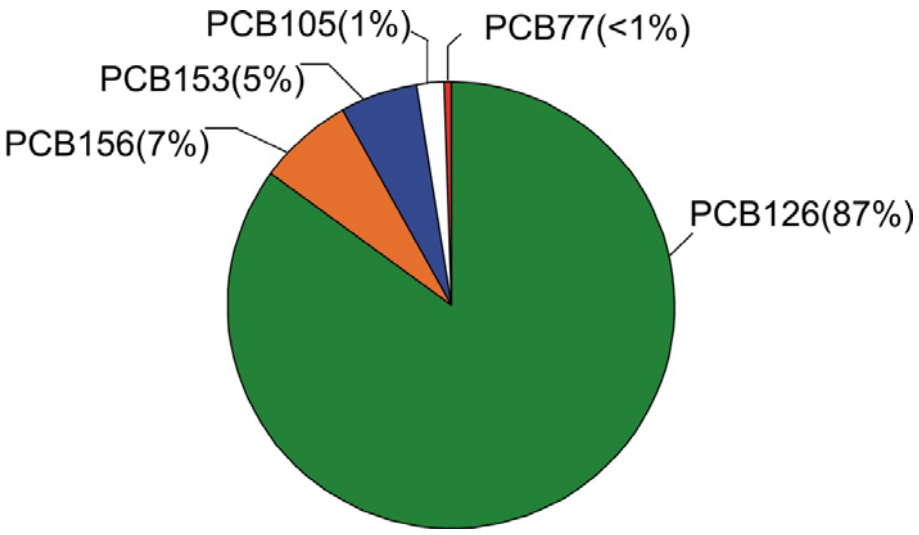

Figure 6. The relative contribution of individual PCB congeners (in terms of mean values) to the cumulative dietary PCB intake, in terms of PCB 126 equivalents, for men 17-74 years of age. A similar result was observed for women.

2.4 CONCLUSIONS

This thesis work provides an important input to the evaluation and application of quantitative risk assessment methods thereby contributing to further development in this area.

Establishment of dose-response relationship for different categories of effects, identification of strain and species differences, and assessment of chemical potency differences are elements that can be identified in hazard characterization, which is a central step in the risk assessment process. **Part one** of this thesis has demonstrated such evaluations, and it has discussed the advantages of performing detailed analysis of the entire dose-response data compared to using a NO(A)EL approach. More specifically, it was concluded that:

- For most parameters investigated in **Paper I**, L-E and H/W data differed statistically in their response to TCDD treatment. In general L-E rats were 5-20 times more sensitive than H/W rats, and this difference was in the range of 20 to 40-fold in the case of hepatic foci. The strain differences were less uncertain (shorter confidence interval) for the retinoid parameters and EROD activity.
- The BMD approach is more suitable than the NOAEL approach for evaluation of bone parameters considered in **Paper II**. BMDLs corresponding to a response level of 5% were in the same range as the NOAEL. It was concluded that the two strains are significantly different in their response to TCDD treatment, confirming the distinct role of AhR, with the L-E rat being the most sensitive strain. This difference was most pronounced (about 49-fold) for cross-sectional area of proximal tibia.
- For the PCB congeners considered in **Paper III** the BMD approach resulted in REP values that were in general more reliable compared to the NOEL and ED_{50} approaches. In addition, our results support further development and use of endpoint specific ranking systems based on BMD-derived REP values for assessment of human exposure to mixtures of chemicals with similar as well as different mode of actions.

Part two of this thesis concerned development of quantitative approaches for risk characterization. The cumulative MOE was in this study estimated for a group of PCBs using reduction of hepatic retinoids as an example endpoint. A framework that accounts for variability and uncertainty, with regard to both components of the MOE, was applied. From **Paper IV**, it was concluded that:

- The median of the 0.1st percentile for the cumulative MOE was about 20 for women, and about 70 for men. The cumulative MOE, more or less, reflected the MOE for PCB 126; the other PCB congeners had little contribution to the cumulative exposure, and thus the cumulative MOE.
- The RPF-free approach more completely accounts for variability and uncertainty compared to the RPF-based approach. On the other hand, the RPF-based approach is less data intensive and can be more easily implemented in practice, allowing for a use of historical data on RPFs.
- Compared to conventional MOE approaches, the approaches discussed in **Paper IV** provide an improved tool under which potential health concerns can be assessed by accounting for both variability and various uncertainties involved quantitatively, contributing to improving cumulative risk assessments.

2.5 FUTURE PERSPECTIVES

As has been discussed, the BMD approach is considered to be preferred over the NO(A)EL method for establishment of health-based guidance values or for assessment of differences in sensitivity between species, strains, genders, etc. One of the major advantages of the BMD approach is that it involves uncertainty analysis (e.g. accounting for sample size). So far, there are not many studies that have investigated the influence of study design on the BMD, i.e. the impact of dose placement, and number of animals distributed at each dose level, on the uncertainty in the BMD or BMD ratio. Further investigation of this issue may help to optimize animal use in toxicological studies, resulting in a better allocation of available resources.

In the case of non-parallel dose-response relationships, REPs cannot be established as single values, i.e. the REP will depend on the BMR. How to account for the case of non-parallel dose-response curves when assessing the health impact of cumulative or

combined exposures is a future challenge. In general, further investigation of how to deal with non-dose-additive compounds by analyzing dose-response data of individual compounds as well as mixtures of compounds is of interest.

In the probabilistic risk characterization framework discussed in **Paper IV**, the overall variability and uncertainty was estimated. An additional step is to include an approach that identifies the relative contribution of each source of variability/uncertainty with respect to the overall variability/uncertainty. This will help to identify the most important sources of MOE variation between individuals and/or the most important sources of uncertainty (that perhaps can be reduced e.g. by additional data collection).

3 ACKNOWLEDGEMENTS

The work presented in this thesis was performed at the Institute of Environmental Medicine (IMM), Karolinska Institutet. I wish to use this opportunity to express my sincere thanks to those who I owe my progress to their help and without their assistance this work was not possible!

I deem it my duty to thank Professor Helen Håkansson for taking me as a PhD student and providing me support and guidance during these years! Thank you for your encouragements and patience to improve my research abilities. Thanks for introducing me the field of health risk assessment and thanks for your positive attitude to the new methodologies in this filed and letting me apply them!

I would like to express my gratitude to Dr. Salomon Sand, for all that he has taught me. Thank you for interesting scientific discussions even over the phone, for your constant guidance and for endless patience and support and positive and encouraging attitude.

My mentor Dr. Reza Mohammadi, thanks for being around whenever I needed your help and advice and for your positive and relaxing views in solving the issues.

Many thanks to Ali, for your loyal friendship, for always being helpful and supportive! For nice chats and all the funs and for your help in arranging my PhD dissertation!

Emma and Anna, thank you for your great friendship, for sharing your experiences of life in general and feminine experiences in particular. Thank you for all your kindness!

My thanks to the past and present members of the unit; Kina thank you for always being kind and helpful; Daniel for funs and laughs; Mattias for your advices and nice chats; Inga-Lill for all your helps and Swedish-English translations; thanks are extended to Annika, Krister, Maria, Sabina, Lubna, Robert, Fillip, Lina, Lauy, Vaninna, and Katarina for making IMM a nice place to work!

Firoozeh and Shokufeh, thank you for your wonderful friendship, for the fruitful discussions and having lots of fun and for generous love and support!

Dr. Zendedel and Marzieh for your kind friendship, for introducing Stockholm to us and sharing up and downs of the life! I would like to acknowledge you for introducing IMM to me as a nice place to pursue my study!

My sincere thanks to my parents, for all of invaluable supports and encouragements during my life. My family Alireza, Yaser, Leili and Amir, thank you for your continuous caring and support. You are the best!

Last but not the least, I would like to thank my beloved Isaac for all his encouragements and supports throughout the long journey of our common life! My thanks and loves to my girls Fatemeh and Setayesh who light up my life every day! Thank you for your patience when mommy was away for her study!

4 REFERENCES

Ahlborg UG, Becking GC, Birnbaum LS, Brouwer A, Derks HJGM, Feeley M, Golor G, Hanberg A, Larsen JC, Liem AKD, Safe S, Schlatter C, Waern F, Younes M, Yrj. 1994. Toxic equivalency factors for dioxin-like PCBs, Report on WHO-ECEH and IPCS consultation, December 1993.

Allen BC, Kavlock RJ, Kimmel CA, Faustman EM. 1994a. Dose-response assessment for developmental toxicity. II. Comparison of generic benchmark dose estimates with no observed adverse effect levels. Fundam Appl Toxicol 23(4): 487-495.

Allen BC, Kavlock RJ, Kimmel CA, Faustman EM. 1994b. Dose-response assessment for developmental toxicity. III. Statistical models. Fundam Appl Toxicol 23(4): 496-509.

Allen BC, Strong PL, Price CJ, Hubbard SA, Daston GP. 1996. Benchmark dose analysis of developmental toxicity in rats exposed to boric acid. Fundam Appl Toxicol 32(2): 194-204.

Bailer AJ, Noble RB, Wheeler MW. 2005. Model uncertainty and risk estimation for experimental studies of quantal responses. Risk Anal 25(2): 291-299.

Barton HA, Andersen ME, Allen BC. 1998. Dose-response characteristics of uterine responses in rats exposed to estrogen agonists. Regul Toxicol Pharmacol 28(2): 133-149.

Becker W, Pearson M. 2002. Riksmaten 1997-98. Kostvanor och näringsintag i Sverige.

Berry CL. 1988. The no-effect level and optimal use of toxicity data. Regul Toxicol Pharmacol 8(4): 385-388.

Bokkers BG, Slob W. 2005. A comparison of ratio distributions based on the NOAEL and the benchmark approach for subchronic-to-chronic extrapolation. Toxicol Sci 85(2): 1033-1040.

Bokkers BG, Slob W. 2007. Deriving a data-based interspecies assessment factor using the NOAEL and the benchmark dose approach. Crit Rev Toxicol 37(5): 355-373.

Boobis AR, Ossendorp BC, Banasiak U, Hamey PY, Sebestyen I, Moretto A. 2008. Cumulative risk assessment of pesticide residues in food. Toxicol Lett 180(2): 137-150.

Bosgra S, Bos PM, Vermeire TG, Luit RJ, Slob W. 2005. Probabilistic risk characterization: an example with di(2-ethylhexyl) phthalate. Regul Toxicol Pharmacol 43(1): 104-113.

Bosgra S, van der Voet H, Boon PE, Slob W. 2009. An integrated probabilistic framework for cumulative risk assessment of common mechanism chemicals in food: an example with organophosphorus pesticides. Regul Toxicol Pharmacol 54(2): 124-133.

Budtz-Jorgensen E, Grandjean P, Keiding N, White RF, Weihe P. 2000. Benchmark dose calculations of methylmercury-associated neurobehavioural deficits. Toxicol Lett 112-113: 193-199.

Budtz-Jorgensen E, Keiding N, Grandjean P. 2001. Benchmark dose calculation from epidemiological data. Biometrics 57(3): 698-706.

Budtz-Jorgensen E, Keiding N, Grandjean P. 2004. Effects of exposure imprecision on estimation of the benchmark dose. Risk Anal 24(6): 1689-1696.

Chu I, Nakai J, Yagminas A, Poon R, Valli T, Håkansson H, Bergman Å. 2000. Toxicity of 2,3,3',4,4',5-hexachlorobiphenyl (PCB 156) in rats. Organohalogen compounds 49: 185-188.

Chu I, Poon R, Yagminas A, Lecavalier P, Hakansson H, Valli VE, Kennedy SW, Bergman A, Seegal RF, Feeley M. 1998. Subchronic toxicity of PCB 105 (2,3,3',4,4'-pentachlorobiphenyl) in rats. J Appl Toxicol 18(4): 285-292.

Chu I, Villeneuve DC, Yagminas A, Lecavalier P, Hakansson H, Ahlborg UG, Valli VE, Kennedy SW, Bergman A, Seegal RF, et al. 1995. Toxicity of PCB 77 (3,3',4,4'-tetrachlorobiphenyl) and PCB 118 (2,3',4,4'5-pentachlorobiphenyl) in the rat following subchronic dietary exposure. Fundam Appl Toxicol 26(2): 282-292.

Chu I, Villeneuve DC, Yagminas A, LeCavalier P, Poon R, Feeley M, Kennedy SW, Seegal RF, Hakansson H, Ahlborg UG, et al. 1994. Subchronic toxicity of 3,3',4,4',5-pentachlorobiphenyl in the rat. I. Clinical, biochemical, hematological, and histopathological changes. Fundam Appl Toxicol 22(3): 457-468.

Chu I, Villeneuve DC, Yagminas A, Lecavalier P, Poon R, Feeley M, Kennedy SW, Seegal RF, Hakansson H, Ahlborg UG, Valli VE, Bergman A. 1996a. Toxicity of 2,2',4,4',5,5'-hexachlorobiphenyl in rats: effects following 90-day oral exposure. J Appl Toxicol 16(2): 121-128.

Chu I, Villeneuve DC, Yagminas A, Lecavalier P, Poon R, Hakansson H, Ahlborg UG, Valli VE, Kennedy SW, Bergman A, Seegal RF, Feeley M. 1996b. Toxicity of 2,4,4'-trichlorobiphenyl in rats following 90-day dietary exposure. J Toxicol Environ Health 49(3): 301-318.

Clewell HJ, Crump KS. 2005. Quantitative estimates of risk for noncancer endpoints. Risk Anal 25(2): 285-289.

Clewell HJ, Lawrence GA, Calne DB, Crump KS. 2003. Determination of an occupational exposure guideline for manganese using the benchmark method. Risk Anal 23(5): 1031-1046.

Crump K. 2002. Critical issues in benchmark calculations from continuous data. Crit Rev Toxicol 32(3): 133-153.

Crump KS. 1984. A new method for determining allowable daily intakes. Fundam Appl Toxicol 4(5): 854-871.

Crump KS. 1995. Calculation of benchmark doses from continuous data. Risk Anal 15: 79-89.

Davis JA, Gift JS, Zhao QJ. 2011. Introduction to benchmark dose methods and U.S. EPA's benchmark dose software (BMDS) version 2.1.1. Toxicol Appl Pharmacol 254(2): 181-191.

Dorne JL, Walton K, Renwick AG. 2001. Uncertainty factors for chemical risk assessment. human variability in the pharmacokinetics of CYP1A2 probe substrates. Food Chem Toxicol 39(7): 681-696.

Dorne JL, Walton K, Renwick AG. 2004. Human variability in the renal elimination of foreign compounds and renal excretion-related uncertainty factors for risk assessment. Food Chem Toxicol 42(2): 275-298.

Edler L, Poirier K, Dourson M, Kleiner J, Mileson B, Nordmann H, Renwick A, Slob W, Walton K, Wurtzen G. 2002. Mathematical modelling and quantitative methods. Food Chem Toxicol 40(2-3): 283-326.

EFSA. 2005. Opinion of the Scientific Panel on Contaminants in the Food Chain on a request from the Commission related to the presence of non dioxin-like polychlorinated biphenyls (PCB) in feed and food. The EFSA Journal 284: 1-137.

EFSA. 2009a. Guidance of the Scientific Committee on a request from EFSA on the use of the benchmark dose approach in risk assessments The EFSA Journal 1150: 1 - 72.

EFSA. 2009b. Cadmium in food. Scientific Opinion of the Panel on Contaminants in the Food Chain. The EFSA Journal 980: 1-139.

Falk-Filipsson A, Hanberg A, Victorin K, Warholm M, Wallen M. 2007. Assessment factors--applications in health risk assessment of chemicals. Environ Res 104(1): 108-127.

Faustman EM, Allen BC, Kavlock RJ, Kimmel CA. 1994. Dose-response assessment for developmental toxicity. I. Characterization of database and determination of no observed adverse effect levels. Fundam Appl Toxicol 23(4): 478-486.

Fowles JR, Alexeeff GV, Dodge D. 1999. The use of benchmark dose methodology with acute inhalation lethality data. Regul Toxicol Pharmacol 29(3): 262-278.

Gaylor DW, Aylward LL. 2004. An evaluation of benchmark dose methodology for non-cancer continuous-data health effects in animals due to exposures to dioxin (TCDD). Regul Toxicol Pharmacol 40(1): 9-17.

Gaylor DW, Slikker W, Jr. 1990. Risk assessment for neurotoxic effects. Neurotoxicology 11(2): 211-218.

Jacobson JL, Janisse J, Banerjee M, Jester J, Jacobson SW, Ager JW. 2002. A benchmark dose analysis of prenatal exposure to polychlorinated biphenyls. Environ Health Perspect 110(4): 393-398.

Kavlock RJ, Allen BC, Faustman EM, Kimmel CA. 1995. Dose-response assessments for developmental toxicity. IV. Benchmark doses for fetal weight changes. Fundam Appl Toxicol 26(2): 211-222.

Kim AH, Kohn MC, Portier CJ, Walker NJ. 2002. Impact of physiologically based pharmacokinetic modeling on benchmark dose calculations for TCDD-induced biochemical responses. Regul Toxicol Pharmacol 36(3): 287-296.

Kodell RL, Howe RB, Chen JJ, Gaylor DW. 1991. Mathematical modeling of reproductive and developmental toxic effects for quantitative risk assessment. Risk Anal 11(4): 583-590.

Kortenkamp A. 2007. Ten years of mixing cocktails: a review of combination effects of endocrine-disrupting chemicals. Environ Health Perspect 115 Suppl 1: 98-105.

Krewski D, Van Ryzin J. 1981. Dose response models for quantal response toxicity data. In: Statistics and Related Topics, (Csorgo D, Dawson D, Rao JNK, Saleh E, eds). North-Holland, Amesterdam, 201-231.

Kroes R, Muller D, Lambe J, Lowik MR, van Klaveren J, Kleiner J, Massey R, Mayer S, Urieta I, Verger P, Visconti A. 2002. Assessment of intake from the diet. Food Chem Toxicol 40(2-3): 327-385.

Kuljus K, von Rosen D, Sand S, Victorin K. 2006. Comparing experimental designs for benchmark dose calculations for continuous endpoints. Risk Anal 26(4): 1031-1043.

Kupper LL, Portier C, Hogan MD, Yamamoto E. 1986. The impact of litter effects on dose-response modeling in teratology. Biometrics 42(1): 85-98.

Lecavalier P, Chu I, Yagminas A, Villeneuve DC, Poon R, Feeley M, Hakansson H, Ahlborg UG, Valli VE, Bergman A, Seegal RF, Kennedy SW. 1997. Subchronic toxicity of 2,2',3,3',4,4'-hexachlorobiphenyl in rats. J Toxicol Environ Health 51(3): 265-277.

Muller AK, Bosgra S, Boon PE, van der Voet H, Nielsen E, Ladefoged O. 2009. Probabilistic cumulative risk assessment of anti-androgenic pesticides in food. Food Chem Toxicol 47(12): 2951-2962.

Murrell JA, Portier CJ, Morris RW. 1998. Characterizing dose-response: I: Critical assessment of the benchmark dose concept. Risk Anal 18(1): 13-26.

Pieters MN, Kramer HJ, Slob W. 1998. Evaluation of the uncertainty factor for subchronic-to-chronic extrapolation: statistical analysis of toxicity data. Regul Toxicol Pharmacol 27(2): 108-111.

Price PS, Keenan RE, Swartout JC, Gillis CA, Carlson-Lynch H, Dourson ML. 1997. An approach for modeling noncancer dose responses with an emphasis on uncertainty. Risk Anal 17(4): 427-437.

Pufulete M, Battershill J, Boobis A, Fielder R. 2004. Approaches to carcinogenic risk assessment for polycyclic aromatic hydrocarbons: a UK perspective. Regul Toxicol Pharmacol 40(1): 54-66.

Rai K, Van Ryzin J. 1985. A dose-response model for teratological experiments involving quantal responses. Biometrics 41(1): 1-9.

Renwick AG. 1993. Data-derived safety factors for the evaluation of food additives and environmental contaminants. Food Addit Contam 10(3): 275-305.

Renwick AG, Barlow SM, Hertz-Picciotto I, Boobis AR, Dybing E, Edler L, Eisenbrand G, Greig JB, Kleiner J, Lambe J, Muller DJ, Smith MR, Tritscher A, Tuijtelaars S, van den Brandt PA, Walker R, Kroes R. 2003. Risk characterisation of chemicals in food and diet. Food Chem Toxicol 41(9): 1211-1271.

Renwick AG, Dorne JL, Walton K. 2000. An analysis of the need for an additional uncertainty factor for infants and children. Regul Toxicol Pharmacol 31(3): 286-296.

Sand S, Filipsson AF, Victorin K. 2002. Evaluation of the benchmark dose method for dichotomous data: model dependence and model selection. Regul Toxicol Pharmacol 36(2): 184-197.

Sand S, Portier CJ, Krewski D. 2011. A Signal-to-Noise Crossover Dose as the Point of Departure for Health Risk Assessment. Environ Health Perspect 119(12): 1766-1774.

Sand S, Victorin K, Filipsson AF. 2008. The current state of knowledge on the use of the benchmark dose concept in risk assessment. J Appl Toxicol 28(4): 405-421.

Sand S, von Rosen D, Eriksson P, Fredriksson A, Viberg H, Victorin K, Filipsson AF. 2004. Dose-response modeling and benchmark calculations from spontaneous behavior data on mice neonatally exposed to 2,2',4,4',5-pentabromodiphenyl ether. Toxicol Sci 81(2): 491-501.

Sand S, von Rosen D, Victorin K, Filipsson AF. 2006. Identification of a critical dose level for risk assessment: developments in benchmark dose analysis of continuous endpoints. Toxicol Sci 90(1): 241-251.

Slob W. 1999. Thresholds in toxicilogy and risk assessment International Journal Of Toxicology 18(4): 259-268.

Slob W. 2002. Dose-response modeling of continuous endpoints. Toxicol Sci 66(2): 298-312.

Slob W. 2007. What is a practical threshold? Toxicol Pathol 35(6): 848-849.

Slob W, Moerbeek M, Rauniomaa E, Piersma AH. 2005. A statistical evaluation of toxicity study designs for the estimation of the benchmark dose in continuous endpoints. Toxicol Sci 84(1): 167-185.

Slob W, Pieters MN. 1998. A probabilistic approach for deriving acceptable human intake limits and human health risks from toxicological studies: general framework. Risk Anal 18(6): 787-798.

Swartout JC, Price PS, Dourson ML, Carlson-Lynch HL, Keenan RE. 1998. A probabilistic framework for the reference dose (probabilistic RfD). Risk Anal 18(3): 271-282.

Toyoshiba H, Walker NJ, Bailer AJ, Portier CJ. 2004. Evaluation of toxic equivalency factors for induction of cytochromes P450 CYP1A1 and CYP1A2 enzyme activity by dioxin-like compounds. Toxicol Appl Pharmacol 194(2): 156-168.

USEPA. 1995. The use of the benchmark dose (BMD) approach in health risk assessment. (EPA/630/R-94/007). Risk Assessment Forum, U.S. Environmental Protection Agency, Washington, DC.

USEPA. 2000. Supplementary guidance for conducting health risk assessment of chemical mixtures EPA/630/R-00/002.

Walton K, Dorne JL, Renwick AG. 2004. Species-specific uncertainty factors for compounds eliminated principally by renal excretion in humans. Food Chem Toxicol 42(2): 261-274.

Van den Berg M, Birnbaum LS, Denison M, De Vito M, Farland W, Feeley M, Fiedler H, Hakansson H, Hanberg A, Haws L, Rose M, Safe S, Schrenk D, Tohyama C, Tritscher A, Tuomisto J, Tysklind M, Walker N, Peterson RE. 2006. The 2005 World Health Organization reevaluation of human and Mammalian toxic equivalency factors for dioxins and dioxin-like compounds. Toxicol Sci 93(2): 223-241.

van der Voet H, Slob W. 2007. Integration of probabilistic exposure assessment and probabilistic hazard characterization. Risk Anal 27(2): 351-371.

van der Voet H, van der Heijden GW, Bos PM, Bosgra S, Boon PE, Muri SD, Bruschweiler BJ. 2009. A model for probabilistic health impact assessment of exposure to food chemicals. Food Chem Toxicol 47(12): 2926-2940.

Vermeire T, Stevenson H, Peiters MN, Rennen M, Slob W, Hakkert BC. 1999. Assessment factors for human health risk assessment: a discussion paper. Crit Rev Toxicol 29(5): 439-490.

Wheeler MW, Bailer AJ. 2007. Properties of model-averaged BMDLs: a study of model averaging in dichotomous response risk estimation. Risk Anal 27(3): 659-670.

Wheeler MW, Bailer AJ. 2009. Benchmark dose estimation incorporating multiple data sources. Risk Anal 29(2): 249-256.

WHO. 1999. Principles of the assessment of risks to human health from exposure to chemicals. (Environmental Health Criteria 210, World Health Organization, Geneva).

WHO/IPCS. 2004. Harmonization document no. 1. IPCS risk assessment terminology. World Health Organization, Geneva.

Zhou T, Ross DG, DeVito MJ, Crofton KM. 2001. Effects of short-term in vivo exposure to polybrominated diphenyl ethers on thyroid hormones and hepatic enzyme activities in weanling rats. Toxicol Sci 61(1): 76-82.