4. CHRONIC ORAL REFERENCE DOSE

This section presents U.S. Environmental Protection Agency (EPA)'s response to the National Academy of Sciences (NAS) recommendations that EPA discuss more explicitly the modeling of noncancer endpoints and develop a Reference Dose (RfD) to address noncancer effects associated with oral 2,3,7,8-tetrachlorodibenzo-*p*-dioxin (TCDD) exposures. Section 2 details the selection of the animal bioassays with the lowest TCDD doses associated with the development of adverse noncancer effects and the selection of relevant epidemiologic studies of adverse noncancer health effects. Section 3 discusses the kinetic modeling and estimation of human equivalent daily oral doses that are used in TCDD RfD development in this section. This section discusses the modeling of noncancer health effects data associated with TCDD exposure and the derivation of an RfD. Specifically, Section 4.1 summarizes the NAS comments on TCDD dose-response modeling and EPA's response, including justification of selected noncancer effects and statistical characterization of modeling results. Section 4.2 presents the TCDD dose-response modeling undertaken for identification of candidate points of departure (PODs) for derivation of an RfD. In Section 4.3, EPA derives an RfD for TCDD. Section 4.4 describes the qualitative uncertainties in the RfD. Finally, Section 4.5 presents two separate quantitative analyses of uncertainty for the TCDD RfD. The first focuses on three data sets (from two epidemiologic studies and one animal bioassay) and quantifies the consequences of alternative decisions in the development of PODs based on these studies. The second develops POD estimates for several Seveso cohort studies that did not qualify for consideration for RfD derivation in the study selection process, but could be considered in the context of investigating uncertainty limits for the RfD.

4.1. NAS COMMENTS AND EPA'S RESPONSE ON IDENTIFYING NONCANCER EFFECTS OBSERVED AT LOWEST DOSES

The NAS recommended that EPA identify the noncancer effects associated with low-dose TCDD exposures and discuss its strategy for identifying and selecting PODs for noncancer endpoints, including biological significance of the effects.

> With respect to noncancer end points, the committee notes that EPA does not use a rigorous approach for evaluating evidence from studies... (p. 47 NAS, 2006b)
>
> The Reassessment should describe clearly the following aspects:
>
> 1. The effects seen at the lowest body burdens that are the primary focus for any risk assessment—the "critical effects."
>
> 2. The modeling strategy used for each noncancer effect, paying particular attention to the critical effects, and the selection of a point of comparison based on the biological significance of the effect; if the ED_{01} is retained, then the biological significance of the response should be defined and the precision of the estimate given... (p. 187, NAS, 2006b).

In this document, EPA has developed a strategy for identifying the noncancer data sets and PODs that represent the most sensitive and toxicologically-relevant endpoints for derivation of an RfD for TCDD. EPA began this process by using the animal bioassays and epidemiologic studies that met its study inclusion criteria as sources of these data sets.

For all epidemiologic studies that were identified as suitable for further quantitative dose-response analyses in Section 2.4.1, EPA has chosen to use NOAELs and LOAELs to identify PODs; benchmark dose (BMD) modeling was not feasible given the nature of the data presented in these studies. Figure 4-1 shows EPA's process for determination of PODs from these key epidemiologic studies. EPA first evaluated the dose-response information in the study to determine whether it provided an estimate of TCDD exposure and an observed health outcome that was toxicologically relevant[1] for RfD derivation. If such data were available, EPA identified a NOAEL or LOAEL as a POD. For each of these, EPA applied a toxicokinetic model to estimate the continuous oral daily intake associated with the POD that could be used in the derivation of an RfD (see Section 4.2). If all of this information was available, the result was included as a POD for derivation of a candidate RfD.

Figures 4-2 and 4-3 together present the strategy EPA used to evaluate the study/endpoint combinations found in the animal bioassays that met EPA's study inclusion criteria, estimate PODs, and develop a final set of candidate RfDs for TCDD. Figure 4-2 summarizes the

[1] RfDs are based on health endpoints that are inherently adverse or clearly linked to downstream functional or pathological alterations (U.S. EPA, 2002).

disposition of the 78 animal noncancer studies selected for TCDD dose-response analyses. Of these studies, 16 were eliminated because EPA determined that they contained no toxicologically-relevant endpoints that could be used to derive a candidate RfD (discussed further in Section 4.2.1). EPA then identified PODs from the remaining bioassays; at this point, Figure 4-2 refers to Figure 4-3, which is a flow chart of the iterative process used to estimate PODs and compare them within and across the remaining studies to arrive at a final set of PODs from these bioassays (see additional details below). From this final set of PODs, Figure 4-2 shows that EPA then eliminated 13 studies from further analysis with both a human equivalent dose (HED) $LOAEL_{HED}$ >1 ng/kg-day and a $NOAEL_{HED}$ or $BMDL_{HED}$ >0.32 ng/kg-day (see Table 4-3); one additional study was also not carried forward because of the lack of toxicokinetic information for estimation of an HED. These dose limits were imposed to limit the size of the analysis yet ensure representation of all important health effects associated with TCDD exposure. From the final list of 48 studies, EPA derived 37 candidate RfDs, with 11 studies presented as supporting information.

Figure 4-3 summarizes the strategy employed for identifying and estimating PODs from the 62 animal bioassays with at least one toxicologically relevant endpoint for RfD derivation. For the noncancer endpoints within these studies, EPA first evaluated the toxicological relevance of each endpoint, rejecting those judged not to be relevant for RfD derivation. Next, initial PODs based on the first-order body burden metric (see Section 3.3.4.2) and expressed as HEDs (i.e., $NOAEL_{HED}$, $LOAEL_{HED}$, $BMDL_{HED}$) were determined for all relevant endpoints (summarized in Table 4-3). Because there were very few NOAELs and BMDL modeling was largely unsuccessful due to data limitations (see Section 4.2), the next stage of evaluation was carried out using LOAELs only. Within each study, effects not observed at the LOAEL (i.e., reported at higher doses) with $BMDL_{HED}$s greater than the $LOAEL_{HED}$ were eliminated from further analysis, as they would not be considered as candidates for the final POD on either a BMDL or NOAEL/LOAEL basis (i.e., the POD would be higher than the PODs of other relevant endpoints). In addition, all endpoints with $LOAEL_{HED}$ estimates beyond a 100-fold range of the lowest identified $LOAEL_{HED}$ across all studies were (temporarily) eliminated from further consideration, as they would not be POD candidates either (i.e., the POD would be higher than the PODs of other relevant endpoints). For the remaining endpoints, EPA then determined final potential PODs based on TCDD whole-blood concentrations obtained from the Emond rodent

PBPK models. HEDs were then estimated for each of these PODs using the Emond human PBPK model. At this point, if the PBPK modeling results suggested considering additional endpoints at higher doses, the process was repeated. From the final set of HEDs, a POD was selected[2] for each study, to which appropriate uncertainty factors (UFs) were applied following EPA guidance (see Section 4.3.3 following). The resulting candidate RfDs were then considered in the final selection process for the RfD. Other endpoints occurring at slightly higher doses representing additional effects associated with TCDD exposure (beyond the 100-fold $LOAEL_{HED}$ range) were evaluated, modeled, and included in the final candidate RfD array[3] to examine endpoints not evaluated by studies with lower PODs. In addition, Benchmark Dose (BMD) modeling based on administered dose was performed on all endpoints for comparison purposes. The final array of selected endpoints is shown in Table 4-4 (summary of BMD analysis) and Table 4-5 (candidate RfDs).

The NAS recommended that EPA better justify the selection of response levels for endpoints used to develop risk estimates. The NAS commented on EPA's decision to estimate an ED_{01} (effective dose eliciting a 1% response) for noncancer bioassay/data set combinations as a comparative tool across studies, suggesting that EPA identify and evaluate the levels of change associated with adverse effects to define the benchmark response (BMR) level for continuous noncancer endpoints.

> The committee notes that the choice of the 1% response level as the POD substantially affects … the noncancer analyses…. The committee recommends that the Reassessment use levels of change that represent clinical adverse effects to define the BMR level for noncancer continuous end points as the basis for an appropriate POD in the assessment of noncancer effects (p. 72, NAS, 2006b).
>
> The committee concludes that EPA did not adequately justify the use of the 1% response level (the ED_{01}) as the POD for analyzing epidemiological or animal bioassay data for … noncancer effects (p. 18, NAS, 2006b).

[2] In the standard order of consideration: BMDL, NOAEL, and LOAEL.
[3] However, studies with a lowest dose tested greater than 30 ng/kg-day were not included in the expanded evaluation.

In the 2003 Reassessment (U.S. EPA, 2003), EPA was not attempting to derive an RfD when it conducted TCDD dose-response modeling. The 2003 Reassessment developed ED_{01} estimates for noncancer effects in an attempt to compare disparate endpoints on a consistent response scale. Importantly, the 2003 Reassessment defined the ED_{01} as 1% of the maximal response for a given endpoint, not as a 1% change from control. Because RfD derivation is the primary goal of noncancer health effects assessment in this document, the noncancer modeling effort undertaken here differs substantially from the modeling in the 2003 Reassessment.

The NAS committee was concerned with the statistical power to determine the shape of the dose-response curve at doses far below observed dose-response information. EPA agrees that the shape of the dose-response curve in the low-dose region cannot be determined confidently when based on higher-dose information. An observed response above background near (or below) the BMR level is needed for discrimination of the shape of the curve and for accurate estimation of an ED_x or BMDL. Although many of the ED_{01}s presented in the 2003 Reassessment were near the lowest dose tested, responses at the lowest doses were often high and much greater than a 1% response (i.e., 1% of the maximum response). The lack of an observed response near the BMR level is often a problem in interpretation of BMD modeling results.

In this document, EPA has used a 10% BMR for dichotomous data for all endpoints; there were no developmental studies that accounted for litter effects, for which a 5% BMR would be used (U.S. EPA, 2000). For continuous endpoints in this document, EPA has used a BMR of 1 standard deviation from the control mean whenever a specific toxicologically-relevant BMR could not be defined. For the vast majority of continuous endpoints, EPA could not establish unambiguous levels of change representative of adversity, which EPA defines as "a biochemical change, functional impairment, or pathologic lesion that affects the performance of the whole organism, or reduces an organism's ability to respond to an additional environmental challenge" (U.S. EPA, 2009a). For body and organ weight change, EPA has previously established a BMR of 10% change, which also is used in this document.

The NAS commented on EPA's development of ED_{01} estimates for numerous study/data set combinations in the 2003 Reassessment, suggesting that EPA had not appropriately characterized the statistical confidence around such model predictions in the low-response region of the model.

> It is critical that the model used for determining a POD fits the data well, especially at the lower end of the observed responses. Whenever feasible, mechanistic and statistical information should be used to estimate the shape of the dose-response curve at lower doses. At a minimum, EPA should use rigorous statistical methods to assess model fit and to control and reduce the uncertainty of the POD caused by a poorly fitted model. The overall quality of the study design is also a critical element in deciding which data sets to use for quantitative modeling (NAS, 2006b, p. 18).
>
> EPA should … assess goodness-of-fit of dose-response models for data sets and provide both upper and lower bounds on central estimates for all statistical estimates. When quantitation is not possible, EPA should clearly state it and explain what would be required to achieve quantitation (NAS, 2006b, p. 10).

The NAS also commented that EPA report information describing the adequacy of dose-response model fits, particularly in the low response region. For those cases where biostatistical modeling was not possible, NAS recommended that EPA identify the reasons.

> The Reassessment should also explicitly address the importance of statistical assessment of model fit at the lower end and the difficulties in such assessments, particularly when using summary data from the literature instead of the raw data, although estimates of the impacts of different choices of models would provide valuable information about the role of this uncertainty in driving the risk estimates (NAS, 2006b, p. 73).

To address this concern, in this document EPA has reported the standard suite of goodness-of-fit measures from the benchmark dose modeling software (BMDS 2.1). These include chi-square *p*-values, Akaike's Information Criterion (AIC), scaled residuals at each dose level, and plots of the fitted models. For the multistage model, when restricted lower-order coefficients hit the lower bound (zero), EPA used likelihood ratio tests to evaluate whether the improvement in fit afforded by estimating successively higher-order coefficients could be justified. Goodness-of-fit measures are reported for all key data sets in Appendix G. (Section 4.2.4.2 discusses the BMD modeling criteria for model evaluation.)

4.2. NONCANCER DOSE-RESPONSE ASSESSMENT OF TCDD

This section describes EPA's evaluation of TCDD dose response for noncancer endpoints from studies that met the study inclusion criteria. Discussions include BMD modeling procedures, kinetic modeling, and development of PODs for derivation of the RfD. Section 4.2.1 discusses the types of endpoints that are considered relevant by EPA for derivation of toxicity values (U.S. EPA, 2005a, b, 1998, 1996, 1994, 1991) and lists the study/endpoint combinations that were not considered for the TCDD RfD derivation, with supporting text in Appendix H. Section 4.2.2 describes how EPA has used PBPK modeling to estimate effective internal exposures as an alternative to using administered doses or body burdens based on first-order kinetics. Section 4.2.3 details the dose-response analysis of the epidemiologic data, with supporting information on kinetic modeling in Appendix F. Section 4.2.4 details the dose-response analysis for the animal bioassay data; Appendix G provides the BMDS input tables (Section G.1) and output for all modeling, including blood concentrations (Section G.2) and administered dose (Section G.3).

4.2.1. Determination of Toxicologically Relevant Endpoints

The NAS committee commented on the low-dose model predictions and the need to discuss the biological significance of the noncancer health effects modeled in the 2003 Reassessment. In selecting POD candidates from the animal bioassays for derivation of the candidate RfDs, EPA considered the toxicological relevance of the identified endpoint(s) from any given study. Some endpoints/effects may be sensitive, but lack general toxicological significance because of lack of inherent adversity[4], being an adaptive response, or not being clearly linked to downstream functional or pathological alterations. Endpoints not considered to be toxicologically relevant for TCDD include Cytochrome P-450 (CYP) induction, oxidative stress measures, mRNA induction, protein phosphorylation, certain immune system responses, gap junction disruption, and epidermal growth factor signaling. As an example, CYP induction alone is not considered a significant toxicological effect given that CYPs are induced as part of the normal hepatic metabolism of xenobiotic agents. Additionally, the role of CYP induction in

[4] An adverse effect is defined in EPA's Integrated Risk Information System glossary as "a biochemical change, functional impairment, or pathologic lesion that affects the performance of the whole organism, or reduces an organism's ability to respond to an additional environmental challenge" (U.S. EPA, 2009a).

the noncancer toxicity of TCDD is unknown, thus, due to the lack of obvious pathological significance, TCDD-induced CYP induction is not considered a relevant endpoint for RfD derivation. Another example is oxidative stress. As an example, TCDD has been shown to induce changes in oxidative stress markers, but no other indicators of brain pathology were assessed (Hassoun et al., 2003; Hassoun et al., 2000; Hassoun et al., 1998). In this case, it is impracticable to link the markers of oxidative stress to a toxicological outcome in the brain; thus, this endpoint is not considered relevant for RfD derivation. Studies otherwise meeting the study inclusion criteria, but with no toxicologically-relevant endpoints that were considered suitable for derivation of a candidate RfD are described and discussed in Appendix H.

4.2.2. Use of Toxicokinetic Modeling for TCDD Dose-Response Assessment

Because relevant toxicokinetic models for TCDD disposition in rodents and humans are available, EPA has not applied the standard uncertainty factor approach in the derivation of the TCDD RfD. In addition, because of the much slower elimination of TCDD in rodents than in humans, EPA has determined that the standard uncertainty factor approach can underestimate the interspecies toxicokinetic extrapolation factor by an order of magnitude or more (U.S. EPA, 2003). The toxicokinetic models chosen by EPA are the rodent and human PBPK models described by Emond et al. (2006; 2005; 2004)[5] and modified by EPA for this assessment as described in Section 3.3.4 (hereafter referred to as the "Emond [rodent or human] PBPK model"). Both the rodent and human models have a gestational component, which allow for more relevant exposure comparisons between general adult exposures and the numerous gestational exposure studies. Ideally, a relevant tissue concentration for each effect would be estimated. However, at present, no models exist for estimation of all relevant tissue concentrations. As virtually all TCDD is found in the adipose fraction of tissues, or bound to specific proteins, a preferred approach to developing a dose metric would be to account for the fat fraction of each tissue and protein binding; however, EPA has decided that the modeling of such estimates is too uncertain, and EPA has not found sufficient data to implement this approach. Therefore, EPA has decided to use the concentration of TCDD in whole blood as a surrogate for tissue concentrations, assuming that tissue concentrations are proportional to

[5]The Emond PBPK models are three-compartment dynamic models.

whole-blood concentrations. Furthermore, because the RfD is necessarily expressed in terms of average daily exposure, the blood concentrations are expressed as averages over the relevant period of exposure for each endpoint. For the animal bioassays, the relevant period of exposure is the duration of dosing, starting at the age of the animals at the beginning of the study. For humans, the relevant period of exposure is generally a lifetime, which is defined as 70 years. However, EPA varied the averaging time for the equivalent human blood concentrations to correspond to the test-animal exposure duration in the following manner.

- For correspondence with animal chronic exposures,[6] the human-equivalent TCDD blood concentration is assumed to be the 70-year average.
- For correspondence with animal gestational exposures, the human-equivalent TCDD blood concentration is assumed to be the average over 45 years for a female, beginning at birth, plus 9 months of gestational exposure.[7] Forty five years of age is considered here as an upper limit on the age at which a typical human female can conceive and bear a child.
- For correspondence with any other animal exposure duration, the human-equivalent TCDD blood concentration is assumed to be the average over the equivalent human exposure duration calculated backward from the peak exposure plateau at or near the end of the 70-year scenario. The average is determined from the terminal end of the human exposure period to be protective of less-than-lifetime exposures occurring at any time in a lifetime; the daily oral intake achieving the target blood concentration is smaller than for the same exposure period beginning at birth. The determination of equivalent exposure durations across species is problematic and somewhat arbitrary, so EPA uses the average peak blood concentration as the human equivalent for all less-than-chronic animal exposures (other than gestational).[8] For the first-order kinetics model, the average peak exposure is close to the theoretical steady-state asymptote (see Section 3.3.4.2). However, for the Emond human PBPK model used by EPA in this assessment, the timing of the peak exposure is dose-dependent and tends to decline after 60 years in some cases. Therefore, the 5-year average TCDD blood concentration that includes the peak ("5-year peak") is used as the relevant dose-metric for the PBPK model applications (see Section 3.3.6 and Figure 3-33).

---

[6]Assumed to be ≥75% of nominal lifetime, or about 550 days in rodents.

[7]See Section 3.3.4.2 for a discussion of this issue, including a comparison of the 45-year old pregnancy scenario to one beginning at age 25 in Table 3-24.

[8]By comparison to a half-lifetime equivalent (1 year in rodents, 35 years in humans), in the 1st-order kinetic model the ratio of body burden to oral intake does not differ significantly from the average-peak scenario; all shorter-term scenarios differ even less (see Section 3.3.4.2). These relationships, with respect to the 5-year peak, hold for the PBPK model results, as well (see Section 3).

4.2.3. Noncancer Dose-Response Assessment of Epidemiological Data

The following four epidemiologic studies describing noncancer endpoints were identified in Section 2.4.1 as studies to be evaluated for development of PODs for derivation of candidate RfDs: Baccarelli et al. (2008), Mocarelli et al. (2008), Alaluusua et al. (2004), and Eskenazi et al. (2002b). Each of these studies described effects observed in the Seveso cohort (see detailed study summaries in Appendix C and Table 2-2). Each study reported individual-level human exposure measures and provided information from which EPA could determine a critical exposure window of susceptibility over which the effective TCDD exposures could be quantified for dose-response assessment. For studies that reported grouped data by TCDD exposure ranges, the representative values for the ranges were determined by taking the geometric mean of the range limits, assuming that the TCDD concentration distribution in the population is more likely to be skewed (e.g., lognormal) than symmetrical (e.g., normal or uniform). A sufficient number of significant digits are carried through intermediate results to avoid round-off error in the final value. EPA used toxicokinetic modeling (Emond human PBPK model) to estimate daily TCDD intake rates for the exposure groups presented in these studies (see Appendix F for details). The exposure scenario in all of these studies, except Baccarelli et al. (2008), entailed an initial high pulse exposure at the time of the plant explosion followed by low-level background exposure over a period of several years across the critical exposure window, resulting in internal exposure profiles characterized by a 5 to 10-fold difference in initial and final TCDD serum concentrations (as lipid adjusted serum concentrations [LASC]). For these scenarios, EPA modeled both the peak TCDD LASC and the average LASC over the critical window, then estimated daily average continuous TCDD intakes over the critical-window duration corresponding to each of the peak and critical-window average serum concentrations. Estimation of LASC and intakes was accomplished using the Emond human PBPK model. EPA considered the critical-window average exposures to be important, although they were much lower than the peak exposures, because the relatively slow elimination of TCDD engenders concerns for an ongoing contribution of residual TCDD body burdens to the adverse health outcomes during the period of susceptibility. However, the overall average exposure does not reflect the influence of the much

higher peak exposure, which may be a significant factor in TCDD toxicity (Kim et al., 2003).[9] That is, EPA is uncertain as to whether the health outcomes, often observed many years beyond the period of susceptibility, are a result of permanent damage from the initial high exposure or more gradual impairment from longer-term ongoing exposure. For these reasons, EPA derived the PODs for RfD consideration by averaging the TCDD intakes for the peak exposure and critical-window exposure average, essentially treating each as equally likely. EPA focused on identifying NOAELs and LOAELs for these studies. EPA did not conduct BMD modeling because the covariates identified by the study authors could not be incorporated by modeling the grouped response data. EPA's development of PODs for these studies is described in this section and Appendix F; the results are shown in Table 4-1.

4.2.3.1. *Baccarelli et al. (2008)*

For Baccarelli et al. (2008), EPA was able to define a LOAEL in terms of the maternal TCDD serum levels corresponding to neonatal thyroid stimulating hormone (TSH) level above 5 µ-Units TSH per mL of serum (5 µU/mL) (See Appendix C, Section C.1.2.1.5.7, and Table 2-2 for study details). The adversity benchmark of 5 µU/mL is based on the WHO (1994) indicator for follow up examination for potential hypothyroidism (see following discussion in Section 4.3.4.1). Baccarelli et al. (2008) performed regression modeling of neonatal TSH against maternal TCDD LASC but did not estimate the equivalent oral intake. The regression model related the level of TSH in 3-day-old neonates to TCDD concentrations in maternal plasma at birth (given as LASC). The authors extrapolated maternal plasma concentrations from previous measurements using a simple first-order pharmacokinetic model. The study authors also reported group average neonatal TCDD serum levels for infants above and below the 5 µU/mL limit. However, because there is limited information regarding the relationship between maternal and neonatal serum TCDD levels, EPA determined that there was too much uncertainty in estimating maternal intake from neonatal TCDD serum concentrations directly. Therefore, EPA determined the maternal intake at the LOAEL from the maternal serum-TCDD/TSH regression model by finding the maternal TCDD LASC at which neonatal TSH exceeded 5 µU/mL. EPA then used the Emond PBPK model under the human gestational scenario (see

[9] Kim et al. (2003) found a significantly higher fraction of altered hepatic foci in rats treated with a single high TCDD dose than those administered a continuous dose over 15 weeks, yielding similar terminal liver TCDD concentrations.

Section 4.2.2) to estimate the continuous daily oral TCDD intake that would result in a TCDD LASC corresponding to a neonatal TSH of 5 µU/mL at the end of gestation; EPA established the resulting maternal intake (0.020 ng/kg-day) as the LOAEL, shown in Table 4-1 as a POD for derivation of candidate RfDs (PBPK modeling details are shown in Appendix F).

4.2.3.2. *Mocarelli et al. (2008)*

Mocarelli et al. (2008) reported decreased sperm concentrations (20%) and decreased motile sperm counts (11%) in men who were 1−9 years old in 1976 at the time of the accident (initial TCDD exposure event) (see Appendix C, Section C.1.2.1.5.8, and Table 2-2 for study details). Men who were 10−17 years old in 1976 were not adversely affected. Serum (LASC) TCDD levels were measured within 1 year of the initial exposure. Serum TCDD levels and corresponding responses were reported by quartile, with a reference group of less-exposed individuals assigned a TCDD LASC value of 15 ppt (which was the mean of individuals outside the contaminated area). The lowest exposed group mean was 68 ppt (1st quartile). Because effects were detected only among boys under the age of 10, EPA assumes there is a maximum 10-year critical exposure window for elicitation of these effects. However, for the exposure profile, with a high initial pulse followed by an extended period of elimination with only background exposure, the estimation of an average exposure resulting in the effect is somewhat problematic. Therefore, EPA implemented a procedure for the estimation of the continuous daily TCDD intake associated with the LOAEL in the Mocarelli et al. (2008) study using the following 5-step process:

1. Using the Emond human PBPK model, the initial (peak) serum TCDD concentrations (LASC) associated with the accident were back-calculated based on the time that had elapsed between the explosion and the serum collection. As serum measurements were taken within 1 year after the event, a lag time to measurement of 0.5 years was assumed. The group average peak serum concentration for the 1st quartile was estimated to be 249 ppt.
2. The oral exposure associated with the peak serum TCDD concentration (peak exposure) was calculated using the Emond PBPK model.
3. Starting with the peak exposure and accounting for background TCDD intake, the average daily serum TCDD concentration experienced by a representative individual in the susceptible lifestage (boys under 10 years old) was estimated using the Emond PBPK model. The average subject age at the time of the event was 6.2 years. Consequently, a

critical exposure window for the cohort was estimated to be, on average, 3.8 years (i.e., a boy aged 6.2 years would remain in this exposure window for 3.8 more years until he was 10 years of age). The critical window average serum concentration for the 1st quartile group was estimated to be 57.7 ppt (45 ppt at 10 years).

4. Using the Emond PBPK model, the average daily TCDD intake rate needed to attain the 3.8-year average serum TCDD concentration in a boy 10 years old was calculated.

5. The LOAEL POD was calculated as the average of the peak exposure intake (0.032 ng/kg-day) and the 3.8-year average exposure intake (0.0080 ng/kg-day), resulting in LOAEL of 0.020 ng/kg-day, shown in Table 4-1 as a POD for derivation of a candidate RfD.

The PBPK modeling details are shown in Appendix F.

4.2.3.3. *Alaluusua et al. (2004)*

For Alaluusua et al. (2004), the approach for estimation of daily oral TCDD intake is virtually identical to the approach used for the Mocarelli et al. (2008) data. (See Appendix C, Section C.1.2.1.5.5, and Table 2-2 for study details.) Alaluusua et al. (2004) reported dental effects in male and female adults who were less than 5 years of age at the time of the initial exposure (1976). For the 75 boys and girls who were less than 5 years old at the time of the accident, 25 (33%) were subsequently diagnosed with some form of dental enamel defect. For the 38 individuals who were older than 5, only 2 (5.3%) suffered dental enamel defects at a later date. In addition, the incidence of missing permanent teeth (lateral incisors and second premolars) was 3 times as prevalent in zone ABR subjects compared with zone non-ABR residents. A window of susceptibility of approximately 5 years is assumed. Serum measurements for this cohort were taken within a year of the accident. Serum TCDD levels and corresponding responses were reported by tertile, with a reference group of less-exposed individuals assigned a TCDD LASC value of 15 ppt (ng/kg); the tertile group geometric means were 72.1, 365.4, and 4,266 ppt. The incidence of dental effects for the reference group was 26% (10/39). The incidence of dental effects in the 1st, 2nd, and 3rd tertile exposure groups was 10% (1/10), 45% (5/11), and 60% (9/15), respectively. EPA judged that the NOAEL and LOAEL were 72.1 and 365.4 ppt TCDD in serum (LASC), in the 1st tertile and 2nd tertile, respectively. Following the same procedure used for the Mocarelli et al. (2008) study (see Section 4.2.3.2), EPA estimated the continuous daily human oral TCDD intake associated with each of the tertiles for both peak and average exposure across the critical exposure window,

assuming that the average age of the susceptible cohort at the time of the accident was 2.5 years. Separate estimates for boys and girls were developed based on both the peak intake and average intake across the critical exposure window (PBPK modeling details are shown in Appendix F). The estimated averaged daily oral intakes for the tertiles, averaged for boys and girls, are 0.0655, 1.65, and 111 ng/kg-day for the peak exposure and 0.0156, 0.149, and 4.81 ng/kg-day for the critical exposure window average. The LOAEL for this study was determined to be 0.897 ng/kg-day, which is the average of the peak exposure and window average exposure for the second tertile. A study NOAEL of 0.0406 ng/kg-day for the first tertile was determined similarly and serves as a POD for derivation of a candidate RfD in Table 4-1.

4.2.3.4. *Eskenazi et al. (2002b)*

The approach used to estimate daily TCDD intake in Eskenazi et al. (2002b) combines the approaches EPA used for Baccarelli et al. (2008), Mocarelli et al. (2008), and Alaluusua et al. (2004). Eskenazi et al. (2002b) reported menstrual effects in female adults who were premenarcheal in 1976 at the time of the initial exposure (see Appendix C, Section C.1.2.1.4.1 and Table 2-2 for study details). In Rigon et al. (2010) , the median age at menarche was shown to be 12.4 in Italian females with intergenerational decreases in age at menarche. Thus, EPA established a window of susceptibility of approximately 13 years for this analysis. The average age of the premenarcheal girls at the time of the initial exposure in 1976 was 6.8 years, establishing an average critical-window exposure duration of 6.2 years for this cohort. Serum samples were collected within a year of the accident from this cohort. However, serum TCDD levels and corresponding responses were not reported by percentile, and no internal reference group was identified. As for Baccarelli et al. (2008), Eskenazi et al. (2002b) developed a regression model relating menstrual cycle length to plasma TCDD concentrations (LASC) measured in 1976. The model estimated that menstrual cycle length was increased 0.93 days for each 10-fold increase in TCDD LASC, with a 95% confidence interval of −0.01 to 1.86 days. The determination of a LOAEL is somewhat arbitrary, with no independent measure of an adversity threshold to establish the toxicological significance of a given increase in menstrual cycle length. The study authors did not present data for unexposed premenarcheal girls (in 1976), so an appropriate reference population is not available. EPA did not conduct BMD modeling because of the lack of a reference population and the inability to include the covariates

considered by the study authors in their analysis. However, an approximate LOAEL can be estimated from Figure 1 in Eskenazi et al. (2002b), noting that both the length of the menstrual cycle and its variance increases above TCDD concentrations of about 1,000 ppt. The highest measured concentration is 16,500 ppt. Consistent with the previously established method for determining representative values for age limits (Sections 4.2.3.2 and 4.2.3.3), the geometric mean of 4,060 ppt for this range is assigned as a LOAEL. The lower range of TCDD concentrations is too large to treat as a single group for estimating a NOAEL, but using the study authors' regression model, a TCDD LASC of about 50 corresponds to a menstrual cycle length of 28 days, generally considered to be the average normal length. These two (1976) serum levels were then modeled by EPA using the Emond human PBPK model in the same manner as for Mocarelli et al. (2008) and Alaluusua et al. (2004), but with a 6.2-year exposure window for the premenarcheal girls. The resulting peak and window-average TCDD intakes for the 50 ppt exposure are 0.0168 and 0.00364 ng/kg-day, respectively; the average of the two intakes is 0.0102 ng/kg-day. The peak and window-average TCDD intakes for the LOAEL exposure (4,060 ppt) are 60.0 and 1.52 ng/kg-day, respectively; the average of the two intakes of 30.8 ng/kg-day defines the LOAEL POD. Further details of the PBPK modeling can be found in Appendix F. Although 0.0102 ng/kg-day could be considered to be a NOAEL, there is too much uncertainty in the upper end of the NOAEL range, given the very large (3,000-fold) difference between it and the LOAEL, for using it as a NOAEL POD. The LOAEL of 30.8 ng/kg-day, also uncertain in magnitude and toxicological significance, is 1,540-fold higher than the LOAEL PODs for Mocarelli et al. (2008) and Baccarelli et al. (2008), and will not be a factor in the derivation of the RfD. Therefore, the LOAEL for this study is not considered further in this assessment except in the context of the RfD uncertainty analysis presented in Section 4.5.

4.2.4. Noncancer Dose-Response Assessment of Animal Bioassay Data

EPA followed the strategy illustrated in Figures 4-2 and 4-3 to evaluate the animal bioassay data for TCDD dose response. For the administered average daily doses (ng/kg-day) in each animal bioassay, EPA identified NOAELs and/or LOAELs based on the original data presented by the study author. Section 2.4.2 identifies these values in Table 2-4 and in the study summaries found in Appendix D. These became PODs for consideration in the derivation of an RfD for TCDD. The candidate RfD values associated with these PODs are presented in

Table 4-5. All PODs were converted to HEDs using the Emond PBPK models, with whole-blood TCDD concentration as the effective dose metric. The remainder of this section describes the steps in this process and concludes with the PODs from the animal bioassay data that were considered for derivation of the RfD.

4.2.4.1. *Use of Kinetic Modeling for Animal Bioassay Data*

Whole-blood TCDD concentrations corresponding to the administered doses in each mouse or rat bioassay qualifying as a final RfD POD were estimated using the appropriate Emond rodent PBPK model. In each case, the simulation was performed using the exposure durations, body weights, and average daily doses from the original studies. For all multiple-exposure protocols, the time-weighted average blood TCDD concentrations over the exposure period were used as the relevant dose metric. For single (gestational and nongestational) exposures, the initial peak blood TCDD concentrations were considered to be the most relevant exposure metric. Gestational exposures were modeled using the species-specific gestational component of the Emond rodent PBPK model. Bioassays employing exposure protocols spanning gestational and postpartum life stages were modeled by sequential application of the gestational and nongestational models.

The Emond PBPK models do not contain a lactation component, so exposure during lactation was not modeled explicitly. Only one bioassay (Shi et al., 2007) considered as a POD for RfD derivation included exposure during lactation. In Shi et al. (2007), pregnant animals were exposed weekly to TCDD throughout gestation and lactation. Exposure was continued in the offspring following weaning for 10 months. For assessment of maternal effects, the Emond gestational model was used, terminating at parturition. For assessment of long-term exposure in the offspring, the Emond nongestational model was used, ignoring prior gestational and lactational exposure, with the assumption that the total exposure during these periods was small relative to exposure in the following 10 months. The assumption is conservative in that effects observed in the offspring would be attributed entirely to adult exposure, which is somewhat less than the actual total exposure.

The model code, input files, and PBPK modeling results for each bioassay are reported in Appendix E. The modeled TCDD blood concentrations were used for BMD modeling of bioassay response data and determination of NOAELs and LOAELs. BMD modeling was

performed, as described in Section 3.5.2.2.1, by substituting the modeled blood concentrations for the administered doses and calculating the corresponding BMDL. For each of these LOAEL, NOAEL, or BMDL blood-concentration equivalents, corresponding HEDs were estimated using the Emond human PBPK model for the appropriate gestational or nongestational scenario as described previously (see Section 4.2.2).

4.2.4.2. *Benchmark Dose Modeling of the Animal Bioassay Data*

BMD modeling was performed for each study/endpoint combination using BMDS 2.1 to determine BMDs and BMDLs. The input data tables for these noncancer studies are shown in Appendix G, Section G.1, including both administered doses (ng/kg-day) and blood concentrations (ng/kg [ppt]) and either incidence data for the dichotomous endpoints or mean and standard deviations for the continuous endpoints (See Section 4.2.4.1 and Sections 3.3.4 and 3.3.5 for a description of the development of TCDD blood concentrations using kinetic modeling).

Evaluation of BMD modeling performance, goodness-of-fit, dose-response data, and resulting BMD and BMDL estimates included statistical criteria as well as professional judgment of their statistical and toxicological properties. For the continuous endpoints, all available models were run separately using both the assumption of constant variance and the assumption of modeled variance. Saturated (0 degrees of freedom) model fits were rejected from consideration. Parameters in models with power or slope parameters were constrained to prevent supralinear fits, which EPA considers not to be biologically plausible and which often have undesirable statistical properties (i.e., the BMDL converges on zero). Table 4-2 shows each model and any restrictions imposed.

For the quantal/dichotomous endpoints, all primary BMDS dichotomous models were run. The alternative dichotomous models were fit to several data sets, but the results were very sensitive to the assumed independent background response and the fits were not accepted. The confidence level was set to 95%, and all initial parameter values were set to their defaults in BMDS. For the continuous endpoints, 1 standard deviation was chosen as the default for the

BMR when a specific toxicologically-relevant BMR could not be defined. For the dichotomous endpoints, a BMR of 10% extra risk was used for all endpoints.[10]

The model output tables in Appendix G show all of the models that were run, both restricted and unrestricted, goodness-of-fit statistics, BMD and BMDL estimates, and whether bounds were hit for constrained parameters. After all models were run, the one giving the best fit was selected using the selection criteria in the draft BMD Technical Guidance (U.S. EPA, 2000). Acceptable model fits were those with chi-square goodness-of-fit *p*-values greater than 0.1. For continuous endpoints, the preference was for models with an asymptote term (plateau for high-dose-response) because continuous measures do not continue to rise (or fall) with dose forever; this phenomenon is particularly evident for TCDD. Unbounded models, such as the power model, must account for the plateauing effect entirely in the shape parameter, generally resulting in a supralinear fit. Also, for the continuous endpoints, the *p*-value for the homogenous variance test (Test 2) was used to determine whether constant variance ($p > 0.1$) or modeled variance ($p < 0.1$) should be used. As BMDS offers only one variance model, model fits for modeled variance models were not necessarily rejected if the variance model did not fit well (Test 3 *p*-value < 0.05). Within the group of models with acceptable fits, the selected model was generally the one with the lowest AIC. If the AICs were similar, the model with the lowest BMDL was selected. However, particularly for continuous models, the fit of the model to the control-group response and in the lower response range was assessed. Models with higher BMDLs or AICs but much better fit to the lower response data were often chosen over the nominally best-fitting model.

For many data sets, no models satisfied the acceptance criteria, and no clear BMD/BMDL selection could be made. In these cases, model fits were examined on an individual basis to determine the reasons for the poor fits. On occasion, high doses were dropped, and the models were refit. Also, if a poor fit to the control mean was evident, the model was refit to the data after fixing the control mean by specifying the relevant parameter in BMDS. However, these techniques rarely resulted in better fits. If the fit was still not acceptable, the NOAEL/LOAEL approach was applied to the study/data set combination. Most of the problems with BMD modeling were a consequence of lack of response data near the

[10]There were no developmental studies that accounted for litter effects, for which a 5% BMR would be used.

BMR; many of the TCDD data sets failed to show a response near the BMR, whether it was a 10% dichotomous relative change or a continuous 1 standard deviation change. Responses at the lowest doses were generally much higher than the BMR, resulting in a lack of "anchoring" at the critical response levels of interest, resulting in insufficient information for precise numerical estimation of BMDLs.

4.2.4.3. *PODs from Animal Bioassays Based on HED and BMD Modeling Results*

Table 4-3 summarizes the PODs that EPA estimated for each key animal study included for TCDD noncancer dose-response modeling that also contained toxicologically relevant endpoints (see Section 4.2.1 and Appendix H for excluded studies). After estimating the blood TCDD concentration associated with a particular toxicity measure (NOAEL, LOAEL, or BMDL) obtained from a rodent bioassay, EPA estimated a corresponding HED using the Emond human PBPK model (described in Section 3). Table 4-3 summarizes the NOAEL, LOAEL, or BMDL based on the administered animal doses for each key bioassay/data set combination. Table 4-3 also summarizes the continuous daily HED corresponding to these administered doses as 1st order body burdens and as whole-blood concentrations. The doses in Table 4-3 are defined as follows, all in units of ng/kg-day:

- Administered Dose NOAEL: Average daily dose defining the NOAEL for the test species in the animal bioassay
- Administered Dose LOAEL: Average daily dose defining the LOAEL for the test species in the animal bioassay
- Administered Dose BMDL: BMDL for the test species based on modeling of the administered doses from the animal bioassay
- First-Order Body Burden HED NOAEL: Average daily dose defining the NOAEL for humans derived from the animal bioassay using the first-order kinetics body-burden model
- First-Order Body Burden HED LOAEL: Average daily dose defining the LOAEL for humans derived from the animal bioassay using the first-order kinetics body-burden model
- First-Order Body Burden HED BMDL: Human-equivalent BMDL from BMD modeling of the animal bioassay data using first-order body burdens
- Blood Concentration HED NOAEL: Average daily dose defining the NOAEL for humans derived from the animal bioassay using the Emond human PBPK model

- Blood Concentration HED LOAEL: Average daily dose defining the LOAEL for humans derived from the animal bioassay using the Emond human PBPK model
- Blood Concentration HED BMDL: Human-equivalent BMDL from BMD modeling of the animal bioassay data using the Emond human PBPK model

An evaluation of key BMD analyses is presented in Table 4-4. Tables showing the best model fit for each study/endpoint combination and the associated BMD/BMDL are shown in Appendix G. As described in Section 4.2.4.2, the BMD modeling was largely unsuccessful, primarily because of a lack of response data near the BMR, poor modeled representation of control values, or nonmonotonic responses yielding poor fits. The comments column in Table 4-4 lists reasons for poor results.

4.3. RFD DERIVATION

Table 4-5 lists all the studies and endpoints considered for derivation of the RfD in order of candidate RfD from lowest to highest (The selection process was previously described in Section 4.1). The range of studies includes three of the four human studies.[11] Figure 4-4 (exposure-response array) shows all of the endpoints listed in Table 4-5 graphically in terms of PODs in human-equivalent intake units (ng/kg-day). The human study endpoints are shown at the far left of the figure, and the animal bioassay endpoints are arranged by category to the right. Figure 4-5 demonstrates the same endpoints, arrayed by RfD value, showing the POD, applicable UFs, and candidate RfD.

Table 4-5 illustrates the study, species, strain and sex, study protocol, and toxicological endpoints observed at the lowest TCDD doses. The table also identifies the human-equivalent BMDLs (when applicable), NOAELs, and LOAELs, as well as the composite UF that applies to the specific endpoint and the corresponding candidate RfD.[12] The NOAELS, LOAELs, and BMDLs are presented as HEDs, based on the assumption that whole-blood concentration is the toxicokinetically equivalent TCDD dose metric across species and serves as a surrogate for tissue concentration.[13] For rats and mice, these estimates relied on the two Emond PBPK

[11]The RfD derived from the study of Eskenazi et al. (2002b) was outside the RfD range presented in Table 4-5.
[12]Extra digits are retained for transparency and comparison prior to rounding to one significant digit for the final RfD.
[13]The procedures for estimating HEDs based on TCDD blood concentration are described in the preceding section.

models—one for the relevant rodent species and one for the human—as described previously (see Section 3.3.4.3). The guinea pig and monkey studies that are included in Table 4-5 are given in HED units based on the first-order body burden model (described in Section 3.3.4.2) because there are no published PBPK models to estimate TCDD disposition in guinea pigs and monkeys. The values listed for guinea pigs and monkeys are not directly comparable to those for rats and mice but are probably biased low, as first-order body burden HED estimates for rats and mice are generally 2- to 5-fold lower than the corresponding PBPK model estimates. The LOAELs for the human studies also rely on the Emond PBPK model, as described in Sections 4.2.2 and 4.2.3.

As is evident from Table 4-5, very few NOAELs and even fewer BMDLs have been established for low-dose TCDD studies. BMD modeling was unsuccessful for all of the endpoints without a NOAEL, primarily because of the lack of dose-response data near the BMR (see discussion in Section 4.2). Therefore, the RfD assessment rests largely on evaluation of LOAELs to determine the POD.

4.3.1. Toxicological Endpoints

As can be seen in Table 4-5, a wide array of toxicological endpoints has been observed following TCDD exposure, ranging from subtle developmental effects to overt toxicity. Developmental effects in rodents include embryotoxicity, neonatal mortality, dental defects, delayed puberty in males, and several neurobehavioral effects. Reproductive effects reported in rodents include altered hormone levels in females and decreased sperm production in males. Immunotoxicity endpoints, such as decreased response to SRBC challenge in mice and decreased delayed-type hypersensitivity response in guinea pigs, are also observed. Longer durations of TCDD exposure in rodents are associated with organ and body weight changes, renal toxicity, hepatotoxicity, and lung lesions. Adverse effects in human studies are also observed, which include both male and female reproductive effects, increased TSH in neonates, and dental defects in children. Other outcomes including diabetes (Michalek and Pavuk, 2008) and hepatic effects (Michalek et al., 2001b) have also been associated with adult human TCDD exposures, but EPA was unable to quantify the exposure-response relationship (see Appendix C). All but three of the study/endpoint combinations from animal bioassays listed in Table 4-5 are on TCDD-induced toxicity observed in mice and rats; the other three study/endpoint combinations are effects in

guinea pigs and monkeys. Although the effects of TCDD also have been investigated in hamsters and mink, those studies were not included for final POD consideration because the effect levels were greater than those in Table 4-5, or because effective oral intakes could not be estimated.

Three human studies were also included for final POD consideration in the derivation of an RfD and are presented in Table 4-5 as candidate RfDs. All three human study/endpoint combinations are from studies on the Seveso cohort. The developmental effects observed in these studies were associated with TCDD exposures either in utero or in early childhood between 1 and 10 years of age. Baccarelli et al. (2008) reported increased levels of TSH in newborns exposed to TCDD in utero, indicating a possible dysregulation of thyroid hormone metabolism. Mocarelli et al. (2008) reported decreased sperm concentrations and decreased motile sperm counts in men who were 1−9 years of age in 1976 at the time of the Seveso accident (initial TCDD exposure event). Alaluusua et al. (2004) reported dental effects in adults who were less than 5 years of age at the time of the initial exposure (1976).

4.3.2. Exposure Protocols of PODs

The studies in Table 4-5 represent a wide variety of exposure protocols, involving different methods of administration and exposure patterns across virtually all exposure durations and life stages. Both dietary and gavage administration have been used in rodent studies, with gavage being the predominant method. Gavage dosing protocols vary quite widely and include single gestational exposures, multiple daily exposures (for up to 2 weeks, intermittent schedules that include 5 days/week, once weekly, or once every 2 weeks), and loading/maintenance dose protocols, in which a relatively high dose is initially administered followed by lower weekly doses. The intermittent dosing schedules require dose-averaging over time periods as long as 2 weeks, which introduces uncertainty in the effective exposures. In other words, the high unit dose may be more of a factor in eliciting the effect than the average TCDD tissue levels over time. Although the loading/maintenance dose protocols are designed to maintain a constant internal exposure, these protocols are somewhat inconsistent with the constant daily TCDD dietary exposures associated with human ingestion patterns.

The epidemiologic studies conducted in the Seveso cohort represent exposures over different life stages including gestation, childhood, and young adulthood. The Seveso exposure

profile is essentially a high initial pulse TCDD exposure followed by a 20−30 year period of elimination with only background exposures to TCDD and other dioxin-like compounds (DLCs). While the exposures were measured soon after the initial pulse, health outcomes were realized, or measured, 10−20 years following the initial exposure; the biologically-relevant critical exposure window for susceptibility varies with effect and may be unknown. Therefore, the effective exposure profiles for the Seveso cohort studies vary considerably. For the Mocarelli et al. (2008) and Alaluusua et al. (2004) studies, where early childhood exposures proximate to the initial event are associated with the outcomes, there is some uncertainty as to the magnitude of the effective doses. Although the effects are associated with TCDD exposure in the first 10 years of life, it is not clear to what extent the initial peak exposure is primarily responsible for the effects. It is also not clear if averaging exposure over the critical window is appropriate given the fairly large (sixfold) difference between initial TCDD body burden and body burden at the end of the critical exposure window. Because of the uncertainty in the influence of the peak exposure relative to the average exposure over the entire window of susceptibility, the LOAELs for both Mocarelli et al. (2008) and Alaluusua et al. (2004) are calculated as the average of the peak exposure and average exposure across the critical exposure window (see Section 4.2 for details).

For the gestational exposure study (Baccarelli et al., 2008), the critical exposure window is strictly defined and relatively short (9 months) and occurs long after the initial maternal exposure (20−30 years). The maternal serum TCDD concentrations were measured 16−22 years after the initial exposure when internal exposures were falling off less steeply; consequently, there is less uncertainty in the toxicokinetic extrapolation between time of measurement and time of birth. The narrow critical exposure window at a much later time than the initial exposure (where the TCDD elimination curve is flattening) is assumed to lead to a relatively steady-state exposure over the critical time period with much less uncertainty in the magnitude of the effective dose. With the exception of Eskenazi et al. (2002b) (see Section 4.2.4), the effective exposures for other effects reported for the Seveso cohort (see Section 2.4.1.1.1.4) have not been quantified for consideration as an RfD POD and are not represented in Table 4-5 because either critical exposure windows cannot be identified, unequivocal adverse effect levels cannot be determined, or individual exposure estimates were not reported. Several of these studies, however, are included in the uncertainty analysis presented in Section 4.5.

4.3.3. Uncertainty Factors (UFs)

Based on U.S. EPA (2002), UFs address five areas of uncertainty. Table 4-5 summarizes the composite (total) UF applied to the POD for each endpoint.

For the PODs based on animal bioassays, the following UFs were applied:

- *Interspecies extrapolation (UF_A).* A factor of 3 ($10^{0.5}$) was applied for interspecies extrapolation. The factor of 3 represents the residual uncertainty for toxicodynamics after accounting for toxicokinetic differences with kinetic modeling. Although there are in vitro studies (Budinsky et al., 2010; Silkworth et al., 2005) that report higher rodent sensitivities than humans for AhR-dependent enzyme induction, EPA believes that there is insufficient information on subsequent toxicological processes to conclude that rodents are more sensitive than humans for downstream adverse effects.
- *Human interindividual variability (UF_H).* A factor of 10 was applied to account for human interindividual variability in susceptibility to TCDD because there is insufficient information on sensitive populations to justify a lower value.
- *LOAEL-to-NOAEL (UF_L).* For all PODs based on the animal bioassay endpoints lacking a NOAEL, a factor of 10 was applied to account for LOAEL-to-NOAEL uncertainty. The factor of 10 is the standard value in the absence of information suggesting a lower value; the magnitude of the effects for most of the LOAELs is relatively high compared to controls.
- *Subchronic-to-chronic (UF_S).* A UF for study duration was not applied, because chronic effects for animal bioassays are well represented in the database.
- *Database factor (UF_D).* A UF for database deficiencies was not applied because the database for TCDD contains an extensive range of human and animal studies that examine a comprehensive set of endpoints. There is no evidence to suggest that additional data would result in a lower RfD.

For the PODs based on epidemiologic studies, the following UFs were applied:

- *Interspecies extrapolation (UF_A).* A UF for interspecies extrapolation was not applied because human data were utilized for derivation of the RfD.
- *Human interindividual variability (UF_H).* A factor of 3 was selected for interindividual variability to account for human-to-human variability in susceptibility. The individuals evaluated in the two principal studies included infants (exposed in utero) and adults who were exposed when they were less than 10 years of age, groups that are considered to represent sensitive lifestages. These studies considered together associate TCDD exposures with health effects in potentially vulnerable lifestage subgroups. A UF of 1

was not applied because the sample sizes for the lifestages studied were relatively small, which, combined with uncertainty in exposure estimation, may not fully capture the range of interindividual variability. In addition, potential chronic effects were not fully elucidated for humans and could possibly be more sensitive.

- *LOAEL-to-NOAEL (UF_L).* A factor of 10 was applied to account for LOAEL-to-NOAEL uncertainty. The factor of 10 for UF_L is the standard value in the absence of information suggesting a lower value.
- *Subchronic-to-chronic (UF_S).* A UF for study duration was not applied, because, although chronic effect levels are not well defined for humans, animal bioassays indicate that duration of exposure is not likely to be a determining factor in toxicological outcomes. Developmental effects and other short-term effects occur at doses similar to effects noted in chronic studies.
- *Database factor (UF_D).* A UF for database deficiencies was not applied because the database for TCDD contains an extensive range of human and animal studies that examine a comprehensive set of endpoints. There is no evidence to suggest that additional data would result in a lower RfD.

4.3.4. Choice of Human Studies for RfD Derivation

For selection of the POD, the human studies are preferred, as EPA favors human data over animal data of comparable quality. The human studies included in Table 4-5 (Baccarelli et al., 2008; Mocarelli et al., 2008; Alaluusua et al., 2004) each evaluate a segment of the Seveso civilian population (i.e., not an occupational cohort) exposed directly to TCDD released from an industrial accident. (The identification of PODs from these studies is detailed in Sections 4.3.4.1, 4.3.4.2, and 4.3.4.3.) Thus, exposures were primarily to TCDD, with apparently minimal DLC exposures beyond those associated with background intake,[14] qualifying these studies for use in RfD derivation for TCDD. In addition, health effects associated with TCDD exposures were observed in humans, eliminating the uncertainty associated with interspecies extrapolation. The cohort members who were evaluated included infants (exposed in utero) and adults who were exposed when they were less than 10 years of age. These studies considered together associate TCDD exposures with health effects in potentially vulnerable lifestages. Finally, the two virtually identical RfDs from different

[14]As an example, note the lack of statistically significant effects reported by Baccarelli et al. (2008) (Figures 2 C and D) in regression models based on either maternal plasma levels of noncoplaner PCBs or total TEQ on neonatal TSH levels.

endpoints in different studies provide an additional level of confidence in the use of these data for derivation of the RfD for TCDD.

Although the human data are preferred, Table 4-5 presents a number of animal studies with RfDs that are lower than the human RfDs. Two of the rat bioassays among this group of studies—Bell et al. (2007b) (RfD = 1.4E−9 mg/kg-day based on delay in the onset of puberty) and NTP (RfD = 4.6E−10 mg/kg-day based on liver and lung lesions) (2006a)—are of particular note. Both studies were recently conducted. Both were very well designed and conducted, using 30 or more animals per dose group (see Table 4-6 for a discussion of these studies' strengths and weaknesses); both also are consistent with and, in part, have helped to define the current state of practice in the field. Bell et al. (2007b) evaluated several reproductive and developmental endpoints, initiating TCDD exposures well before mating and continuing through gestation. NTP (2006a) is the most comprehensive evaluation of TCDD chronic toxicity in rodents to date, evaluating dozens of endpoints at several time points in all major tissues. Thus, proximity of the RfDs derived from these two recent high-quality studies provides additional support for the use of the human data for RfD derivation.

There are several animal bioassay candidate RfDs at the lower end of the RfD range in Table 4-5 that are more than 10-fold below the human-based RfDs. Two of these studies report effects that are analogous to the endpoints reported in the three human studies and support the RfDs based on human data. Specifically, decreased sperm production in Latchoumycandane and Mathur (2002) is consistent with the decreased sperm counts and other sperm effects in Baccarelli et al. (2008), and missing molars in Keller et al. (2008a; 2008b; 2007) are similar to the dental defects seen in Alaluusua et al. (2004). Thus, because these endpoints have been associated with TCDD exposures in humans, these animal studies were not selected for RfD derivation in preference to human data showing the same effects.

Another characteristic of the remaining studies in the lower end of the candidate RfD distribution is that they are dominated by mouse studies (comprising 7 of the 9 lowest candidate RfDs). EPA has less confidence in the candidate RfD estimates based on mouse data than those based on either the rat or human data. EPA has less confidence in the use of the Emond mouse PBPK model to estimate the PODs because of the lack of key mouse-specific data, particularly for the gestational component (see Section 3.3.4.3.2.5). The toxicokinetic interspecies extrapolation factors used for mice are very large, introducing a potential for large errors. The

ratio of administered dose to HED (D_a:HED) ranges from 65 to 1,227 depending on the duration of exposure. The D_a:HED for mice is, on average, about four times larger than that used for rats. In addition, each one of the mouse studies has other qualitative limitations and uncertainties (discussed above and in Table 4-6) that further reduce confidence in using them as the basis for the RfD.

4.3.4.1. *Identification of POD from Baccarelli et al. (2008)*

Baccarelli et al. (2008) reported increased levels of TSH in newborns exposed to TCDD in utero, indicating a possible dysregulation of thyroid hormone metabolism. The study authors related TCDD concentrations in maternal plasma to neonatal TSH levels using a multivariate linear regression model adjusting for a number of covariates (gender, birth weight, birth order, maternal age, hospital, and type of delivery). Based on this regression modeling, EPA has defined the LOAEL for Baccarelli et al. (2008) to be the maternal TCDD LASC of 235 ppt (at delivery) corresponding to a neonatal TSH level of 5 µU/mL.

The World Health Organization (WHO/UNICEF/ICCIDD, 1994) established the 5 µU/mL standard as an indicator of potential iodine deficiency and potential thyroid problems in neonates. Increased TSH levels are indicative of decreased thyroid hormone (T4 and/or T3) levels. The 5 µU/mL limit for TSH measurements in neonates was recommended by WHO (1994) for use in population surveillance programs as an indicator of iodine deficiency disease (IDD). In explaining this recommendation, WHO (1994) stated that

> While further study of iodine replete populations is needed, a limit of 5µU/ml whole blood… may be appropriate for epidemiological studies of IDD [iodine deficiency disease.] Populations with a substantial number of newborns with TSH levels above the limit could indicate a significant IDD problem.

For TCDD, the toxicological concern is not likely to be iodine uptake inhibition, but rather increased metabolism and clearance of T4, as evidenced in a number of animal studies (see discussion in Section 4.3.6.1). Baccarelli et al. (2008) discount iodine status in the population as a confounder, as exposed and referent populations all lived in a relatively small geographical area. It is unlikely that there was iodine deficiency in one population and not in the other population based on iodine levels in the soil.

Clinically, a TSH level of >4 µU/mL in a pregnant woman is followed up by an assessment of free T4, and treatment with L-thyroxine is prescribed if T4 levels are low (Glinoer and Delange, 2000). This is to ensure a sufficient supply of T4 for the fetus, which relies on maternal T4 exclusively during the 1st half of pregnancy (Chan et al., 2005); (Calvo et al., 2002; Morreale de Escobar et al., 2000). Adequate levels of thyroid hormone also are essential in the newborn and young infant as this is a period of active brain development (Zoeller and Rovet, 2004; Glinoer and Delange, 2000). Smaller reserves, higher demand, and shorter half-life of thyroid hormones in newborns and young infants also could make this lifestage more susceptible to the impact of insufficient levels of T4 (Savin et al., 2003; Greer et al., 2002; Van Den Hove et al., 1999). Thyroid hormone disruption during pregnancy and in the neonatal period can lead to neurological deficiencies, particularly in the attention and memory domains (Oerbeck et al., 2005). While such altered hormone levels are associated with decreased IQ scores. (e.g., 2009), report such associations among adolescents), the exact relationship between TSH increases and adverse neurodevelopmental outcome is not well defined. A TSH level above 20 µU/L in a newborn infant is cause for immediate intervention to prevent mental retardation, often caused by a malformed or ectopic thyroid gland in the newborn (WHO, 2007; Rovet, 2002; Glinoer and Delange, 2000). Recent epidemiological data indicate concern for even lower level thyroid hormone perturbations during pregnancy. For example, Haddow et al. (1999) reported that women with subclinical hypothyroidism, with a mean TSH of 13.2 µU/L had children with IQ deficits of up to 4 IQ points on the Wechsler IQ scale. Neonatal TSH within the first 72 hours of birth [as was evaluated by Baccarelli et al. (2008)] is a sensitive indicator of both neonatal and maternal thyroid status (Delange et al., 1983). Animal models have recently indicated that very modest perturbations in thyroid status for even a relatively short period of time can lead to altered brain development (Sharlin et al., 2010; Royland et al., 2008; Sharlin et al., 2008; Ausó et al., 2004; Lavado-Autric et al., 2003). Rodent bioassay results also suggest that elevated TSH levels in neonates can affect sperm development as adults (Anbalagan et al., 2010); this study also reported reduced fertility among adult males and females with increased neonatal TSH levels.

EPA has defined the LOAEL for Baccarelli et al. (2008) to be the maternal TCDD LASC of 235 ppt corresponding to a neonatal TSH level of 5 µU/mL, determined by the regression modeling performed by the study authors. Using the Emond human PBPK model, the daily oral

intake at the LOAEL is estimated to be 0.020 ng/kg day (see Section 4.2.3.1). A NOAEL is not defined because it is not clear what maternal intake should be assigned to the group below 5 μU/mL.

4.3.4.2. *Identification of POD from Mocarelli et al. (2008)*

Mocarelli et al. (2008) reported decreased sperm concentrations and decreased motile sperm counts in men who were 1−9 years old in 1976 at the time of the Seveso accident (initial TCDD exposure event). The sperm concentrations and motile sperm counts of men who were 10−17 years old in 1976 were not decreased. Serum (LASC) TCDD levels were measured in samples collected within 1 year of the initial exposure. Serum TCDD levels and corresponding responses were reported by quartile, with a reference group of less-exposed individuals assigned a TCDD LASC value of 15 ppt (which was the mean of the TCDD LASC reported in individuals outside the contaminated area). In the reference group, mean sperm concentrations and motile sperm counts were approximately 74 million sperm/mL and 41%, respectively[6]. The lowest exposed group (1st-quartile) TCDD LASC mean was 68 ppt. In the 1st quartile, mean sperm concentrations of approximately 55 million sperm/mL[15] and motile sperm counts of approximately 36% were reduced about 25 and 12%, respectively, from the reference group. Further decrease in these measures in the groups exposed to more than 68 ppt was minimal. Relative to the reference population, the percent decreases in sperm concentrations were approximately 25, 21, and 33% in the 2nd, 3rd, and 4th quartiles, respectively, and the percent decreases in progressive sperm motility were approximately 20, 25, and 22% in the 2nd, 3rd, and 4th quartiles, respectively.

Mocarelli et al. (2008) also conducted a separate analysis of all the 22−31 year-old men (combining all quartiles of the men exposed when they were 1−9 years of age). In the exposed men, the mean total sperm concentration was reported by Mocarelli et al. (2008) to be 53.6 million/mL, with a value of 21.8 million/mL at 1 standard deviation below the mean. In the comparison group that consisted of men not exposed to TCDD by the Seveso explosion and of the same age as the exposed men, the mean total sperm concentration was 72.5 million/mL (31.7 million/mL at 1 standard deviation below the mean).

[15] This estimate is based on Figure 3 in Mocarelli et al. (2008).

There is no clear adverse effect level indicating male fertility problems for either of these sperm effects. As sperm concentration decreases, the probability of pregnancy from a single ejaculation also decreases; infertile conditions arise when the number of normal sperm per ejaculate is consistently and sufficiently low. Previously, the incidence of male infertility was considered increased at sperm concentrations less than 20 million sperm/mL (WHO, 1980). More recently, Cooper et al. (2010) suggested that the 5th percentile for sperm concentration (15 million/mL) could be used as a limit by clinicians to indicate needed follow-up for potential infertility. Skakkeback (2010) suggests the following two limits for human sperm concentrations: 15 million sperm/mL, based on Cooper et al. (2010) and 40 million sperm/mL. Skakkeback justifies the upper level of 40 million sperm/mL citing a study by Bonde et al. (1998) of couples planning to become pregnant for the first time; in the Bonde study, pregnancy rates declined when sperm concentrations were below 40 million sperm/mL. Skakkeback suggests that 15 million sperm/mL may be too low of a limit off for normal fertility and that sperm concentrations between 15 million sperm/mL and 40 million sperm/mL may indicate a range of reduced fertility. For fertile men, between 50% and 60% of sperm are motile (Swan et al., 2003; Slama et al., 2002; Wijchman et al., 2001). Any impacts on these reported levels could become functionally significant, leading to reduced fertility. Low sperm counts are typically accompanied by poor sperm quality with respect to morphology and motility (Slama et al., 2002).

EPA judged that the impact on sperm concentration and quality reported by Mocarelli et al. (2008) is biologically significant given the potential for functional impairment. Although a decrease in sperm concentration of 25% likely would not have clinical significance for a typical individual, EPA's concern with the reported decreases in sperm concentration and total number of motile sperm (relative to the comparison group) is that such decreases associated with TCDD exposures could lead to shifts in the distributions of these measures in the general population. Because male fertility is susceptible to reductions in both the number and quality of sperm produced, such shifts in the population could result in decreased fertility in men at the low ends of these population distributions. Further, in the group exposed due to the Seveso accident, individuals 1 standard deviation below the mean had sperm concentrations of 21.8 million/mL; this concentration falls at the low end of the range of reduced fertility (15 million and 40 million sperm/mL) suggested by (Skakkebaek, 2010).

EPA has designated the lowest exposure group (68 ppt) as a LOAEL, which translates to a continuous daily oral intake of 0.020 ng/kg-day (see Section 4.2.3.2). The reference group is not designated as a NOAEL because the serum levels were not measured for this group, directly, and background exposures to dioxin-like compounds (DLCs) are relatively large by comparison to TCDD in this group, introducing too much uncertainty in quantifying the full NOAEL exposure (see discussion in Section 4.5). Also, there is no clear zero-exposure measurement for any of these endpoints, complicating the interpretation of the reference group response as a true "control" response (see discussion in Section 4.4). However, males less than 10 years old can be designated as a sensitive lifestage as compared to older males who were not affected.

4.3.4.3. *Identification of POD from Alaluusua et al. (2004)*

Alaluusua et al. (2004) reported dental enamel defects and missing permanent teeth in male and female adults who were less than 5 years of age, but not older, at the time of the initial exposure (1976) in Seveso. EPA used the same approach to estimate daily TCDD intake as was used for the Mocarelli et al. (2008) data; a window of susceptibility of about 5 years was established. Serum measurements for this cohort were taken within a year of the accident. Serum TCDD levels and corresponding responses were reported by tertile, with a reference group of less-exposed individuals assigned a TCDD LASC value of 15 ppt (ng/kg); the tertile group means were 130, 383, and 1,830 ppt. Both a NOAEL and LOAEL can be defined for this study. The NOAEL is 0.12 ng/kg-day, corresponding to the TCDD LASC of 130 ppt at the first tertile. The LOAEL is 0.93 ng/kg-day at the second tertile. The children in this cohort less than 5 years old can be designated as a sensitive lifestage as compared to older individuals who were not affected relative to the reference group.

4.3.5. Derivation of the RfD

The two human studies, Baccarelli et al. (2008) and Mocarelli et al. (2008), have identical LOAELs of 0.020 ng/kg-day. Together, these two studies define the most sensitive health effects in the epidemiologic literature and constitute the best foundation for establishing a POD for the RfD, and are designated as co-principal studies. Therefore, increased neonatal TSH levels in Baccarelli et al. (2008) and male reproductive effects (decreased sperm count and motility) in Mocarelli et al. (2008) are designated as co-critical effects. A composite UF of 30 is

applied to the LOAEL of 0.020 ng/kg-day to account for lack of a NOAEL ($UF_L = 10$) and human interindividual variability ($UF_H = 3$); the resulting RfD in standard units is 7×10^{-10} mg/kg-day. Table 4-7 presents the details of the RfD derivation.

4.3.6. Studies Reporting Outcomes Comparable to the Principal Studies Used to Derive the RfD

Other animal and human epidemiological studies report associations between TCDD exposures and effects similar to those reported by Baccarelli et al. (2008) and Mocarelli et al. (2008).

4.3.6.1. *Dysregulation of Thyroid Hormone Metabolism Associated with Dioxin Exposure in Neonates*

One of the principal studies for the dioxin noncancer RfD, Baccarelli et al. (2008), reported increased levels of TSH in newborns exposed to TCDD in utero, indicating a possible dysregulation of thyroid hormone metabolism. No other human studies that met the selection criteria of this analysis reported similar effects.

However, based on an analysis of over 20 epidemiology studies, Goodman et al. (2010) concluded that DLC exposures were not clearly or consistently correlated with differences in thyroid hormone levels in neonates and children less than 12 years of age. Focusing on neonatal TSH for direct comparison to Baccarelli et al. (2008), Goodman et al. (2010), in Table 3 of their analysis, identify 13 different studies, including Baccarelli et al. (2008), which measured infant TSH levels within 1 week of birth. Of these studies, only Baccarelli et al. (2008) was TCDD-specific and evaluated exposures well above ambient exposure levels. The other studies examined total TEQ or individual DLCs near background exposure levels. The LOAEL derived by EPA from Baccarelli et al. (2008) is approximately sixfold higher than the ambient total TEQ exposure levels at the time of the exposures for the general Seveso population[16] and more than 30-fold above an estimate of current TEQ levels (Lorber et al., 2009). In the other studies, the exposures appear to have been largely to DLCs, with TCDD as a minor component. Because the equivalent TCDD exposure for DLCs is derived from TEF methodology, which is conservative in nature (TEFs are higher than the median), the total TEQ concentrations would likely be over-

[16]Estimated by EPA to be 3.5×10^{-3} ng/kg-day on a total TEQ basis (see Section 4.5.1.1 and Appendix F).

estimated (relative to TCDD) and uncertain. In addition, only 2 of the other 12 studies evaluated by Goodman et al. (2010) reported TSH measures 3 days after birth, which is an international standard and would be most comparable to those in Baccarelli et al. (2008). TSH levels generally peak about 2 hours after birth then decline rapidly to typical long-term levels over the next few days (Steinmaus et al., 2010). Several of the studies included in Table 3 of Goodman et al. (2010) evaluated cord-blood TSH measurements, which represent early high TSH concentrations and are not directly comparable to 3-day measurements. Given these considerations, particularly the relatively low ambient exposures and differences in the timing of TSH measures, it would be unlikely that any consistent pattern would be detected across these studies.

Several animal studies that met the selection criteria evaluated the effects of TCDD on the thyroid or thyroid hormone levels. Overall, this set of studies show that TCDD affects thyroid hormone levels and the thyroid gland. The studies of Sewall et al. (1995), Seo et al. (1995), Van Birgelen et al. (1995a; 1995b), Crofton et al. (2005), and NTP (2006a) each reported decreases in T4 levels. In response to TCDD treatment, NTP (2006a) reported increases in total T3 concentrations, and both NTP (2006a) and Sewall et al. (1995) reported increased TSH concentrations. Sewall et al. (1995) and Chu et al. (2007) reported reductions in thyroid follicles, with Chu et al. (2007) noting that, of the health effects observed in their study, thyroid effects were the most sensitive to TCDD exposures. Although none of these studies address in utero or neonatal exposure, they show that TCDD can affect the level of thyroid hormones and the thyroid organ in adult animals.

4.3.6.2. *Male Reproductive Effects associated with Dioxin Exposures*

The other principal study for the dioxin noncancer RfD, Mocarelli et al. (2008), reported decreased sperm concentrations and decreased motile sperm counts in men who were aged 1−9 years at the time of the Seveso accident (initial TCDD exposure event). The sperm concentrations and motile sperm counts of men who were 10−17 years old in 1976 were not adversely affected. While no other human studies that met the selection criteria of this analysis reported similar effects, a newly published study, Mocarelli et al. (2011), also reports male reproductive effects. Several animal studies that met the study selection criteria also reported male reproductive effects.

Mocarelli et al. (2011) examined the relationship between maternal serum TCDD levels and semen quality in male offspring. Analyses were based on 39 of the 78 men aged 18−26 years born to women residing in the areas most heavily polluted by dioxin after the explosion in Seveso, Italy, in 1976 and age-matched controls (58 out of 123 recruited) born to women residing in noncontaminated areas of Italy. In the exposed group of women, pregnancies occurred between 9 months and 6 years after the accident (March 1977−January 1984). The male offspring of these women were categorized based on whether they were breastfed (n = 21, born to 20 mothers) or formula-fed (n = 18, born to 17 mothers) as infants. In the comparison group, 36 were breastfed, and 22 were formula-fed. Sons born to dioxin-exposed women whose spouses were also exposed to TCDD, as well as all men with reported diseases, were excluded.

TCDD exposures were based on estimated maternal serum concentration at conception. To estimate these levels in the exposed group, the authors relied on maternal serum measures, all of which were collected shortly after the accident in 1976−1977, and a biokinetic model (Kreuzer et al., 1997) that estimated TCDD elimination from the time of the accident to conception for individual women (average half-life = 4 years). Mothers of sons in the comparison group were assumed to be exposed to average background TCDD levels of 10 ppt based on measurements reported in Eskenazi et al. (2004).

Semen samples were collected from all participants. These samples were maintained at 37°C and examined within an hour of ejaculation. For serum inhibin B and FSH analyses, fasting blood samples were obtained the morning of semen collection. Statistical analyses were performed on sperm properties, serum hormone levels, and TCDD levels using a "general linear model" (Mocarelli et al., 2011). Model covariates included age, duration of abstinence prior to semen collection, smoking status, exposures to organic solvents, adhesives or paints, BMI, alcohol use, educational level, and employment status.

Relative to the comparison group, men born to exposed mothers had decreased sperm concentration (46 million vs. 81 million sperm/mL; p = 0.01), total sperm count (144 million vs. 231 million sperm; p = 0.03), and total number of motile sperm (51 million vs. 91 million; p = 0.05). Relative to the breastfed comparison group, breastfed sons born to exposed mothers exhibited decreased sperm concentrations (36 million vs. 86 million sperm/mL; p = 0.002), total sperm counts (117 million vs. 231 million sperm; p = 0.02), and motile sperm counts (39 million vs. 98 million; p = 0.01). Relative to the breastfed comparison group, breastfed sons born to

exposed mothers also exhibited increased FSH concentrations (4.1 vs. 2.6 IU/L; $p = 0.03$) and decreased inhibin B levels (70.2 million vs. 101.8 pg/mL; $p = 0.01$). The formula-fed exposed and comparison groups were not significantly different by any of these measures.

This study was well-designed with well-characterized exposures (for the exposed group), which relied on measured sera TCDD concentrations and a peer-reviewed TCDD elimination model to estimate maternal serum TCDD levels at the time of conception. Exposures in the comparison group relied on estimates from other studies. The study excluded sons of fathers that were likely highly exposed to TCDD, to limit potential influences from highly exposed fathers. The study relies on self-reported recollection of infant feeding (i.e., breastfed vs. formula-fed), which may lead to some misclassification based on recall error. Statistically significant associations were evident for both the exposed men and their comparison group and breastfed men and the breastfed comparison group.

In this study, elevated TCDD exposures during and after pregnancy (via breast-feeding) led to long-term decrements in male reproductive endpoints. These effects included changes in levels of hormones that affect spermatogenesis; they also include decreases in sperm concentration and sperm motility.

In addition, two rodent bioassays also report sperm effects associated with dioxin treatment. Latchoumycandane and Mathur (2002) reported decreased daily sperm production and decreased reproductive organ weights in male albino Wistar rats given daily oral doses of TCDD for 45 days. The LOAEL was 1.0 ng/kg-day, which corresponds to a $LOAEL_{HED}$ of 0.016 ng/kg-day (see Table 4-5); a NOAEL was not identified. Simanainen et al. (2004b) reported a reduction in daily sperm production and cauda epididymal sperm reserves in male rat pups born to dams exposed to 300 ng/kg TCDD or higher on GD 15 by oral gavage. In this case a NOAEL of 100 ng/kg was identified, which corresponds to a $NOAEL_{HED}$ of 0.433 ng/kg-day, with a $LOAEL_{HED}$ of 1.7 ng/kg-day (see Table 4-5). Detailed descriptions of these studies can be found in Appendix D.

4.4. QUALITATIVE UNCERTAINTIES IN THE RFD

Exposure assessment is a key limitation of the epidemiologic studies (of the Seveso cohort) used to derive the RfD. The Seveso cohort exposure profile consists of an initial high

TCDD exposure[17] followed by a drop in body burden to background levels over a period of about 20 years, at which time the effects were observed. This exposure scenario is inconsistent with the constant daily intake scenario addressed by the RfD methodology. The determination of an effective average daily dose from the Seveso exposure scenario requires a consideration of the biologically-relevant critical time-window of susceptibility and the influence of the peak exposure on the occurrence of the observed effects, particularly when the peak exposure is high relative to the average exposure over the critical exposure window. For one of the principal studies (Mocarelli et al., 2008), a maximum susceptibility exposure window can be identified based on the age of the population at risk. However, the influence of the peak exposure on the effects observed 20 years later is unknown, and the biological significance of averaging the exposure over several years, with internal exposure measures spanning a 5.5-fold range, is unknown. EPA has not developed guidance for large interval averaging. Furthermore, because there is an assumption of a threshold level of exposure below which noncancer effects are not expected to occur, averaging over large intervals could include exposures that are below a threshold. The process used by EPA to estimate the LOAEL exposure for the Mocarelli et al. (2008) study is a compromise between the most- and least-conservative alternatives; as such, there is some uncertainty in the estimate, perhaps in the range of 3- to 10-fold in either direction. This uncertainty also applies to the LOAEL determined for the developmental dental effects reported in Alaluusua et al. (2004) and the increased menstrual cycle length reported in Eskenazi et al. (2002b) (see Section 4.2.3.4); in both of those studies, the uncertainty is greater, as the difference between peak and average internal exposures is an order of magnitude or more. The LOAEL for increased TSH in neonates (Baccarelli et al., 2008), however, is less uncertain because the critical exposure window is much narrower (9 months), and the developmental exposures occurred 20 to 30 years after the initial exposure, when internal TCDD concentrations for the pregnant women likely were leveling off; that is, exposure over the critical window was more constant and estimation of the relevant exposures was less uncertain. However, there is some uncertainty in the magnitude of the exposures because they were estimated from

[17]Mocarelli (2001) reported the release from the Seveso plant to contain a mixture of TCDD, ethylene glycol, and sodium hydroxide. Because these chemicals are not thought to persist in the environment or in the body, coexposure to these additional contaminants along with TCDD would not have a significant impact on longer-term TCDD dose-response. For acute exposure, male reproductive or thyroid hormone effects are not evident for ethylene glycol (U.S. EPA, 2009a). It is unlikely that sodium hydroxide, being primarily a caustic agent, would cause these effects.

measurements in sera taken several years prior to pregnancy and do not take into account changing patterns of exposure during pregnancy.

Another source of uncertainty using human epidemiologic data is the lack of completely unexposed populations. The available TCDD epidemiologic data were obtained by comparing populations that experienced elevated TCDD exposures to populations that experienced lower exposures, rather than to a population with no TCDD exposure. An additional complicating factor is coexposure to DLCs, which can act toxicologically in the same way as TCDD. Although the accidental exposure to the Seveso women's cohort was primarily to TCDD, background exposure was largely to DLCs. Eskenazi et al. (2004) reported that TCDD comprised only 20% of the total toxicity equivalence (TEQ) in the serum of the reference group that was not exposed as a result of the Seveso factory explosion, which implies that the effective background TEQ exposure was approximately fivefold higher than exposure to TCDD. WHO (1998) estimated that TCDD may comprise only 5–20% of background exposures to dioxin and DLCs. The higher background exposure could be significant at the lower TCDD exposure levels, with the effect diminishing as TCDD exposure increased. For dose-response modeling, the effect of a higher background dose (i.e., total TEQ), if included, would be to shift the response curve to the right, with responses now being associated with higher exposures. Adding a constant to all exposures would also reduce the proportional spread of the exposures, which would tend to alter the shape of the dose-response curve towards sublinear. Both the right shift and the more sublinear shape would result in higher POD estimates. In addition, the response in the reference population is not a true zero-exposure (TEQ-free) response. The actual magnitude of the impact of the DLC background exposure is impossible to assess without knowing the zero-exposure background response. The (TEQ-free) background response cannot be assessed as no TEQ-free population exists. Ideally, an independent absolute measure of adversity in terms of the response variable, such as the 5 μU/mL neonatal TSH benchmark, is needed for dose-response modeling.

As part of the uncertainty analysis for the TCDD RfD, the possible influence of different background DLC exposure assumptions on the POD estimates derived from the two principal studies, Baccarelli et al. (2008) and Mocarelli et al. (2008), and one comprehensive animal bioassay, NTP (2006a), is examined quantitatively in Section 4.5. In addition, the range of

possible PODs for other epidemiologic studies that did not pass all the selection criteria in comparison to the principal studies is presented in Section 4.5.

A primary strength of the TCDD database is that analogous effects have been observed in animal bioassays for most of the human endpoints, increasing the overall confidence in the relevance to humans of the effects reported in rodents and the association of TCDD exposure with the health outcomes reported in humans. Table 4-5 shows that low-dose TCDD exposures are associated with a wide array of toxicological endpoints in rodents including developmental effects, reproductive effects, immunotoxicity, and chronic toxicity. Effects reported in human studies are similar, including male reproductive effects, increased TSH in neonates, and dental defects in children; other human health effects such as female reproductive effects and chloracne have been observed at higher exposures (see Appendix C). Severe liver toxicity, which is a consistently reported effect in rodents, has not been observed in humans; Michalek et al.(2001c), however, reported slightly elevated liver enzyme levels in serum and other nonspecific liver effects for the Ranch Hand cohort, suggestive of mild liver toxicity. Overt immunological endpoints, reported in the rodent bioassays, also have not been reported in human studies. However, with respect to immunological effects, Baccarelli et al. (2004; 2002) evaluated immunoglobin and complement levels in the sera of TCDD-exposed individuals from the Seveso cohort and found reduced immunoglobulin in the highest exposure groups but no effect on other immunoglobulins or on C3 or C4 complement levels and no indication of compromised immune response. The latter finding indicates that at least one immunological measure in humans is not a sensitive endpoint, as it is for mice, with large reductions in serum complement at low exposure levels (White et al., 1986).

Although there is a substantial amount of qualitative concordance of effects between rodents and humans, quantitative concordance is not as strong, with reference to Table 4-5. The differential sensitivity of mice and humans for the serum complement endpoint is one example. Other examples of differential sensitivity are developmental dental effects and thyroid hormonal dysregulation. Developmental dental defects are relatively sensitive effects in rodents, appearing at exposure levels in mice (Keller et al., 2008a; Keller et al., 2008b; Keller et al., 2007) more than an order of magnitude lower than effect levels in humans (Alaluusua et al., 2004). In contrast, thyroid hormone effects are seen in rats (Crofton et al., 2005) at 30-fold higher exposures than for humans (Baccarelli et al., 2008). Male reproductive effects (sperm

production) occur in rats (Latchoumycandane and Mathur, 2002) and humans (Mocarelli et al., 2008) at about the same dose. To what extent these differential sensitivities depend on specifics of the comparison, such as species (mouse vs. rat), life-stage (e.g., fetal vs. adult), endpoint measure (e.g., thyroxine [T4] vs. TSH), or magnitude of the lowest dose tested, cannot be determined, so strong conclusions about quantitative concordance cannot be made.

A more detailed tabular and graphical presentation of qualitative and quantitative cross-species comparisons of selected toxicological endpoints for all the animal and human studies that met the EPA selection criteria is given in Appendix D.3. The endpoints include male and female reproductive effects, thyroid hormone levels, and developmental dental effects, all of which have been reported for humans. In addition, immunological and neurological effects are shown because they are sensitive effects in experimental animal studies, although not evident in humans. Hepatic effects, which are not shown in Appendix D.3, are evident in virtually all rodent studies that looked for them and are often severe, but are not severe in humans. The analysis presented in Appendix D.3 supports the conclusion that there is a substantial amount of qualitative concordance of effects between rodents and humans, but a much lower quantitative concordance. However, there are no endpoints in the selected animal bioassays that address diabetes or glucose metabolism. There may be other animal studies showing effects of interest at higher doses in those studies that did not meet the dose limit selection criterion.

A number of qualitative strengths and limitations/uncertainties are associated with the animal bioassays listed in Table 4-5, as articulated in Table 4-6. Considering the issue of lowest tested dose, the general lack of NOAELs and acceptable BMDLs is a primary weakness of the rodent bioassay database. None of the eight most sensitive rodent studies in Table 4-5, spanning an 18-fold range of LOAELs, had defined NOAELs or BMDLs. NOAELs or BMDLs were established for only 4 of the next 13 rodent studies. In addition, many of these LOAELs are characterized by relatively high responses with respect to the control population, so it is not certain that a 10-fold lower dose (based on the application of UF_L of 10) would be approximately equivalent to a NOAEL. A major reason for the failure of BMD modeling was that the responses were not "anchored" at the low end (i.e., first response levels were far from the BMR [see Table 4-4]). Another major problem with the animal bioassay data was nonmonotone and flat response profiles. The small dose-group sizes and large dose intervals probably contributed to many of these response characteristics that prevented successful BMD modeling. Larger study

sizes with narrower dose intervals at lower doses are still needed to clarify rodent response to TCDD.

Lower TCDD doses have been tested in rodents but almost entirely for investigation of specialized biochemical endpoints[18] that EPA does not consider to be toxicologically relevant for the derivation of a noncancer RfD (see Appendix H). There is, however, a fundamental limit to the lowest dose of TCDD that can be tested meaningfully, as TCDD is present in feed stock and accumulates in unexposed animals prior to the start of any study. This issue is illustrated by the presence of TCDD in tissues of unexposed control animals, often at significant levels relative to the lowest tested dose in low-dose studies (Bell et al., 2007b; Ohsako et al., 2001; Vanden Heuvel et al., 1994; Vanden Heuvel et al., 1994 see Text Box 4-1). Some DLCs also have been measured in animal feeds (Bell et al., 2007b; NTP, 2006a) and are anticipated to accumulate in unexposed test animals, further complicating the interpretation of low-dose studies.

[18] Enzyme induction, oxidative stress indicators, mRNA levels, etc.

Text Box 4-1. Background levels of TCDD in Control Group Animals

TCDD tissue levels in control animals are rarely reported either explicitly or implicitly. Vanden Heuvel et al. (1994), however, reported TCDD concentrations in livers of control animals (10-week-old female Sprague-Dawley rats) of 0.43 ppt (ng/kg) compared to 0.49 ppt in the livers of animals given a single oral TCDD dose of 0.1 ng/kg. Assuming proportionality of liver concentration to total body burden, the body burden of untreated animals was 87.8% of that of treated animals at the lowest dose. The equivalent (single) administered dose for untreated animals (d_0) can be calculated as equal to $0.878 \times (0.1 + d_0)$, assuming proportionality of body burden to administered dose and that all animals started with the same TCDD body burdens. The calculation yields a value of 0.72 ng/kg for d_0, which represents the accumulated TCDD from all sources in these animals prior to being put on and during test. This value would raise the nominal 0.1 ng/kg TCDD dose 8-fold to 0.82 ng/kg. The next higher dose of 1 ng/kg would be nearly doubled to 1.72 g/kg. The impact on higher doses would be negligible, because the ratio of treatment dose to apparent background exposure levels increases with higher treatment levels. Bell et al. (2007b) reported slightly higher levels (0.66 ppt) in the livers of slightly older untreated pregnant female Sprague-Dawley rats (mated at 16–18 weeks of age and tested 17 days later).

Ohsako et al. (2001) reported TCDD concentrations in the fat of offspring of untreated pregnant Holtzman rats that were 46% of the TCDD fat concentrations in animals exposed in utero to 12.5 ng/kg (single exposure on GD 15). This level of TCDD would imply a very large background exposure, but quantitation based on simple kinetic assumptions probably would not reflect the more complicated indirect exposure scenario.

Bell et al. (2007b) also reported concentrations of 0.1 and 0.6 ppt TCDD measured in two samples of feed stock. Assuming that the average of 0.35 ppt is representative of the entire supply of feed stock and a food consumption factor of 10% of body weight per day, the average daily oral exposure from feed to these animals would be 0.035 ng/kg. Discrimination of outcomes from longer-term repeated exposures might be problematic at exposure levels around 0.1 ng/kg-day. Background exposure was not much of an issue for Bell et al. (2007b), as the lowest TCDD exposure level was 2.4 ng/kg-day (28-day dietary exposure).

NTP (2006c) reported TCDD concentrations in the liver and fat of untreated female Sprague-Dawley rats after 2 years on test that were 1% and 2.5% of the levels in the liver and fat of the low-dose TCDD treatment group (2.14 ng/kg-day) (NTP, 2006a), respectively. Assuming proportionality of fat concentration and oral intake, control animal exposure would have been approximately 0.05 ng/kg-day, similar to the estimate from Bell et al. (2007b). As for the latter study, background intake for the NTP (2006a) study animals would not have a large effect on the dose-response assessment given the lowest exposure level of 2.14 ng/kg-day.

In all of these studies, except the 28-day exposure in Bell et al. (2007b), control animals were gavaged with corn oil vehicle. TCDD concentrations in corn oil were not reported in any of the studies.

4.5. QUANTITATIVE UNCERTAINTY IN THE RFD

The development of each candidate RfD in Sections 4.1 through 4.3 required the analysis of numerous kinetic, toxicologic, and epidemiologic data sets. These analyses included interpretive decisions that were made considering different sources of uncertainty in each study and EPA's methods for developing RfDs. This section quantifies the impacts of some sources of uncertainty encountered in the development of candidate RfDs (Sections 1.1 and 1.3 describe the the NAS and SAB comments pertaining to uncertainty analysis for the RfD). In Section 4.5.1, the impacts of some sources of uncertainty encountered in the development of candidate RfDs based on Baccarelli et al. (2008), Mocarelli et al. (2008) and NTP (2006a) are elucidated using "*variable* sensitivity" trees depicting the sensitivity of the POD value to choices made for PBPK model variables and inputs. In Section 4.5.2, an additional range of potential PODs is presented

as a bounding analysis considering background DLC exposures and several epidemiologic studies that did not qualify for RfD consideration, but for which limiting NOAEL and LOAEL values can be estimated.

4.5.1. Development of Variable Sensitivity Trees for the Principal Epidemiological Studies that were the basis of the RfD and the NTP (2006a) Rodent Bioassay

In Section 4.5.1, the impacts of some sources of uncertainty encountered in the development of candidate RfDs based on Baccarelli et al. (2008), Mocarelli et al. (2008) and NTP (2006a) are elucidated using "variable sensitivity" trees depicting the sensitivity of the POD value to choices made for PBPK model variables and inputs. Baccarelli et al. (2008) and Mocarelli et al. (2008) are the principal studies used to develop the RfD. NTP (2006a) is among the most recent and comprehensive rodent bioassay studies of TCDD. For each of the three PODs used to develop candidate RfDs from these studies, EPA generated plausible alternative interpretations of information used to define judgment-based inputs for specific model variables. The goal of this analysis is to provide quantitative insights on critical uncertainties encountered in the development of the RfD by illustrating the consequences (quantified as alternative PODs at the end of each branch in each tree) of plausible alternative interpretations of these key data sets.

Previously, in their examination of low-dose carcinogenicity associated with formaldehyde and chloroform exposures, Evans et al. (1994a; 1994b) assigned subjective weights to each branch of a probability tree and calculated probability masses for population risks associated with alternate interpretations of toxicological and pharmacokinetic data and exposure information.[19] In the examination of uncertainty undertaken in this report, EPA utilizes the development of sensitivity trees; subjective probability weights are not developed for any of the branches, and there is no propagation of probabilities across branches. Further, these trees do not present a comprehensive analysis of quantitative uncertainty of the three candidate RfDs; rather, EPA has focused on the impacts of key interpretive decisions largely dealing with exposure and kinetic modeling uncertainties. However, it should be noted that because POD values do not vary greatly across each of the three trees (less than threefold in either direction), it

[19] Small (2008) discusses other studies of distributional approaches in risk assessment by Sielken and collaborators that are similar to those of Evans and colleagues. These include the following: Sielken (1993, 1990), Holland and Sielken (1993), Sielken and Valdez Flores (1999, 1996), and Sielken et al. (1995).

is unlikely that the distribution of probability mass resulting from specific probability assignments would result in a significant amount of mass away from the chosen PODs. To extend this analysis further, candidate RfDs can be estimated by dividing the POD values EPA has generated by the appropriate uncertainty factors. The latter is largely a judgment call and cannot be modeled, per se. However, the impact of the magnitude of uncertainty factors on the RfD is proportional and relatively trivial to compute.

In this analysis, the structure of the decisions and the resulting POD estimates are presented as sensitivity trees in graphical form (see Figures 4-6 through 4-8). In these figures, the left-hand columns depict the variables considered in the sensitivity analysis. The values used for these variables were either directly specified in the literature or were based on judgment using exposure information provided in related papers. Each variable was assessed one at a time, while fixing all the other variables at the values used in the primary POD estimation that was used to develop the RfD in Section 4.3, termed hereafter the "standard pathway," and indicated in Figures 4-6 through 4-8 by the bolded lines. Up to three significant digits are shown for the PODs that are presented so that differences among the PODs across analytic choices can be readily discerned.

4.5.1.1. *Epidemiological Sensitivity Analyses*

In estimating the PODs for the principal studies for the RfD (Baccarelli et al., 2008; Mocarelli et al., 2008), a series of assumptions were made to model the exposure history of the cohorts and to estimate an intake leading to the observed effect. In this section, the series of trees highlights the effects of choosing alternative assumptions on the POD estimates.

4.5.1.1.1. *Mocarelli et al. (2008)*

To examine the impacts of potential uncertainties associated with the assumptions made in estimating the standard pathway LOAEL POD in Mocarelli et al. (2008) (see Section 4.2.3.2), EPA evaluated the impact of several alternate exposure assumptions on the oral intakes associated with the POD, as shown in Figure 4-6. The left side of the figure depicts the variables of the exposure analysis considered in the sensitivity analysis (i.e., background exposure, exposure duration, measurement lag, and age at exposure). The values used for these variables were not directly specified in the literature but were based on judgment of the exposure

information provided in Mocarelli et al. (2008) and related papers. All of these variables are inputs to the Emond human PBPK model, which was used to estimate the actual exposures to the affected population and the corresponding continuous intakes for determining the RfD POD; all modeling for this analysis was carried out using the Emond human PBPK model. Each variable was assessed one at a time, while fixing all the other variables at the values used in the standard pathway analysis. The sensitivity analysis begins with the reported LASC of 68 ppt TCDD in the LOAEL group. The terminal nodes at the bottom of the figure show the daily oral intakes (ng/kg-day) resulting from each alternative value for the variables examined.

In Figure 4-6 and in the text that follows, the following abbreviations are used:

- "*P*" identifies the intake associated with peak LASC exposure estimates.
- "*W*" identifies the intake associated with the average LASC over the actual exposure window.
- "AVG" is the average of the intakes associated with "*P*" and "*W*." Intakes associated with either "*P*" or "*W*" conceivably could have been selected as the primary POD.

Because of the relatively large differences between peak exposures and average exposures decreasing over a relatively long time span,[20] and the uncertainty of the relative influence of acute high exposures vs. lower longer-term averages on the toxicological outcome, EPA elected to use the average of the peak exposure intake (*P*) and the critical-window exposure average intake (*W*) as the basis for the POD, giving equal weight to both (see discussion in Section 4.2.3); these values are labeled as "AVG" across all terminal nodes in the tree.

For Figure 4-6, background exposures in the population (labeled "Background") were estimated in several ways, taking into account background exposures of TCDD only or the presence of other DLCs. The Emond human PBPK model was used to estimate all background intakes by assuming a constant exposure from birth to time of measurement for each scenario (see Appendix F for modeling details). The background value used in the standard pathway analysis was based on an LASC of 15 ppt used by Mocarelli et al. (2008) in their analysis as the TCDD level in the comparison group; this value was reported by Needham et al. (1998) to be the median TCDD concentration in the unexposed reference adult population (25 years or older)

[20] The modeled TCDD LASC decreased by a factor of 5.5 from peak exposure to the terminal value at 10 years.

(designated "Needham" in Figure 4-6). EPA estimated a corresponding daily TCDD intake of 3.5×10^{-4} ng/kg-day from birth, assuming that 15 ppt was obtained at age 35. The alternative is an age-specific background intake based on an average TCDD concentration of 40.5 ppt for girls less than 12 years of age (designated "Eskenazi" in Figure 4-6) (Eskenazi et al., 2004).[21] Assuming that background TCDD concentrations were similar for boys and girls in the Seveso cohort, EPA estimated an average TCDD intake of 3.52×10^{-3} ng/kg-day corresponding to the same average 40.5 ppt LASC for boys of similar age. The 10-fold higher value than for the adult background is likely a result of higher food consumption in children and a higher average environmental concentration for the relevant childhood exposure period (1964−1976) than for the adult exposures (ca. 1941−1976) (Lorber, 2002; Pinsky and Lorber, 1998).

The other alternate background scenarios take into account the presence of DLCs (i.e., other than TCDD) in the background exposure. Because DLCs are presumed to behave in the same manner as TCDD (for AhR induction), the magnitude of the background DLC exposure is an important concern in establishing the POD.

Both the "Needham" and "Eskenazi" background exposure scenarios are evaluated for DLCs. For this analysis, the total DLC-TEQ, whether reported by the authors or modeled herein, is assumed to be applicable for estimation of equivalent TCDD intake. However, the reported TEQ values are based on serum concentrations, while the TEFs on which the TEQ values are based are largely derived from oral dosing studies conducted in experimental animals. The outcomes from such studies implicitly account for DLC toxicokinetics (i.e., absorption, distribution, metabolism, and elimination). Applications of TEFs to DLC tissue concentrations do not account for toxicokinetics. Whole body half-life estimates for the DLCs vary from about 6 months to 20 years (Ogura et al., 2004; Flesch-Janys et al., 1996), so the equivalence of internally estimated TEQ with ingested quantities is not strictly valid. Currently, there is no human PBPK model capable of addressing all the DLC congeners, although both EPA (U.S. EPA, 2003) and Lorber (2002) have used DLC half-life estimates and tissue concentrations to estimate intake rates of individual DLCs in humans; however, the dioxin-like PCBs were not included in either Lorber (2002) or EPA (2003). In addition, the TEF methodology is designed

[21] Table 3 in Eskanazi et al. (2004) reports the results of two pools of sera collected from girls aged 0−12 years, who did not reside in areas affected by the Seveso accident and were presumably exposed only to background levels of TCDD. EPA estimated the mean of these reported sera concentrations of 47.6 ppt TCDD and 33.4 ppt TCDD.

to be health protective, in that the TEFs are not central tendency estimates but biased high by design (Van den Berg et al., 2006). Two different approaches for estimating background DLC exposures are presented.

The first approach models the exposure directly, by matching the total TEQ (TCDD included) at the time of measurement with the corresponding intake using the Emond model. The total TEQ for the Eskenazi background scenario is estimated from Table 3 in Eskenazi et al. (2004). The average TEF_{05} DLC-TEQ contribution was estimated by multiplying the 0−12 year old average of 76.05 ppt (based on TEF_{98} values) by a factor of 0.7. The factor of 0.7 is an approximation based on a ratio of 0.72 for the TEF_{05} to TEF_{98} background DLC-TEQ values for the Ranch Hand cohort (Pavuk et al., 2007) and a ratio of 0.65 based on serum collected in 1998 for 78 Seveso women (Warner et al., 2005). The Ranch Hand value was determined by Pavuk et al. (2007) and reported directly. The 0.65 ratio for the Seveso women was determined by EPA by calculating the total TEQ using both the 1998 and 2005 TEF values from the median congener concentrations reported by Warner et al. (2005). Figure 4-6 shows the results of modeling total TEQ directly under this approach, labeled as "Modeled" under the "Total TEQ" branches for both the "Needham" and "Eskenazi" background exposure scenarios.

The second approach for estimating DLC background exposure is a simple additive one, in which an estimate of background DLC-TEQ intake is added to the modeled TCDD intake. This is accomplished by assuming that TCDD comprises 10% of the total background TEQ, which is about the proportion of TCDD to total TEQ in serum as estimated by WHO (1998). In addition, TCDD is about 10% of the total serum TEQ as calculated by EPA from the NHANES (2001/2002) data reported by Lorber et al. (2009). However, the same qualifier holds here as for modeling total TEQ directly, in that the TEFs are based on oral exposures. If the proportional relationship (i.e., TCDD is 10% of total TEQ) is assumed for oral exposure, the modeled TCDD intake is simply multiplied by nine to get the corresponding DLC-TEQ intake. The TCDD background exposures for the Needham and Eskenazi background scenarios are 3.5×10^{-4} ng/kg-day and 3.5×10^{-3} ng/kg-day, respectively (see Appendix F for details); the corresponding DLC-TEQ intakes for the additive background approach are 3.15×10^{-3} ng/kg-day and 3.15×10^{-2} ng/kg-day, respectively. Figure 4-6 shows the additive approach, labeled as "DLC-TEQ added" under the "Total TEQ" branches for both the "Needham" and "Eskenazi" background exposure scenarios.

"Exposure duration" refers to the duration of the elevated (external) TCDD exposures immediately following the Seveso accident, which is not known with certainty. In the standard pathway analysis, the "exposure duration" of the TCDD exposures due to the Seveso accident was modeled using the Emond model as a single pulse on 1 day (i.e., 24 hours). The alternative also uses the Emond model but models the exposures following the Seveso accident using pulse doses on two consecutive days (i.e., 48 hours).

"Measurement lag" refers to the period of time between TCDD exposure following the Seveso accident and the collection of blood for future TCDD analyses. Within the Seveso cohort, serum samples were collected in 1976 and 1977, so in the standard pathway analysis, an average measurement lag time of 6 months was assumed for exposure to TCDD. The alternative analyses simulate lag times of 1 month and 1 year.

"Age at exposure" is the average age of the susceptible lifestage (boys, 1−9 years old) at the time of the Seveso accident. Within the cohort, the average age at exposure was reported to be 6.2 years, which was used in standard pathway analysis. The alternative analysis considers individuals who would have been 1 year or 9 years of age at the time of the Seveso accident, representing the bounds of the susceptible age range. This category is included to show the potential range of exposures across the cohort rather than to evaluate plausible alternatives to the mean age of 6.2 years. That is, the intakes associated with ages 1 or 9 would not be considered as PODs.

Overall, excluding the age-at-exposure variable, the daily intakes (TCDD or total TEQ) based on the alternative assumptions in this tree vary between 0.007 ng/kg-day (*W* for 1-month measurement lag) and 0.05 ng/kg-day (*P* for modeled total TEQ, Needham background). This range spans the LOAEL for the standard pathway analysis of 0.020 ng/kg-day by less than a factor of three on each side. The AVG values vary over a smaller range from 0.013 ng/kg-day (TCDD-only, Eskenazi background) to 0.0335 ng/kg-day (modeled total TEQ, Needham background), bracketing the LOAEL for the standard pathway by less than a factor of two.

The ratio of peak intake to window-average intake (*P*:*W* ratio) is of interest in evaluating the range of exposures over which an average is taken. The *P*:*W* ratio is 4 for the standard pathway POD. In general, the *P*:*W* ratios are greater than three across the terminal nodes. However, the higher the background exposure, the lower the peak intake and the lower the *P*:*W* ratio and the lower the impact of averaging *P* and *W*. The *P*:*W* ratio is lowest for all the

Eskenazi background scenarios, decreasing to about a factor of 1.3 for the TEQ analyses. The higher background exposure scenario had the largest impact on the TCDD-only intakes, with a 35% lower AVG than for the standard pathway RfD LOAEL POD. The next largest variation was for the 48-hour exposure time, with a 24% lower AVG than for the 24-hour scenario. However, the modeled exposures on each of the 2 days were equal when, in reality, they would be decreasing with time, such that the peak is somewhat underestimated in this exercise; longer exposure scenarios assuming constant levels would not be realistic. The largest differences in the other direction were obtained for the modeled total TEQ scenarios, with a 67% higher AVG for the Needham background assumption (compared to the standard pathway RfD POD) and a 30% higher AVG for the Eskenazi background assumption. Note that any DLC background exposure estimate based on TEQ will be an over-estimate because of the conservative nature of the TEF methodology. All the other alternative assumptions resulted in a 16% or lower change in the AVG values. Although not a consideration for defining the POD, the TCDD AVG intakes across the susceptible age range (1−9 years) were within 5% of the standard pathway RfD POD, but with a large *P*:*W* ratio (10) for 1-year-olds.

In summary, the quantitative uncertainties evaluated here for the RfD POD based on Mocarelli et al. (2008) span less than a 3-fold range in either direction. The largest differences are those between peak and window-average exposures, which decrease when considering the alternative Eskenazi background. Using the latter, the AVG POD is about half of the RfD POD, but is more impacted by background DLC exposure; considering the TEQ contribution from this background exposureresults in approximately the same value as the RfD POD with additive background DLC. Using the directly-modeled approach, background DLC exposure has a larger impact on the standard RfD POD, increasing it by 67%. At this time, EPA cannot recommend any approach for incorporating background DLC exposure directly into the POD for the RfD. Overall, given the bidirectional nature and relatively small magnitude of the uncertainties, EPA believes that this sensitivity analysis provides support for the magnitude of the RfD.

4.5.1.1.2. *Baccarelli et al. (2008)*

To examine the impacts of potential uncertainties associated with the assumptions made in estimating the standard pathway POD for Baccarelli et al. (2008) (see Sections 4.2 and 4.3), EPA analyzed alternate assumptions about exposure and the level of change in neonatal TSH

levels associated with the designation of a LOAEL or a NOAEL from this study as shown in Figure 4-7. For the NOAEL in Figure 4-7, the equivalent LOAEL (by multiplying by 10[22]) is also shown for direct comparison to the LOAEL estimates. The uncertainty considerations and the approach presented in Figure 4-7 are similar to those depicted in Figure 4-6, but the variables are different. There are several ways in which a POD could be derived from the Baccarelli et al. (2008) study. In the standard pathway RfD analysis, EPA used the study authors' regression model results from their Figure 2A (designated the "Regression Model") to determine a LOAEL based on the maternal plasma concentration corresponding to neonatal TSH levels of 5 μU/mL. The regression model was used to account for covariates that influenced the dose-response relationship. Three alternative values are examined by selecting specific points or ranges from the figures in the Baccarelli paper, without consideration of the regression modeling results (the "graphical method"). The alternative values, therefore, do not account for the covariates. The first assumes a NOAEL of 40 ppt maternal LASC, which is essentially the highest TCDD concentration above which neonatal TSH levels are consistently above 5 μU/mL [see Figure 2A in Baccarelli et al. (2008)]. The figure (2A) shows that 5 of the 6 neonates born to women who had TCDD concentrations above 40 ppt had TSH levels above 5 μU/mL; among the 45 women who had TCDD concentrations below 40 ppt, only two had babies with TSH levels above 5 μU/mL. The second alternative assumes that the 6 neonates born to women with TCDD LASC above 40 ppt comprise a LOAEL group, with a median maternal LASC of 90 ppt. The third alternative assumes a LOAEL at the highest neonatal TSH level (8.5 μU/mL) shown in Figure 2A, which corresponds to a maternal TCDD LASC of 312 ppt.

Background exposures in the population were estimated in several ways. The background TCDD exposure used in the standard pathway RfD analysis was based on continuous intake necessary to obtain 15 ppt at 30 years for females (the "Needham" background); the modeled TCDD intake was 3.9×10^{-4} ng/kg-day, slightly higher than that for males. To examine the maternal TEQ exposures associated with a LOAEL based on a neonatal TSH level of 5 μU/mL, EPA relied on the regression results reported in Baccarelli et al. (2008). Baccarelli et al. (2008) reported maternal plasma TEQ concentrations in the following two ways: (1) PCDDs, PCDFs, coplanar PCBs, without noncoplanar PCBs (see Figure 2B) and (2) PCDDs,

[22] A tenfold factor is used because the LOAEL POD is divided by a UF_L of 10 in the RfD derivation. The "equivalent" LOAEL is not meant to be an alternative LOAEL but is used strictly for comparison.

PCDFs, coplanar PCBs, and noncoplanar PCBs, termed total TEQ (see Figure 2D). The concentrations in their Figures 2B and 2D are reported as TEQs and were modeled as TCDD for this analysis. Excluding the noncoplanar PCBs, maternal TEQ levels of 219 ppt in serum are associated with neonatal TSH level of 5 μU/mL. For the total TEQ, maternal TEQ levels of 485 ppt in serum are associated with a neonatal TSH level of 5 μU/mL. Confidence in the total TEQ estimate is lower than that for the one without the noncoplanar PCBs because of the lower significance of the total TEQ regression coefficient (p = 0.14) than the one without the noncoplanar PCBs (p = 0.005).

For the standard pathway RfD analysis, the maternal "age at conception" was set at 30 years, which was the average reported in Baccarelli et al. (2008). The alternative assumes the maternal age at conception to be 45 years of age; this is the standard gestational scenario used in estimating the human equivalent doses for the animal bioassays reporting reproductive or developmental effects and is considered to be a reasonable upper end of female fertility.

The alternative LOAELs based on this analysis of Baccarelli et al. (2008) vary between 0.005 and 0.059 ng/kg-day. These two values are roughly a factor of 4 lower and a factor of 3 larger, respectively, than the LOAEL estimate of 0.020 ng/kg-day that was the basis of the standard pathway RfD. The TCDD intake of 0.0016 ng/kg-day corresponding to the alternative NOAEL is slightly more than an order of magnitude lower than the standard pathway RfD LOAEL POD and would yield a slightly lower RfD estimate than the current RfD after eliminating the 10-fold UF_L factor. EPA has much less confidence in the NOAEL estimate than in the selected LOAEL because the NOAEL does not take into account the covariates and falls in a lower concentration range where the background DLC exposures are a much more significant component. The largest downward impact on the standard pathway LOAEL POD results from grouping the highest exposures independent of the modeling results (POD = 0.005), which decreases the LOAEL by a factor of four; however, analogous to the NOAEL alternative, the approach ignores the contribution of covariates.

The largest upward impact on the standard pathway LOAEL POD is the inclusion of modeled total TEQ (POD = 0.059), which increases the LOAEL by a factor of three. However, the model fit is poor, and the result can be compared with an analogous calculation to the additive DLC approach used for the Mocarelli analysis in Figure 4-6. An additive DLC-TEQ background of 3.5×10^{-3} ng/kg-day can be estimated for the women in the Baccarelli analysis by

multiplying the TCDD background intake of 3.9×10^{-4} ng/kg-day by 9 (not shown in Figure 4-7). Adding the estimated DLC background to the standard pathway RfD LOAEL POD of 0.020 gives a corresponding total-TEQ intake of 0.023 ng/kg-day. This is 18% higher than the standard pathway RfD POD but 2.6-fold lower than the modeled total-TEQ POD. Leaving out the noncoplanar PCBs greatly improves the model fit, which could suggest that the noncoplanar PCBs do not contribute to the effect as much as the PCDDs and PCDFs or that there is greater uncertainty in the TEQ estimates for the noncoplanar PCBs. In either case, as for the Mocarelli analysis, any estimate of background DLC exposure based on TEQ is likely an over-estimate because of the conservative nature of TEFs. Overall, although background DLC exposures will effectively increase the POD to some degree, EPA believes that the effect is relatively small in the range of the estimated standard pathway TCDD LOAEL.

In summary, the quantitative uncertainties evaluated here for the RfD POD based on Baccarelli et al. (2008) span a 3- to 4-fold range in either direction. The alternative LOAELs at either extreme are not strong POD candidates; the lowest value (from the graphical method) does not account for covariates and there is greater uncertainty in the (total TEQ) regression model for the highest value than for the other regression models. All the other alternative LOAELs are within a factor of 1.5 of the RfD POD. Overall, as for Mocarelli et al. (2008) analysis, EPA believes that this sensitivity also supports the magnitude of the RfD.

4.5.1.2. *NTP (2006a) Sensitivity Analysis*

To examine the impacts of some of the uncertainties associated with estimating the POD from the NTP (2006a) study (see Section 4.2), EPA analyzed two different approaches for estimating dose and alternate choices of rodent kinetic model and background. Figure 4-8 depicts this analysis, which relied on an approach similar to those used in characterizing some of the uncertainties in the RfDs derived from Mocarelli et al. (2008) and Baccarelli et al. (2008). The lowest administered dose was determined to be the animal LOAEL based on liver and lung lesions in the rats. In the standard pathway candidate RfD analysis, the $LOAEL_{HED}$ was the POD.

Exposures were estimated either based on a kinetic model of the administered TCDD dose or on the measured concentrations of TCDD and DLCs in the rat adipose tissue after terminal sacrifice. NTP reported concentrations of TCDD, 2,3,4,7,8-pentachlorodibenzofuran

(PeCDF), and 3,3N,4,4N,5-pentachlorobiphenyl (PCB-126) in the adipose and liver tissues obtained from the rats after terminal sacrifice. The 2005 WHO TEF values for PeCDF and PCB-126 are 0.3 and 0.1, respectively (Van den Berg et al., 2006).

To predict average tissue concentrations based on the administered TCDD dose, EPA used both the Emond and CADM kinetic models. EPA also used the first-order body burden model to predict whole body TCDD concentrations; this model uses a constant half-life to simulate the elimination of TCDD from the body. Section 3 describes all of these models.

EPA used several alternative dose metrics based on the modeling approach and measured tissue concentrations. The first-order body burden model estimates the TCDD concentration in the whole body. When using the Emond model to evaluate the disposition of TCDD, EPA evaluated both whole-blood TCDD concentrations and LASC. For the CADM model, EPA simulated TCDD concentrations in the adipose compartment following the administered TCDD dose. EPA also used the TCDD (see Table 13 in the NTP report) or DLC concentrations (see Tables 10 and 11 in the NTP (2006c) report) measured in the adipose tissue collected at study termination.

Using the DLC concentration information, EPA estimated TEQ in two ways. In the first approach, based on an analysis of DLCs in the adipose tissue that was reported in another NTP study on DLC mixtures (NTP, 2006c), EPA initially estimated the ratio of the adipose tissue TEQ concentration to the adipose tissue TCDD concentration, then applied this ratio to the Emond whole-blood TCDD estimates assuming proportionality (resulting in a LOAEL whole blood concentration of 2.75 ppt instead of the TCDD-only concentration of 2.56 ppt).

In the second approach, EPA estimated TEQ dose based on adipose tissue TCDD levels reported by NTP; the reported TCDD concentration in the fat given in the study at the lowest dose was used to estimate a LOAEL using the Emond model. Finally, using the 2005 WHO TEF values (Van den Berg et al., 2006), EPA converted the reported concentrations of TCDD, PeCDF, and PCB-126 measured in the fat of the control rats in the NTP mixtures study (NTP, 2006c) to TEQ using eq. 4-1.

$$Chemical_i(B) = \frac{Chemical_i\left(fat_{MC}\right)\ \ TEF_i}{TCDD\left(fat_{TCDD}\right)}\ \ Dose_{TCDD} \qquad \text{(Eq. 4-1)}$$

Where

$Chemical_i(B)$ = estimate of background exposure to Chemical *i* in ng/kg units of TCDD blood concentrations at 105 weeks, for *i* = TCDD, PeCDF, and PCB126.

$Chemical_i(fat_{MC})$ = mean pg/g of Chemical *i* in the fat tissues of the control animals at 105 weeks in mixtures study (NTP, 2006c).

$TCDD(fat_{TCDD})$ = mean pg/g of TCDD in the fat tissues of the 3 ng/kg dose group at 105 weeks in the TCDD study (NTP, 2006a).

$Dose_{TCDD}$ = 2.56 ng/kg TCDD blood concentration for the 3 ng/kg dose group in the TCDD study (NTP, 2006a).

TEF_i = Toxicity Equivalence Factor for Chemical *i* [from Van den berg et al. (2006)].

Assuming simple proportionality of blood TCDD concentrations between controls and low-dose (2.14 ng/kg-day) animals, the TEF-adjusted ratio of each congener (Chemical *i*) in control animal fat to low-dose-animal fat is multiplied by the modeled TCDD blood concentration for the low-dose animals to obtain an equivalent background exposure in the dose metric (ppt whole blood). For total TEQ, the estimates of all three congeners are summed. Total TEQ estimates likely are biased somewhat high because they are based on terminal (2-year) measurements rather than representing lifetime averages.

Overall, the alternative LOAEL estimates in this tree (see Figure 4-8) vary between 0.023 and 0.44 ng/kg-day. The LOAEL for the standard pathway RfD was estimated to be 0.14 ng/kg-day and is at the lower end of the range. The alternative LOAEL based on first order body burden (0.023 ng/kg-day) is the lowest value in the range, approximately 85% lower than the LOAEL based on the standard pathway approach. The difference between these two estimates is consistent with the more conservative approach used in modeling first-order TCDD body burdens. The alternative LOAEL based on the TEQ in whole blood is less than 10% greater than the LOAEL from the standard pathway RfD. The alternative candidate LOAEL based on the TCDD in lipid-adjusted serum is approximately 120% greater than the LOAEL for the standard pathway RfD. The use of the CADM model to estimate adipose tissue concentration based on administered dose resulted in a 35% increase in the LOAEL estimate relative to the LOAEL based on the standard pathway approach. The LOAELs based on measured TCDD or TEQ levels in rodent adipose tissue were greater than the LOAEL from the standard pathway RfD by approximately a factor of three.

4.5.2. Evaluation of Range of Alternative PODs for Additional Epidemiological Endpoints

In addition to the principal studies depicted in Figures 4-6 and 4-7, EPA evaluated a number of endpoints presented in seven other Seveso cohort studies to estimate the range of potential PODs based on uncertainties in exposure duration, exposure averaging protocols, and DLC background exposures. Included in those study/endpoint combinations are the following: two that passed all the selection criteria, developmental dental effects (Alaluusua et al., 2004) and duration of menstrual period (Eskenazi et al., 2002b); a new developmental study on semen quality (Mocarelli et al., 2011) that was published after the study selection process was completed but is useful in this uncertainty analysis of the POD ranges; and four studies that did not pass all the criteria for qualification as POD candidates (Warner et al., 2007; Eskenazi et al., 2005; Warner et al., 2004; Mocarelli, 2000), but for which limiting NOAEL and LOAEL values can be estimated. Descriptions and evaluations of each of these studies can be found in Appendix C. Tables 4-8 through 4-10 and Figure 4-9 present the exposure values modeled using the Emond human PBPK model for potential POD ranges for 7 additional endpoints studied in the Seveso cohort.[23] For most of the studies that did not pass all the criteria, the major uncertainties are the definition of the critical exposure window and the corresponding relevant exposure-averaging time, and the determination of adverse effect levels. Eskenazi et al. (2002b) passed the selection criteria because a critical exposure window could be identified, but the determination of an adverse effect level for length of menstrual cycle is somewhat arbitrary. A critical exposure window can be identified also for Warner et al. (2004) (age at menarche), but no TCDD-related adverse health outcomes were observed. However, with some additional assumptions, NOAELs and LOAELs at nominal group-exposure levels can be determined for each of these studies. The critical exposure window is assumed to be the entire duration from exposure in 1976 to time of interview (i.e., end of follow-up period) when a critical window cannot be identified. Tentative NOAELs and LOAELs are designated for those endpoints where adversity levels are difficult to define. Given these assumptions, TCDD and total TEQ intakes can be modeled but must be considered to be lower bounds on the effective exposures, given the

[23] The details of the kinetic modeling for these endpoints and the corresponding background exposures can be found in Appendix F.

conservative nature of the assumptions; EPA does not consider these estimates suitable for use in the derivation of the TCDD RfD.

Additional endpoints reported in the epidemiologic literature were considered in the context of this uncertainty analysis but were excluded based on large uncertainties in defining adversity or plausible exposure profiles over time. All the Ranch Hand studies[24] were excluded because of the inability to construct effective exposure profiles with any confidence, given the 20-year lag between the actual TCDD exposures and measurement of serum levels. For the Seveso cohort, several studies[25] were eliminated from consideration because uncertainties in defining plausible NOAELs or LOAELs were too large.

For modeling of the endpoints in Tables 4-8 to 4-10, grouped exposure ranges were represented by the geometric mean of the range limits. The average daily intakes for exposures (LASC) in the background range were estimated as the continuous exposure from birth resulting in the reported serum concentrations (TCDD or total TEQ) at the average subject age at time of measurement. Peak and critical-window average exposures (as LASC) were modeled for measured LASC values greater than background using the actual exposure scenarios. Because all exposure durations were less than lifetime, average daily intakes for all modeled peak and window-average LASC were estimated using the terminal 5-year-peak average as described in Section 3.3.6. Precision is expressed to the nearest 10^{-5} ng/kg-day for all intake estimates to avoid rounding errors when adding DLC background intakes. Values less than or equal to 10^{-3} are shown in scientific notation for readability.

Figure 4-9 shows the range of NOAELs and LOAELs and exposures for all of the endpoints considered in this uncertainty analysis, the endpoints on which they are based, and the study citation. The study/endpoint combinations are separated into two groups representing either those chosen for RfD POD consideration ("Candidate RfD") or those not otherwise qualifying ("Uncertainty Analysis Only"). The NOAELS and LOAELS are indicated for each study, as appropriate, and the vertical lines through these PODs represent the range of possible PODs based on Emond PBPK results using alternative exposure scenarios. The limits—indicated by symbols of the same type—for each POD type (NOAEL or LOAEL) for

[24] (Michalek and Pavuk, 2008; Pavuk et al., 2003; Michalek et al., 2001a; Michalek et al., 2001b; Michalek et al., 2001c; Longnecker and Michalek, 2000)
[25] (Eskenazi et al., 2007; Baccarelli et al., 2005; Baccarelli et al., 2004; Eskenazi et al., 2003; Landi et al., 2003; Baccarelli et al., 2002; Eskenazi et al., 2002a)

each endpoint cover the full range of alternative PODs in Tables 4-8 to 4-10, without distinction of the relative plausibility of each one. That is, all the PODs are treated equally without considering the relative confidence held in each one, individually. The low end of most of the ranges is the critical-window average exposure, which does not take into account the influence of the much higher peak exposure. Conversely, the upper end of the range is generally the peak exposure, which does not account for the potential effect of longer-term continuous exposure. On the "uncertainty analysis only" side of Figure 4-9, most of the NOAELs and many of the LOAELs are somewhat speculative and would not be considered as strong candidates for the RfD POD. The range limits are themselves uncertain, as constraints were applied to the lower and upper limits to keep them in the range of the data. The same DLC modeling issues presented in Section 4.5.1 apply to all the TEQ results here, so the TEQ results are approximations and are unlikely to be very accurate. Also, the lowest POD estimates are more affected by background DLC exposure than are the PODs closer to the RfD POD; generally, TCDD is a minor component of the total TEQ for the lower PODs, subjecting the lowest alternative PODs to the greatest uncertainty. The RfD LOAEL POD (0.02 ng/kg-day) and its equivalent NOAEL estimate (0.002 ng/kg-day, with the 10-fold UF), along with the RfD (7×10^{-4} ng/kg-day), are shown on the figure for comparison to the alternative POD ranges.

The LOAEL ranges for the two principal studies (Baccarelli et al., 2008; Mocarelli et al., 2008) span the RfD LOAEL POD, whether based on TCDD alone or total TEQ. The single NOAEL estimate for Baccarelli et al. (2008) is only slightly below the equivalent RfD NOAEL POD. The NOAEL and the lowest alternative LOAELs for Baccarelli et al. (2008) are not strong POD candidates because they are based on the raw observations and do not take into account the covariates that affect the exposure-response relationship, as does the regression model on which the RfD LOAEL POD is based. In general, background DLC exposure has a small impact on the LOAEL PODs for the co-principal studies, raising the effective exposure level by 15% for the Mocarelli et al. (2008) RfD LOAEL POD and yielding essentially the same value for the Baccarelli et al. (2008) RfD LOAEL POD, if noncoplanar PCBs are excluded (see Figure 4-7). Including the noncoplanar PCBs from the Baccarelli et al. (2008) regression modeling results has a much bigger impact, raising the LOAEL by a factor of 3; however, the significance of the modeled slope is relatively poor ($p = 0.14$), so EPA does not place much biological significance on the outcome. The POD ranges for the other candidate RfD endpoints are well above their

respective comparison NOAEL/LOAEL benchmarks. The NOAEL for Eskenazi et al. (2002b) is somewhat arbitrary, based simply on a continuous average exposure over a 13-year window corresponding to a normal 28-day menstrual cycle, without considering the possible range of normal durations.

Of the endpoints that were not selected as RfD POD candidates, there are three whose LOAEL ranges are wholly or mostly below the RfD LOAEL POD. The sperm effects in men who were exposed in utero and by lactation reported by Mocarelli et al. (2011) are very similar to those in men exposed as boys in one of the principal studies (Mocarelli et al., 2008). The maternal exposures associated with the effects reported by Mocarelli et al. (2011) are very low with the TCDD-only LOAEL being 12-fold lower than the RfD LOAEL POD for the 30-year exposure scenario. For this study, a TCDD-only NOAEL can be established at 2.9×10^{-4} ng/kg-day (for the reference population), which is sevenfold below the equivalent RfD NOAEL POD. Both the TCDD-only NOAEL and LOAEL are much lower than the estimated DLC background exposure; however, assuming a simple TEQ additive model, and with the aforementioned uncertainties concerning DLC-TEQ estimation, a TEQ NOAEL and LOAEL of 2.9×10^{-3} and 5.4×10^{-3} ng/kg-day can be estimated (Table 4-8). Although the TEQ LOAEL is still well below that for the RfD POD, the TEQ NOAEL is in the range of the equivalent RfD NOAEL POD. Given the large amount of uncertainty in the modeled NOAEL and LOAEL for this endpoint, EPA elected not to consider either as a POD.

The second endpoint with lower LOAELs than the RfD POD is age at menopause reported by Eskenazi et al. (2005). The figure for this endpoint includes two separate LOAEL candidates because of uncertainty in determining adversity at the lower exposure level in question (3rd quintile). For that reason, the daily intakes associated with the critical-window average and peak exposures are labeled ("W" and "P," respectively). The intakes associated with the peak are in the range of the RfD LOAEL benchmark, while the window-average TCDD intakes are closer to the NOAEL benchmark. Considering background DLC intake, the window-average TEQ intakes are considerably higher, the DLC exposures being larger than the TCDD intakes, themselves, but still below the LOAEL benchmark. The range of the TEQ P/W average of 0.01−0.031 ng/kg-day (see Table 4-10), however, straddles the RfD LOAEL benchmark of 0.02 ng/kg-day. Uncertainty in the NOAEL is similar to that for the LOAEL, depending on whether the 1st or 2nd quintile can be called a NOAEL. Although the response in

the 2nd quintile is not significant compared to the 1st quintile, the NOAEL determination is complicated by the lack of an absolute measure of "normal." EPA considered the quantitative and qualitative uncertainties to be too large to consider this endpoint as an RfD POD candidate.

The NOAELs and LOAELs for altered sex ratio reported by Mocarelli et al. (2000) span their respective RfD POD benchmarks and are above the benchmarks when considering the peak/window exposure averages or background DLC exposures. The uncertainties for lack of an identifiable critical exposure window also apply to this endpoint. The other two endpoints, age at menarche (Warner et al., 2004) and ovarian function (Warner et al., 2007), are unbounded NOAELs at the highest exposures. The ovarian function endpoint also is uncertain for lack of an identifiable critical exposure window.

Additional uncertainties not covered explicitly in this analysis include exposure to other AhR agonists, either naturally occurring in food-stuffs (Connor et al., 2008) or by-products of combustion or manufacturing processes (e.g., poly-aromatic hydrocarbons), and choice of uncertainty factor. As a final note on background DLC exposure, the background DLC intake estimates for the standard scenario (Needham) used in this assessment are somewhat crude, in that they are simple multiples of modeled TCDD intake based on an approximation of the proportion of TCDD to total TEQ. TCDD exposures are modeled over durations of up to 35 years (1941−1976) using a single fixed background intake term (a model limitation). However, background TCDD/TEQ exposures are thought to have varied widely over that time period, increasing gradually in the United States from the early 20th century to a peak in 1965, then decreasing rapidly to near current levels in the early 1980s (Lorber, 2002). Based on a digitization of Figure 6 in Lorber (2002), depicting the estimated TEQ intake over the course of the 20th century, a time-weighted average total TEQ intake for the period 1941−1976 of 4.6×10^{-3} ng/kg-day can be estimated. Adjusting the TEF_{98}-based Lorber (2002) TEQ intakes to TEF_{05}-based values, assuming a 10% TCDD fraction and using the 0.7 TEF_{05}:TEF_{98} factor described previously (see Section 4.5.1), yields a DLC-TEQ intake estimate of 3.4×10^{-3} ng/kg-day for that time period, which is similar to the estimated DLC background intake of 3.33×10^{-3} ng/kg-day for the standard scenario using the simple scaling model.

However, the DLC intake estimate based on Lorber (2002) is somewhat of an underestimate because it does not include dioxin-like PCBs. Pinsky and Lorber (1998) estimated a TCDD intake of 4×10^{-4} ng/kg-day for the U.S. population in the 1970s, which is almost the

same as the modeled TCDD background intake for the Seveso population. However, there is no information on comparative environmental exposures for the United States and Italy during this period, and TCDD exposures before 1970 for these populations were not necessarily the same, on average. Higher TCDD background exposures have been estimated by others. Pinsky and Lorber (1998) estimated an average TCDD-only intake of 1.4×10^{-3} to 1.9×10^{-3} ng/kg-day for the U.S. population in the late 1960s and early 1970s using a 1st-order kinetics model with a variable intake term and a TCDD half-life of 7.1 years. Aylward and Hays (2002) estimated a TCDD intake of at least 1.3×10^{-3} ng/kg-day for the United States, Canada, Germany, and France prior to 1972 using a 1st-order kinetics model assuming a TCDD half-life of 7.5 years. These estimates are 3.5−5 times higher than the background TCDD intake estimated by EPA using the Emond PBPK model for this assessment. Total TEQ background would increase proportionally. However, none of these estimates, including EPA's, is based on actual intake measurements and are all dependent on modeling assumptions. Raising the background DLC exposure would obviously increase the effective PODs. However, increasing the background TCDD intake for modeling purposes would decrease the contribution of the actual TCDD exposures experienced by the Seveso population in 1976, resulting in a lower TCDD POD, as can be seen in the Eskenazi background scenario for Mocarelli et al. (2008) (see Figure 4-6). The overall result would be a slightly higher POD (ca., 0.032 ng/kg-day) based on TEQ.

This analysis highlights several important research needs. While the disposition of TCDD following high exposures is reasonably understood and simulated in current models, the current scientific understanding of disposition following TCDD exposures that are closer to current background dietary intakes, likely the primary source of TCDD exposure for most of the U.S. population, is not understood as well at present. This uncertainty affects the estimation of TCDD intake rates corresponding to the lower blood TCDD levels associated with LOAELs and NOAELs. The disposition of DLCs following exposures at background levels is similarly not well understood. Furthermore, there is uncertainty in the relationship of DLC tissue concentrations to oral intakes in the current TEF approach. Finally, there is toxicological uncertainty regarding several of the endpoints. Additional studies corroborating these outcomes and their toxicological significance would further increase their utility in refining the TCDD RfD.

Overall, EPA believes that the results of this analysis of alternative endpoints and PODs increase the confidence in the TCDD RfD, both qualitatively and quantitatively. EPA's analyses of some studies show POD estimates higher than the RfD PODs—primarily those analyses that consider background DLCs. Other analyses show POD estimates lower than the RfD POD, such as the use of alternative age-adjusted background TCDD/DLC intake rates and some evaluations of more uncertain endpoints (e.g., age at menopause endpoint in Eskanazi et al. (2005). The more extreme values on the lower end are also the most uncertain, particularly with respect to the contribution of TCDD relative to total TEQ. In addition, except for the male reproductive effects in Mocarelli et al. (2011), determination of adversity for the lower LOAELs is problematic, leading to lower confidence in the PODs. The TCDD and TEQ LOAELs for semen quality in males exposed in utero and by lactation (Mocarelli et al., 2011) are much lower than the corresponding LOAELs for males exposed between ages 1 and 10 years (Mocarelli et al., 2008). However, the NOAEL established for in utero and lactational exposure is fairly strong in the qualitative sense; that is, there is fairly clear indication that semen quality is unaffected at the corresponding dioxin exposure level. Quantitatively, there is more uncertainty, but considering background DLC exposure, the NOAEL is close to the RfD NOAEL benchmark.

Table 4-1. PODs for epidemiologic studies of TCDD

Study	POD (ng/kg-day)	Critical effects
Alaluusua et al. (2004)	0.0406[a] (NOAEL)	Dental effects in adults exposed to TCDD in childhood
Baccarelli et al. (2008)	0.0199[b] (LOAEL)	Elevated TSH in neonates
Mocarelli et al. (2008)	0.0201[c] (LOAEL)	Decreased sperm count and motility in men exposed to TCDD in childhood

[a]Mean of peak exposure (0.0655 ng/kg-day) and average exposure over 10-year critical window (0.0156 ng/kg-day).
[b]Maternal exposure corresponding to neonatal TSH concentration exceeding 5 µU/mL.
[c]Mean of peak exposure (0.032 ng/kg-day) and average exposure over 10-year critical window (0.0080 ng/kg-day).

Table 4-2. Models run for each study/endpoint combination in the animal bioassay benchmark dose modeling

Model	Restrictions imposed
Continuous models	
Exponential M2−M5, not grouped	Adverse direction specified according to the response data; power ≥1
Hill	Adverse direction is automatic; $n > 1$
Linear	Adverse direction is automatic; degree of polynomial = 1
Polynomial	Adverse direction is automatic; degree of polynomial unrestricted; restrict the sign of the power to nonnegative or nonpositive, depending on the direction of the responses
Power	Adverse direction is automatic; power ≥1
Dichotomous models	
Gamma	Power ≥1
Logistic	None
Log-Logistic	Slope ≥1
Log-Probit	None
Multistage	Beta ≥0, 2nd degree polynomial
Probit	None
Weibull	Power ≥1

Table 4-3. Summary of key animal study PODs (ng/kg-day) based on three different dose metrics: administered dose, first-order body burden HED, and blood concentration

Study	Endpoint	Administered dose[a]			1st-order body burden HED[b]			Blood concentration HED[c]		
		NOAEL	LOAEL	BMDL[d]	NOAEL	LOAEL	BMDL[d]	NOAEL	LOAEL	BMDL[d]
Amin et al. (2000)	Saccharin preference ratio, female	–	2.50E+01	–[e]	–	2.49E−02	–[e]	–	1.71E−01	–[e]
Bell et al. (2007b)	Balano-preputial separation in male pups	–	2.40E+00	2.87E+00	–	1.26E−02	1.50E−02	–	8.85E−02	4.34E−02
Bowman et al.(1989a; 1989b); Schantz and Bowman (1989); Schantz et al. (1986); Schantz et al. (1992)	Neurobehavioral effects	–	1.20E−01	–	–	8.22E-03	–	–	–	–
Cantoni et al. (1981)	Urinary coproporhyrins	–	1.43E+00	–[e]	–	1.24E−02	–[e]	–	6.37E−02	–[e]
Chu et al. (2001)	Tissue-weight changes	2.50E+02	1.00E+03	–	7.55E−01	3.02E+00	–	7.03E+00	2.96E+01	–
Chu et al. (2007)	Liver lesions	2.50E+00	2.50E+01	–	7.55E−03	7.55E−02	–	3.49E−02	5.63E−01	–
Crofton et al. (2005)	Serum T4	3.00E+01	1.00E+02	–[e]	1.92E−02	6.40E−02	–[e]	1.69E−01	7.43E−01	–[e]
Croutch et al. (2005)	Decreased body weight	5.43E+01	2.17E+02	–	2.22E−01	8.89E−01	–	7.81E-01	3.57E+00	–
DeCaprio et al. (1986)	Decreased body weight, organ-weight changes	6.10E−01	4.90E+00	–	4.11E−03	3.30E−02	–	–	–	–
Fattore et al. (2000)	Decreased hepatic retinol	–	2.00E+01	–	–	1.23E−01	–	–	7.82E−01	–
Fox et al. (1993)	Increased liver weight	5.70E−01	3.27E+02	–	1.42E−03	8.12E−01	–	8.08E-04	3.05E+00	–

Table 4-3. Summary of key animal study PODs (ng/kg-day) based on three different dose metrics: administered dose, 1st-order body burden HED and blood concentration HED (continued)

Study	Endpoint	Administered dose[a]			1st-order body burden HED[b]			Blood concentration HED[c]		
		NOAEL	LOAEL	BMDL[d]	NOAEL	LOAEL	BMDL[d]	NOAEL	LOAEL	BMDL[d]
Franc et al. (2001)	Organ-weight changes	1.00E+01	3.00E+01	1.34E+01	6.62E−02	1.99E−01	8.87E−02	4.49E−01	1.41E+00	2.61E−01
Franczak et al. (2006)	Abnormal estrous cycle	–	7.14E+00	–	–	5.95E−02	–	–	3.18E−01	–
Hojo et al. (2002)[f]	DRL response per min	–	2.00E+01	–[e]	–	5.26E−03	–[e]	–	5.51E−02	–[e]
Hochstein et al. (2001)[g]	Kit mortality at 6 wk	–	2.65E+00	–	–	–	–	–	–	–
Hutt et al. (2008)	Embyrotoxicity	–	7.14E+00	–	–	4.67E-02	–	–	2.52E-01	–
Ikeda et al. (2005)	Sex ratio	–	1.65E+01	–	–	1.05E−01	–	–	2.75E+00	–
Ishihara et al. (2007)	Sex ratio	1.00E−01	1.00E+02	–	3.18E−04	3.18E−01	–	4.91E-05	4.96E-01	–
Kattainen et al. (2001)	3rd molar length	–	3.00E+01	–[e]	–	7.89E−03	–[e]	–	9.01E−02	–[e]
Keller et al. (2008a; 2008b; 2007)	Missing mandibular molars	–	1.00E+01	–[e]	–	2.58E−03	–[e]	–	9.88E−03	–[e]
Kociba et al. (1976)	Liver and hematologic effects and body-weight changes	7.14E+00	7.14E+01	–	4.53E−02	4.53E−01	–	2.62E−01	3.03E+00	–
Kociba et al. (1978)	Liver and lung lesions, increased urinary porphyrins	1.00E+00	1.00E+01	–[e]	1.07E−02	1.07E−01	–[e]	6.33E−02	6.34E−01	–[e]
Kuchiiwa et al. (2002)	Immunoreactive neurons	–	7.00E-01	–	–	3.11E-03	–	–	2.75E-03	–[e]
Latchoumycandane and Mathur (2002)[h]	Sperm production	–	1.00E+00	–[e]	–	3.87E−03	–[e]	–	1.62E−02	–[e]
Li et al. (1997)	Increased serum FSH	3.00E+00	1.00E+01	–[e]	7.89E−04	2.63E−03	–[e]	2.90E−03	1.67E−02	–[e]
Li et al. (2006)	Hormone levels (serum estradiol)	–	2.00E+00	–[e]	–	9.85E−04	–[e]	–	1.58E−03	–[e]

Table 4-3. Summary of key animal study PODs (ng/kg-day) based on three different dose metrics: administered dose, 1st-order body burden HED and blood concentration HED (continued)

Study	Endpoint	Administered dose[a]			1st-order body burden HED[b]			Blood concentration HED[c]		
		NOAEL	LOAEL	BMDL[d]	NOAEL	LOAEL	BMDL[d]	NOAEL	LOAEL	BMDL[d]
Markowski et al. (2001)	FR2 revolutions	–	2.00E+01	–[e]	–	6.25E−03	–[e]	–	5.15E−02	–[e]
Maronpot et al. (1993)	Increased relative liver weight	1.07E+01	3.50E+01	–	8.97E−02	2.93E−01	–	5.03E-01	1.71E+00	–
Miettinen et al. (2006)	Cariogenic lesions in pups	–	3.00E+01	–[e]	–	7.89E−03	–[e]	–	8.95E−02	–[e]
Murray et al. (1979)	Fertility index in F2 generation	1.00E+00	1.00E+01	–[e]	9.43E−03	9.43E−02	–[e]	2.89E−02	3.79E−01	–[e]
NTP (1982b)	Liver lesions	–	1.39E+00	–[e]	–	6.47E−03	–[e]	–	2.16E−02	–[e]
NTP (2006a)	Liver and lung lesions	–	2.14E+00	–[e]	–	2.34E−02	–[e]	–	1.36E−01	–[e]
Nohara et al. (2000)	Decreased spleen cellularity	8.00E+02	–	–	2.10E−01	–	–	5.34E+00	–	–
Nohara et al. (2002)	Mortality from influenza virus-A challenge	5.00E+02	–	–	1.29E-01	–	–	1.37E+00	–	–
Ohsako et al. (2001)	Anogenital distance in pups	1.25E+01	5.00E+01	–[e]	3.29E−03	1.32E−02	–[e]	2.74E−02	1.78E−01	–[e]
Schantz et al. (1996)	Maze errors	–	2.50E+01	–[e]	–	–[e]	4.55E-02	–	1.71E-01	–[e]
Seo et al. (1995)	Decreased thymus weight	2.50E+01	1.00E+02	–	2.49E−02	9.96E−02	–	1.67E−01	9.15E−01	–
Sewall et al. (1995)	Serum T4	1.07E+01	3.50E+01	5.16E+00	8.97E−02	2.93E−01	4.33E−02	5.03E−01	1.71E+00	1.80E−01
Shi et al. (2007)	Serum estradiol in female pups	1.43E−01	7.14E−01	2.24E−01	1.23E−03	6.13E−03	1.92E−03	4.47E−03	2.69E−02	4.74E−03

Table 4-3. Summary of key animal study PODs (ng/kg-day) based on three different dose metrics: administered dose, 1st-order body burden HED and blood concentration HED (continued)

Study	Endpoint	Administered dose[a]			1st-order body burden HED[b]			Blood concentration HED[c]		
		NOAEL	LOAEL	BMDL[d]	NOAEL	LOAEL	BMDL[d]	NOAEL	LOAEL	BMDL[d]
Simanainen et al. (2002)	Decreased serum T4	1.00E+02	3.00E+02	–	2.63E−02	7.89E−02	–	4.26E-01	1.67E+00	–
Simanainen et al. (2003)	Decreased thymus weight and change in EROD activity	1.00E+02	3.00E+02	–	2.63E−02	7.89E−02	–	4.26E-01	1.67E+00	–
Simanainen et al. (2004a)	Decreased daily sperm production	1.00E+02	3.00E+02	–	2.63E-02	7.89E-02	–	4.26E-01	1.67E+00	–
Smialowicz et al. (2004)	Decreased antibody response to SRBCs	3.00E+02	1.00E+03	–	7.73E−02	2.58E−01	–	7.23E-01	3.28E+00	–
Smialowicz et al. (2008)	PFC per 10^6 cells	–	1.07E+00	–[e]	–	5.00E−03	–[e]	–	6.26E−03	–[e]
Smith et al. (1976)	Cleft palate in pups	1.00E+02	1.00E+03	1.84E+02	1.59E-01	1.59E+00	2.93E-01	5.24E-01	7.61E+00	9.46E-01
Sparschu et al. (1971)	Decreased fetal body weight	3.00E+01	1.25E+02	–[e]	5.45E-02	2.27E-01	-	3.18E-01	1.73E+00	–[e]
Toth et al. (1979)	Skin lesions	–	1.00E+00	–[e]	–	3.70E−03	–[e]	–	9.91E−03	–[e]
VanBirgelen et al. (1995a)[i]	Decreased liver retinyl palmitate	–	1.35E+01	–[e]	–	8.32E−02	–[e]	–	5.14E−01	–[e]
Vos et al. (1973)	Decreased delayed-type hypersensitivity response to tuberculin	1.14E+00	5.71E+00	–	6.43E−03	3.22E−02	–	–	–	–
Weber et al. (1995)	Increased liver weight	1.00E+03	3.00E+03	–	3.51E-01	1.05E+00	–	3.27E+00	1.18E+01	–
White et al. (1986)	Decreased serum complement	–	1.00E+01	–[e]	–	2.23E−02	–[e]	–	2.77E−02	–[e]
Yang et al. (2000)	Increased endometrial implant survival	1.79E+01	–	–	6.74E−01	–	–	–	–	–

Table 4-3. Summary of key animal study PODs (ng/kg-day) based on three different dose metrics: administered dose, 1st-order body burden HED and blood concentration HED (continued)

[a]Average administered daily dose over the experimental exposure period.
[b]HED based on 1st-order body burden model described in Section 3.2.4.4.
[c]HED based on Emond rodent and human PBPK models described in Section 3.3.6.
[d]BMR = 0.1 for quantal endpoints and 1 standard deviation control mean for continuous endpoints, except for body and organ weights, where BMR = 10% relative deviation from control mean.
[e]BMD modeling unsuccessful (see Table 4-4 and Appendix G for details).
[f]Zareba et al. (2002) is considered to be the same study but report effects at doses above the LOAEL that are not considered further; this study is not carried forward for determination of an RfD POD but is included in the RfD uncertainty analysis presented in Section 4.4.
[g]Hochstein et al. (2001) is not carried forward because of the lack of toxicokinetic information for estimation of an HED.
[h]Latchoumycandane et al. (2002a; 2002b) are considered to be the same study but report effects (not toxicologically relevant) at doses above the LOAEL that are not considered further; these two studies are not carried forward.
[i]Van Birgelen et al. (1995b) is considered to be the same study but reports effects at doses above the LOAEL that are not considered further; this study in not carried forward for determination of an RfD POD but is included in the RfD uncertainty analysis presented in Section 4.4.

− value not established or not modeled.

Table 4-4. TCDD BMDL analysis (NOAEL, LOAEL, BMD, and BMDL values given as animal whole blood concentrations in ng/kg)[a]

Study[b,c]	NOAEL/ LOAEL	Endpoint	Control response	First response[d]	Max response[e]	Model fit detail	BMD/ BMDL	Comments
Amin et al. (2000) (rat)	– 3.38E+00	Saccharin consumed, female, (0.25%) (n = 10)	–	22% ↓ (0.3 SD)	66% ↓	Continuous linear, modeled variance (p = 0.55)	9.15E+00 6.09E+00	BMDL > LOAEL; restricted power model, constrained parameter hit lower bound
						Continuous power, modeled variance, unrestricted (p = NA)	8.37E+00 3.42E+00	Saturated model; supralinear fit (power = 0.74)
		Saccharin consumed, female (0.50%) (n = 10)	–	49% ↓ (0.7 SD)	80% ↓	Continuous linear, modeled variance (p = 0.06)	1.02E+01 6.57E+00	Restricted power model, constrained parameter hit lower bound
						Continuous power, modeled variance, unrestricted (p = NA)	6.57E+00 1.15E+00	Saturated model; supralinear fit (power = 0.40)
		Saccharin preference ratio, female (0.25%) (n = 10)	–	29% ↓ (1.8 SD)	33% ↓	Continuous linear, modeled variance (p = 0.002)	1.16E+01 5.57E+00	BMDL > LOAEL; no response near BMR; near maximal response at LOAEL
		Saccharin preference ratio, female (0.50%) (n = 10)	–	39% ↓ (1.1 SD)	54% ↓	Continuous linear, constant variance (p = 0.14)	8.14E+00 5.11E+00	BMDL > LOAEL; near maximal response at LOAEL; restricted power model, constrained parameter hit lower bound
						Continuous power, constant variance, unrestricted (p = NA)	2.60E+00 1.06E−14	Saturated model; supralinear fit (power = 0.28)

Table 4-4. TCDD BMDL analysis (NOAEL, LOAEL, BMD, and BMDL values given as animal whole blood concentrations in ng/kg)[a] (continued)

Study[b,c]	NOAEL/ LOAEL	Endpoint	Control response	First response[d]	Max response[e]	Model fit detail	BMD/ BMDL	Comments
Bell et al. (2007b) (rat)	– 2.20E+00	Balano-preputial separation in male pups (n = 30 [dams])	1/30	5/30	15/30	Dichotomous log-logistic, restricted (p = 0.78)	2.25E+00 1.39E+00	Adequate fit; constrained parameter bound hit; not litter based; selected
						Dichotomous log-logistic, unrestricted (p = 0.50)	2.00E+00 2.80E−01	Supralinear fit (slope = 0.93); selected
Cantoni et al. (1981) (rat)	– 1.85E+00	Urinary uroporhyrins (n = 4)	—	2.4-fold ↑ (5.7 SD)	87-fold ↑	Continuous exponential (M2), modeled variance (p = 0.0003)	3.76E+00 2.76E+00	No response near BMR; poor fits for all modeled variance models; constant variance poor representation of control SD; BMDL > LOAEL
		Urinary coproporhyrins (n = 4)	—	2.4-fold ↑ (3.1 SD)	4.0-fold ↑	Continuous exponential (M4), modeled variance (p = 0.49)	5.34E−01 1.80E−01	No response near BMR
						Continuous power, modeled variance, unrestricted (p = 0.61)	2.77E−02 2.03E−05	Supralinear fit (n = 0.30); poor model choice for plateau effect
Crofton et al. (2005) (rat)	3.46E+00 9.26E+00	Serum T4, (n = 4−14)	—	29% ↓ (1.9 SD)	51% ↓	Continuous exponential (M4), constant variance (p = 0.94)	5.19E+00 3.03E+00	No response near BMR

Table 4-4. TCDD BMDL analysis (NOAEL, LOAEL, BMD, and BMDL values given as animal whole blood concentrations in ng/kg)[a] (continued)

Study[b,c]	NOAEL/ LOAEL	Endpoint	Control response	First response[d]	Max response[e]	Model fit detail	BMD/ BMDL	Comments
Franc et al. (2001) (rat)	6.59E+00 1.45E+01	S-D Rats, Relative Liver Weight	–	8.1% ↑ (0.58 SD)	55% ↑	Continuous power, constant variance (p = 0.84)	9.47E+00 4.59E+00	Acceptable fit; selected
		L-E Rats, Relative Liver Weight	–	6.3% ↑ (0.63 SD)	22% ↑	Continuous Hill, modeled variance, restricted (p = 0.83)	7.72E+00 1.22E+00	Constrained parameter hit lower bound; poor fit for variance model
						Continuous Hill, modeled variance, unrestricted (p = N/A)	7.22E+00 1.15E+00	Supralinear fit (power = 0.55)
		S-D Rats, Relative Thymus Weight	–	9.0% ↓ (0.11 SD)	77% ↓	Continuous exponential (M4), modeled variance (p = 0.72)	1.88E+00 9.22E−01	Poor fit for responses in controls and lowest exposure group
						Continuous polynomial, modeled variance (p = 0.40)	4.78E+00 3.89E+00	No response near BMR; otherwise acceptable fit
		L-E Rats, Relative Thymus Weight	–	7.7% ↓ (0.15 SD)	66% ↓	Continuous exponential (M4), constant variance (p = 0.23)	2.08E+00 5.93E−01	Poor fit for responses in controls and lowest exposure group; dose-response relationship not significant
		H/W Rats, Relative Thymus Weight	–	3.7% ↓ (0.10 SD)	51% ↓	Continuous exponential (M2), constant variance (p = 0.70)	5.09E+00 3.13E+00	No response near BMR; otherwise acceptable fit

Table 4-4. TCDD BMDL analysis (NOAEL, LOAEL, BMD, and BMDL values given as animal whole blood concentrations in ng/kg)[a] (continued)

Study[b,c]	NOAEL/ LOAEL	Endpoint	Control response	First response[d]	Max response[e]	Model fit detail	BMD/ BMDL	Comments
Hojo et al. (2002) (rat)	– 1.62E+00	DRL reinforce per min (n = 12)	–	55% ↑ (1.0 SD)	80% ↑	Continuous exponential (M4), constant variance (p = 0.054)	1.32E+00 2.37E−03	Poor fit; near maximal response at lowest dose, BMD/BMDL ratio >100
		DRL response per min (n = 12)	–	105% ↓ (2.4 SD)	105% ↓	Continuous exponential (M4), constant variance (p = 0.48)	3.81E−01 1.55E−02	No response data near BMR; maximal response at lowest dose, BMD/BMDL ratio »20
Kattainen et al. (2001) (rat)	– 2.23E+00	3rd molar length in pups (n = 4−8)	–	15% ↓ (4.2 SD)	27% ↓	Continuous Hill, modeled variance, restricted (p = 0.02)	3.13E−01 1.68E−01	No response data near BMR; Constrained parameter lower bound hit
						Continuous Hill, modeled variance, unrestricted (p < 0.001)	1.21E−02 –	BMDL could not be calculated
		3rd molar eruption in pups (n = 4−8)	1/16	3/17	13/19	Dichotomous log-logistic, restricted (p = 0.98)	2.40E+00 1.33E+00	Constrained parameter lower bound hit
						Dichotomous log-logistic, unrestricted (p = 0.95)	1.93E+00 1.84E−01	Supralinear fit (slope = 0.91)
Keller et al. (2008a; 2008b; 2007) (mouse)	– 5.37E−01	Missing molars (n = 23−36)	0/29	2/23	30/30	Dichotomous 1° multistage (p = 0.26)	1.09E+00 7.62E−01	Poor fit at first response level; not most sensitive endpoint; other endpoints not amenable to BMD modeling

Table 4-4. TCDD BMDL analysis (NOAEL, LOAEL, BMD, and BMDL values given as animal whole blood concentrations in ng/kg)[a] (continued)

Study[b,c]	NOAEL/ LOAEL	Endpoint	Control response	First response[d]	Max response[e]	Model fit detail	BMD/ BMDL	Comments
Kociba et al. (1978) (rat)	1.55E+00 7.15E+00	Uroporphyrin per creatinine, females (n = 5)	–	15% ↑ (0.48 SD)	89% ↑	Continuous linear, constant variance (p = 0.79)	1.31E+01 9.29E+00	BMDL > LOAEL; otherwise adequate fit
		Urinary coproporphyrins, females (n = 5)	–	67% ↑ (5.1 SD)	78% ↑	Continuous exponential (M4), modeled variance (p = 0.01)	1.57E+00 7.18E−01	Poor fit; no response near BMR
		Liver lesions (n = 50)						No data presented
		Lung lesions (n = 50)						No data presented
Kuchiiwa et al. (2002) (mouse)	1.42E+02 –	Immunoreactive Neurons in Dorsalis, males (n = 6)	–	42% ↓ (3.5 SD)	64% ↓	Continuous linear, constant variance (p = NA, insufficient degrees of freedom)	6.04E-02 4.27E-02	No response near BMR
		Immunoreactive Neurons in Medianus, males (n = 6)	–	63% ↓ (4.8 SD)	75% ↓	Continuous linear, modeled variance (p = NA, insufficient degrees of freedom)	4.93E-02 3.23E-02	No response near BMR
		Immunoreactive Neurons in B9, males (n = 6)	–	69% ↓ (6.6 SD)	87% ↓	Continuous linear, constant variance (p = NA, insufficient degrees of freedom)	4.17E-02 3.01E-02	No response near BMR
		Immunoreactive Neurons in Magnus, males (n = 6)	–	55% ↓ (7.0 SD)	75% ↓	Continuous linear, modeled variance (p = NA, insufficient degrees of freedom)	3.35E-02 2.05E-02	No response near BMR

Table 4-4. TCDD BMDL analysis (NOAEL, LOAEL, BMD, and BMDL values given as animal whole blood concentrations in ng/kg)[a] (continued)

Study[b,c]	NOAEL/ LOAEL	Endpoint	Control response	First response[d]	Max response[e]	Model fit detail	BMD/ BMDL	Comments
Latchoumy-candane and Mathur (2002) (rat)	– 7.85E−01	Daily sperm production (n = 6)	—	29% ↓ (1.0 SD)	41% ↓	Continuous Hill, constant variance, restricted (p = 0.96)	1.17E−01 1.32E−02	Near maximal response at LOAEL; constrained parameter bound hit; standard deviations given in paper interpreted as standard errors
						Continuous Hill, constant variance, unrestricted (p = N/A)	9.96E−02 1.23E−09	Slightly supralinear fit (n = 0.92)
Li et al. (1997) (rat)	2.66E−01 7.99E−01	FSH in female rats (n = 10)	—	3.6-fold ↑ (2.0 SD)	19-fold ↑	Continuous power, modeled variance, restricted (p < 0.01)	2.00E+02 1.36E+02	Power hit lower bound
						Continuous power, modeled variance, unrestricted (p = 0.003)	1.96E−01 2.48E−02	Supralinear fit (power = 0.31)
Li et al. (2006) (mouse)	– 1.59E−01	Serum estradiol (n = 10)	—	2.0-fold ↑ (0.8 SD)	2.4-fold ↑	Continuous linear, constant variance (p = 0.16)	1.61E+01 5.38E+00	BMDL > LOAEL; high control CV (1.25); near maximal response at low dose; nonmonotonic response; other model fits are step-function-like
		Serum progesterone (n = 10)	—	33% ↓ (2.0 SD)	61% ↓	Continuous Hill, modeled variance (p = 0.39)	9.46E−04 8.01E−11	No response data near BMR; large CVs (>1) for treatment groups; poor fit for variance model; Hill coefficient at lower bound (step-function)

Table 4-4. TCDD BMDL analysis (NOAEL, LOAEL, BMD, and BMDL values given as animal whole blood concentrations in ng/kg)[a] (continued)

Study[b,c]	NOAEL/ LOAEL	Endpoint	Control response	First response[d]	Max response[e]	Model fit detail	BMD/ BMDL	Comments
Markowski et al. (2001) (rat)	– 1.56E+00	FR5 run opportunities (n = 4–7)	–	10% ↓ (0.21 SD)	51% ↓	Continuous Hill, constant variance (p = 0.94)	1.72E+00 9.08E−01	Constrained parameter upper bound hit
						Continuous power, constant variance, unrestricted (p = 0.13)	2.67E+00 1.03E−14	Saturated model; supralinear fit (power = 0.39); BMD/BMDL ratio »100
		FR2 revolutions (n = 4–7)	–	9% ↓ (0.15 SD)	43% ↓	Continuous Hill, constant variance (p = 0.65)	1.84E+00 5.99E−01	Constrained parameter bound hit (upper bound)
						Continuous power, constant variance, unrestricted (p = 0.16)	5.74E+00 1.03E−14	Supralinear fit (power = 0.32)
		FR10 run opportunities (n = 4–7)	–	15% ↓ (0.24 SD)	57% ↓	Continuous exponential (M2) , constant variance (p = 0.30)	8.57E+00 2.89E+00	BMDL > LOAEL
Miettinen et al. (2006) (rat)	– 2.22E+00	Cariogenic lesions in pups (n = 4–8)	25/42	23/29	29/32	Dichotomous log-logistic, restricted (p = 0.60)	1.43E+00 5.17E−01	Constrained parameter lower bound hit; near maximal response at LOAEL; high control response
						Dichotomous log-logistic, unrestricted (p = 0.73)	4.94E−02 –	Supralinear fit (slope = 0.47); BMDL could not be calculated
Murray et al. (1979) (rat)	1.12E+00 5.88E+00	Fertility in F2 gen. (no litters) (n = 20)	4/32	0/20	9/20	Dichotomous multistage (p = 0.08)	2.73E+00 1.37E+00	Poor fit; nonmonotonic response; no response data near BMR

Table 4-4. TCDD BMDL analysis (NOAEL, LOAEL, BMD, and BMDL values given as animal whole blood concentrations in ng/kg)[a] (continued)

Study[b,c]	NOAEL/ LOAEL	Endpoint	Control response	First response[d]	Max response[e]	Model fit detail	BMD/ BMDL	Comments
NTP (1982b) (mouse)	– 7.67E−01	Toxic hepatitis; males (n = 50)	1/73	5/49	44/50	Dichotomous multistage (p = 0.04)	2.78E+00 1.34E+00	No acceptable model fits; lowest BMDL shown
NTP (2006a) (rat)	– 2.56E+00	Hepatocyte hypertrophy (n = 53−54)	0/53	19/54	52/53	Dichotomous multistage (p = 0.02)	9.27E−01 7.91E−01	Poor fits for all models
		Alveolar metaplasia (n = 52−54)	2/53	19/54	46/52	Dichotomous log-logistic (p = 0.72)	6.50E−01 3.75E−01	No response near BMR
		Oval cell hyperplasia (n = 53−54)	0/53	4/54	53/53	Dichotomous probit (p = 0.23)	5.67E+00 4.79E+00	Relatively poor fit for control and low-dose groups; negative response intercept (same for logistic); BMDL > LOAEL
						Dichotomous Weibull (p = 0.08)	5.72E+00 4.09E+00	Marginal fit; BMDL > LOAEL
		Gingival hyperplasia (n = 53−54)	1/53	7/54	16/53	Dichotomous log-logistic, restricted (p = 0.06)	5.85E+00 3.73E+00	Poor fit; constrained parameter bound hit; BMDL > LOAEL
						Dichotomous log-logistic, unrestricted (p = 0.66)	7.05E−01 1.26E−05	Supralinear fit (slope = 0.37)
		Eosinophilic focus, multiple (n = 53−54)	3/53	8/54	42/53	Dichotomous probit (p = 0.46)	5.58E+00 4.86E+00	Relatively poor fit to control response; BMDL > LOAEL
		Liver fatty change, diffuse (n = 53−54)	0/53	2/54	48/53	Dichotomous Weibull (p = 0.72)	3.92E+00 2.86E+00	BMDL > LOAEL; otherwise adequate fit

Table 4-4. TCDD BMDL analysis (NOAEL, LOAEL, BMD, and BMDL values given as animal whole blood concentrations in ng/kg)[a] (continued)

Study[b,c]	NOAEL/ LOAEL	Endpoint	Control response	First response[d]	Max response[e]	Model fit detail	BMD/ BMDL	Comments
NTP (2006a) (rat) (continued)	– 2.56E+00 (continued)	Liver necrosis (n = 53−54)	1/53	4/54	17/53	Dichotomous log-probit, unrestricted (p = 0.80)	7.50E+00 3.50E+00	Adequate fit; slightly supralinear; BMDL > LOAEL
		Liver pigmentation (n = 53−54)	4/53	9/54	53/53	Dichotomous log-probit (p = 0.96)	2.46E+00 1.89E+00	Adequate fit
		Toxic hepatopathy (n = 53−54)	0/53	2/54	53/53	Dichotomous multistage (p = 0.69)	3.98E+00 3.06E+00	BMDL > LOAEL; otherwise adequate fit
Ohsako et al. (2001) (rat)	1.04E+00 3.47E+00	Anogenital distance in male pups (n = 5)	–	12% ↓ (1.0 SD)	17% ↓	Continuous Hill, constant variance, restricted (p = 0.15)	2.88E+00 8.03E−01	Constrained parameter lower bound hit; near maximal response at LOAEL
						Continuous Hill, constant variance, unrestricted (p = 0.056)	3.49E+00 3.05E−01	Supralinear fit (n = 0.59)
Schantz et al. (1996)	- 3.38E+00	Facilitory effect on radial arm maze learning (n = 10)	–	22% ↓ (1.2 SD)	34% ↓	Continuous linear, constant variance (p = 0.16)	7.00E+00 4.60E+00	BMDL > LOAEL; otherwise adequate fit
Sewall et al. (1995) (rat)	7.11E+00 1.66E+01	Serum T4 (n = 9)	–	9.1% ↓ (0.6 SD)	40% ↓	Continuous Hill, constant variance, restricted (p = 0.90)	1.03E+01 3.60E+00	Constrained parameter hit lower bound; otherwise acceptable fit; selected
						Continuous Hill, constant variance, unrestricted (p = 0.86)	9.71E+00 1.97E+00	Supralinear fit (power = 0.57)

Table 4-4. TCDD BMDL analysis (NOAEL, LOAEL, BMD, and BMDL values given as animal whole blood concentrations in ng/kg)[a] (continued)

Study[b,c]	NOAEL/ LOAEL	Endpoint	Control response	First response[d]	Max response[e]	Model fit detail	BMD/ BMDL	Comments
Shi et al. (2007) (rat)	3.42E−01 1.07E+00	Serum estradiol in female pups (n = 10)	—	38% ↓ (0.4 SD)	62% ↓	Continuous exponential (M4), modeled variance (p = 0.69)	8.07E−01 3.54E−01	Adequate fit; selected
Smialowicz et al. (2008) (mouse)	− 4.38E−01	PFC per spleen (n = 15)	—	24% ↓ (0.5 SD)	89% ↓	Continuous power, unrestricted, modeled variance (p = 0.27)	1.19E+01 3.76E+00	BMDL > LOAEL; fit at control and low dose inconsistent with data; constrained parameters in other models hit lower bounds
		PFC per 10^6 cells (n = 8−15)	—	24% ↓ (0.5 SD)	9.3-fold ↓	Continuous power unrestricted, constant variance (p = 0.48)	1.90E+00 2.16E−01	Constant variance test failed; observed control variance underestimated by 35%; poor fits for all modeled variance models
Smith et al. (1976) (mouse)	7.11E+00 5.06E+01	Cleft palate in pups (n = 14−41)	0/34	2/41	10/14	Dichotomous log-logistic, restricted (p = 0.42)	3.52E+01 1.06E+01	Adequate fit; selected
Sparschu et al. (2008; 1971) (rats)	5.09E+00 1.63E+01	Male fetus weight (n = 3−117)	—	2.7% ↑ (0.1 SD)	33% ↓	Continuous exponential (M5), modeled variance (p < 0.0001)	5.46E+02 1.30E+02	BMDL > LOAEL; variance not captured by either variance model; poor fit in region surrounding NOAEL and LOAEL
		Female fetus weight (n = 4−129)	—	2.3% ↑ (0.06 SD)	30% ↓	Continuous exponential (M2), modeled variance (p < 0.028)	1.03E+03 6.48E+02	BMDL > LOAEL; variance not captured by either variance model; poor fit in region surrounding NOAEL and LOAEL

Table 4-4. TCDD BMDL analysis (NOAEL, LOAEL, BMD, and BMDL values given as animal whole blood concentrations in ng/kg)[a] (continued)

Study[b,c]	NOAEL/ LOAEL	Endpoint	Control response	First response[d]	Max response[e]	Model fit detail	BMD/ BMDL	Comments
Toth et al. (1979) (mouse)	– 5.73E−01	Skin lesions (n = 38−44)	0/38	5/44	25/43	Dichotomous log-logistic, restricted (p = 0.08)	6.41E+00 4.02E+00	Constrained parameter lower bound hit
						Dichotomous log-logistic, unrestricted (p = 0.74)	5.97E−01 6.77E−02	Supralinear fit (slope = 0.48)
	– 5.73E−01 (cont.)	Dermal amyloidosis (n = 38−44)	0/38	5/44	17/43	Dichotomous log-logistic, restricted (p = 0.05)	1.50E+01 8.75E+00	Poor fit; constrained parameter lower bound hit; BMDL > LOAEL
						Dichotomous log-logistic, unrestricted (p = 0.90)	4.84E−01 5.31E−03	Supralinear fit (slope = 0.33)
Van Birgelen et al. (1995a) (rat)	– 7.20E+00	Hepatitic retinol (n = 8)	–	44% ↓ (0.74 SD)	96% ↓	Continuous exponential (M4), modeled variance (p < 0.01)	2.49E+01 3.36E+00	Poor fit
						Continuous power, modeled variance, unrestricted (p = 0.01)	3.80E−01 1.39E−02	Poor fit; supralinear fit (power = 0.14)
		Hepatitic retinyl palmitate (n = 8)	–	80% ↓ (1.4 SD)	99% ↓	Continuous exponential (M4), modeled variance (p < 0.01)	1.42E+02 3.65E+01	Poor fit; no response near BMR
						Continuous power, modeled variance, unrestricted (p = 0.24)	5.26E−02 5.89E−05	Supralinear fit (power = 0.06)

Table 4-4. TCDD BMDL analysis (NOAEL, LOAEL, BMD, and BMDL values given as animal whole blood concentrations in ng/kg)[a] (continued)

Study[b,c]	NOAEL/ LOAEL	Endpoint	Control response	First response[d]	Max response[e]	Model fit detail	BMD/ BMDL	Comments
White et al. (1986) (mouse)	– 1.09E+00	Total hemolytic complement activity (CH50) (n = 8)	—	41% ↓ (2.6 SD)	81% ↓	Continuous Hill, modeled variance, restricted (p = 0.002)	8.63E+00 1.50E+00	Poor fit; no response near BMR; constrained parameter bound hit; BMDL > LOAEL
						Continuous Hill, modeled variance, unrestricted (p = 0.07)	1.48E−01 4.35E−03	Supralinear fit (n = 0.25)

[a]Animal whole blood concentrations were used to determine the HEDs in Table 4-3 and Table 4-5.
[b]The following studies previously presented in Table 4-3 are not presented in Table 4-4 because toxicokinetic models for guinea pigs, minks, or monkeys, and were not found: DeCaprio et al. (1986); Hochstein et al (2001); Rier et al. (1995; 1993); Vos et al. (1973); Yang et al. (2000).
[c]The following studies previously presented in Table 4-3 are not presented in Table 4-4 because the data were not amenable to BMD modeling: Chu et al. (2001); Chu et al. (2007); Croutch et al. (2005); Fattore et al. (2000); Fox et al. (1993); Franczak et al. (2006); Hutt et al. (2008); Ikeda et al. (2005); Ishihara et al. (2007); Kociba et al. (1976); Maronpot et al. (1993); Nohara et al. (2000); Nohara et al. (2002) Schantz et al. (1996)Seo et al. (1995); Simanainen et al. (2002); Simanainen et al. (2003); Simanainen et al. (2004a); Smialowicz et al. (2004); Weber et al. (1995).
[d]Magnitude of response at first dose where response differs from control value (in the adverse direction); continuous response magnitudes given as relative to control plus change relative to control standard deviation; quantal response given as number affected/total number.
[e]Magnitude of response maximally differing from control value (in the adverse direction).
SD = standard deviation; S-D = Sprague-Dawley; L-E = Long-Evans; H-W = Han-Wistar.

Table 4-5. Candidate PODs for the TCDD RfD using blood-concentration-based human equivalent doses

Study	Species, strain (sex, if not both)	Protocol	Endpoint	$NOAEL_{HED}$ (N) or $BMDL_{HED}$ (B) (ng/kg-day)	$LOAEL_{HED}$ (ng/kg-day)	UF[a]	RfD (mg/kg-day)
Li et al. (2006)	Mouse, NIH (F)	Gavage GDs 1−3; *n* = 10	Hormone levels in pregnant dams (decreased progesterone, increased estradiol)	–	1.6E−03	300	5.3E−12
Kuchiiwa et al. (2002)	Mouse, ddY	Maternal 8 week-gavage prior to mating; *n* = 3	Decreased serotonin-immunoreactive neurons in raphe nuclei of male offspring (F1)	–	2.7E-03	300	9.2E-12
Smialowicz et al. (2008)	Mouse, $B6C3F_1$ (F)	90-day gavage; *n* = 8−15	Decreased SRBC response	–	6.3E−03	300	2.1E−11
Bowman et al.(1989a; 1989b); others[b]	Rhesus Monkey (F)	Daily dietary exposure, 3.5−4 years *n* = 3−7	Neurobehavioral effects	–	8.2E−03[c]	300	2.7E−11
Keller et al. (2008a; 2008b; 2007)[d]	Mouse, CBA/J and C3H/HeJ	Gavage GD 13; *n* = 23−36 (pups)	Missing molars, mandibular shape changes in pups	–	9.9E−03	300	3.3E−11
Toth et al. (1979)	Mouse, Swiss/ H/Riop (M)	1-year gavage; *n* = 38−44	Dermal amyloidosis, skin lesions	–	9.9E−03	300	3.3E−11
Latchoumy-candane and Mathur (2002); others[e]	Rat, Wistar (M)	45-day oral pipetting; *n* = 6	Decreased sperm production	–	1.6E−02	300	5.4E−11
NTP (1982b)	Mouse, $B6C3F_1$ (M)	2-year gavage; *n* = 50	Liver lesions	–	2.2E−02	300	7.2E−11

Table 4-5. Candidate points of departure for the TCDD RfD using human equivalent doses (continued)

Study	Species, strain (sex, if not both)	Protocol	Endpoint	$NOAEL_{HED}$ (N) or $BMDL_{HED}$ (B) (ng/kg-day)	$LOAEL_{HED}$ (ng/kg-day)	UF[a]	RfD (mg/kg-day)
White et al. (1986)	Mouse, B6C3F$_1$ (F)	14-day gavage; *n* = 6−8	Decreased serum complement	–	2.8E−02	300	9.2E−11
Li et al. (1997)	Rat, S-D (F, 22 day-old)	Single gavage; *n* = 10	Increased serum FSH	2.9E−03 (N)	1.7E−02	30[f]	9.7E−11
DeCaprio et al. (1986)	Guinea pig, Hartley	90-day dietary; *n* = 10	Decreased body weight, organ weight changes (liver, kidney, thymus, brain)	4.1E−03[c] (N)	3.3E−02[c]	30[f]	1.4E−10
Shi et al. (2007)	Rat, S-D (F)	11-month gavage; *n* = 10	Decreased serum estradiol	4.5E−03 (N) 4.7E−03 (B)	2.7E−02	30[f]	1.6E−10
Markowski et al. (2001)	Rat, Holtzman	Gavage GD 18; *n* = 4−7	Neurobehavioral effects in pups (running, lever press, wheel spinning)	–	5.2E−02	300	1.7E−10
Hojo et al. (2002); Zareba et al. (2002)	Rat, S-D	Gavage GD 8; *n* = 12	Food-reinforced operant behavior in pups	–	5.5E−02	300	1.8E−10
Cantoni et al. (1981)	Rat, CD-COBS (F)	45-week gavage; *n* = 4	Increased urinary porhyrins	–	6.4E−02	300	2.1E−10
Vos et al. (1973)	Guinea pig, Hartley (F)	8-week gavage; *n* = 10	Decreased delayed-type hypersensitivity response to tuberculin	6.4E−03[c] (N)	3.2E−02[c]	30[f]	2.1E−10
Miettinen et al. (2006)	Rat, Line C	Gavage GD 15; *n* = 3−10	Cariogenic lesions in pups	–	8.9E−02	300	3.0E−10
Kattainen et al. (2001)	Rat, Line C	Gavage GD 15; *n* = 4−8	Inhibited molar development in pups	–	9.0E−02	300	3.0E−10
NTP (2006a)	Rat, S-D (F)	2-year gavage; *n* = 53	Liver and lung lesions	–	1.4E−01	300	4.5E−10
Amin et al. (2000)	Rat, S-D	Gavage GDs 10−16; *n* = 10	Reduced saccharin consumption and preference	–	1.7E−01	300	5.7E−10
Schantz et al. (1996)	Rat, S-D (F)	Gavage GDs 10-16; *n* = 80-88	Maze errors (facilitatory effect)	–	1.7E−01	300	5.7E−10

Table 4-5. Candidate points of departure for the TCDD RfD using human equivalent doses (continued)

Study	Species, strain (sex, if not both)	Protocol	Endpoint	$NOAEL_{HED}$ (N) or $BMDL_{HED}$ (B) (ng/kg-day)	$LOAEL_{HED}$ (ng/kg-day)	UF[a]	RfD (mg/kg-day)
Mocarelli et al. (2008)	**Human (M)**	**Childhood exposure; *n* = 157**	**Decreased sperm concentration and sperm motility, as adults**	–	**2.0E−02[g]**	**30[h]**	**6.7E−10**
Baccarelli et al. (2008)	**Human infants**	**Gestational exposure; *n* = 51**	**Increased TSH in newborn infants**	–	**2.0E−02[i]**	**30[h]**	**6.7E−10**
Hutt et al. (2008)	Rat, S-D (F)	13-week dietary; *n* = 3	Embryotoxicity	–	2.5E−01	300	8.4E−10
Ohsako et al. (2001)	Rat, Holtzman	Gavage GD 15; *n* = 5	Decreased anogenital distance in male pups	2.7E−02 (N)	1.8E−01	30[f]	9.1E−10
Murray et al. (1979)	Rat, S-D	3-generation dietary	Reduced fertility and neonatal survival (F0 and F1)	2.9E−02 (N)	3.8E−01	30[f]	9.6E−10
Franczak et al. (2006)	Rat, S-D (F)	Gavage GD 14, 21, PND 7, 14; *n* = 7	Abnormal estrous cycle	–	3.2E-01	300	1.1E-09
Chu et al. (2007)	Rat, S-D (F)	28-day gavage, *n* = 5	Liver lesions	3.5E−02 (N)	5.6E−01	30[f]	1.2E−09
Bell et al. (2007b)	Rat, CRL:WI (Han) (M)	17-week dietary; *n* = 30	Delay in onset of puberty	4.3E−02 (B)	8.9E−02	30[f]	1.4E−09
Ishihara et al., (2007)	Mouse, ICR (M)	Weekly gavage for 5 weeks; *n* = 42−43	Decreased male/female sex ratio	–[j]	5.0E-01	300	1.7E-09
VanBirgelen et al. (1995a)[k]	Rat, S-D (F)	13-week dietary; *n* = 8	Decreased liver retinyl palmitate	–	5.1E−01	300	1.7E−09
Kociba et al. (1978)	Rat, S-D (F)	2-year dietary; *n* = 50	Liver and lung lesions, increased urinary porhyrins	6.3E−02 (N)	6.3E−01	30[f]	2.1E−09
Fattore et al. (2000)	Rat, S-D	13-week dietary; *n* = 6	Decreased hepatic retinol	–	7.8E−01	300	2.6E−09
Seo et al. (1995)	Rat, S-D	Gavage GDs 10−16; *n* = 10	Decreased serum T4 and thymus weight	1.7E−01 (N)	9.1E−01	30[f]	5.6E−09
Crofton et al. (2005)	Rat, Long-Evans (F)	4-day gavage; *n* = 4−14	Decreased serum T4	1.7E−01 (N)	7.4E−01	30[f]	5.6E−09

Table 4-5. Candidate points of departure for the TCDD RfD using human equivalent doses (continued)

Study	Species, strain (sex, if not both)	Protocol	Endpoint	$NOAEL_{HED}$ (N) or $BMDL_{HED}$ (B) (ng/kg-day)	$LOAEL_{HED}$ (ng/kg-day)	UF[a]	RfD (mg/kg-day)
Sewall et al. (1995)	Rat, S-D (F)	30-week gavage; *n* = 9	Decreased serum T4	5.0E−01 (N) 1.8E−01 (B)	1.7E+00	30[f]	6.0E−09
Franc et al. (2001)	Rat, Long-Evans (F)	22-week gavage; *n* = 8	Increased relative liver weight; decreased relative thymus weight	4.5E−01 (N) 2.6E−01 (B)	1.4E+00	30[f]	8.7E−09
Kociba et al. (1976)	Rat, S-D	5-days/week gavage for 13 weeks; *n* = 12	Liver and lung lesions, increased urinary porphyrins	2.6E-01 (N)	3.0E+00	30[f]	8.7E-09
Sparschu et al. (1971)	Rat, S-D (F)	Gavage GD 6-15; *n* = 4−129	Decreased fetal body weight	3.2E-01 (N)	1.7E+00	30[f]	1.1E-08
Alaluusua et al. (2004)	Human	Childhood exposure; *n* = 48	Dental defects	4.1E−02[l] (N)	9.0E−01[m]	3[n]	1.4E−08

[a]Except where indicated, $UF_A = 3$ (for dynamics), $UF_H = 10$, $UF_L = 10$.
[b]Schantz and Bowman (1989); Schantz et al. (1986); Schantz et al. (1986).
[c]HED determined from 1st-order body burden model; no PBPK model available for guinea pigs or monkeys; Hochstein et al. (2001) was not presented in the table because no PBPK model exists for minks and 1st-order body burden could not be calculated because a TCDD half-life could not be determined.
[d]Results from three separate studies with identical designs combined.
[e]Latchoumycandane et al. (2002a; 2002b).
[f]$UF_L = 1$ (NOAEL or BMDL).
[g]Mean of peak exposure (0.0321 ng/kg-day) and average exposure over 10-year critical window (0.0080 ng/kg-day).
[h]$UF_H = 3$, $UF_L = 10$.
[i]Maternal exposure corresponding to neonatal TSH concentration exceeding 5 µU/mL.
[j]The NOAEL of 4.9E-5 was excluded from consideration because of the large dose spacing in the study.
[k]Van Birgelen et al. (1995b) is considered to be the same study but reports effects at doses above the LOAEL that are not considered further; this study in not carried forward for determination of an RfD POD but is included in the RfD uncertainty analysis presented in Section 4.4.
[l]Mean of peak exposure (0.0655 ng/kg-day) and average exposure over 10-year critical window (0.0156 ng/kg-day).
[m]Mean of peak exposure (1.65 ng/kg-day) and average exposure over 10-year critical window (0.149 ng/kg-day).
[n]$UF_H = 3$.

S-D = Sprague-Dawley.

Table 4-6. Qualitative analysis of the strengths and limitations/uncertainties associated with animal bioassays providing PODs for the TCDD RfD

Study	Strengths	Limitations	Remarks
Bell et al. (2007b)	• Large sample size of both rat dams and offspring/dose employed • Several developmental effects tested	• Batch-to-batch variation of up to 30% in TCDD concentration in the diet • Longer-term dosing of dams does not accurately define gestational period when fetus is especially sensitive to TCDD-induced toxicity	Study is a significant addition to a substantial database on the developmental toxicity of TCDD in laboratory animals
Cantoni et al. (1981)	• Experiments were designed to test qualitative and quantitative composition and the course of urinary excretion in TCDD-induced porphyria	• Small sample size of rats/dose employed (n = 4) • Concurrent histological changes with tissue porphyrin levels were not examined • TCDD used for dosing was of unknown purity	Early study on porphyrogenic effects of TCDD
DeCaprio et al. (1986)	• Subchronic oral dosing duration up to 90 days • Male and female guinea pigs tested	• Relatively small sample size of guinea pigs/dose employed (n = 10) • No histopathological analyses performed • TCDD used for dosing was of unknown purity	Limited subchronic study; PBPK model not available for estimation of HED
Franc et al. (2001)	• Three different rat strains with varying sensitivities to TCDD were utilized (Sprague-Dawley, Long Evans, Han/Wistar) • Longer-term oral dosing up to 22 weeks	• Relatively small sample size of rats/dose employed (n = 8) • Only female rats were tested • Concurrent liver histopathological changes with liver-weight changes were not examined • Gavage exposure was only biweekly	Limited subchronic study
Hojo et al. (2002)	• Low TCDD dose levels used allowed for subtle behavioral deficits to be identified in rat offspring • Preliminary training sessions in operant chamber apparatuses were extensive • Neurobehavioral effects are exposure-related and cannot be attributed to presence of learning or discrimination deficits	• Relatively small sample size of rat dams/dose employed (n = 12) • Small sample size of rat offspring/dose evaluated (n = 5−6) • Neurobehavioral effects induced by TCDD at earlier or later gestational dosing dates are unknown because of single gavage administration on GD 8 • Although BMD analysis was conducted, the model parameters were not constrained according to EPA guidance, so the results cannot be used	One of a few neurobehavioral toxicity studies; somewhat limited study size

Table 4-6. Qualitative analysis of the strengths and limitations/uncertainties associated with animal bioassays possessing candidate points-of-departure for the TCDD RfD (continued)

Study	Strengths	Limitations	Remarks
Keller et al. (2008a; 2008b; 2007)	• Six different inbred mouse strains were utilized • Large sample size of mouse offspring/dose/strain evaluated • Low TCDD dose levels used compared to typical mouse studies allowed for identification of subtle sensitivity differences in presence of absence of third molars, variant molar morphology, and mandible structure in offspring	• Unknown sample size of mouse dams/dose/strain employed • All inbred strains possessed sensitive *b* allele at the *Ahr* locus (i.e., a potentially resistant subpopulation was not evaluated for comparison purposes) • Morphological dental and mandibular changes induced by TCDD at earlier or later gestational dosing dates are unknown because of single gavage administration on GD 13 • Difficulties breeding A/J mice led to abandonment of that strain in the analysis (Keller et al., 2008a; Keller et al., 2008b)	Endpoint similar to effects observed at higher exposure levels in humans; HED highly uncertain using mouse PBPK model
Latchoumy-candane and Mathur (2002)	• Compared to epididymal sperm counts, the testicular spermatid head count provides better quantitation of acute changes in sperm production and can indicate pathology	• Small sample size of rats/dose employed (n = 6) • Oral pipette administration of TCDD may be a less efficient dosing method than gavage	Endpoint has human relevance, similar to critical effects in principal human study for RfD
Li et al. (2006)	• Female reproductive effects (i.e., early embryo loss and changes in serum progesterone and estradiol) were tested at multiple exposure times—early gestation, preimplantation, and peri- to postimplantation	• Small sample size of dams/dose (n = 10) • Large dose-spacing interval (25-fold at lowest 2 doses)	Endpoint has human relevance but HED highly uncertain using mouse PBPK model
Markowski et al. (2001)	• Low TCDD dose levels used allowed for subtle behavioral deficits to be identified in rat offspring • Several training sessions on wheel apparatuses were extensive • Neurobehavioral effects are exposure-related and cannot be attributed to motor or sensory deficits	• Unknown sample size of rat dams/dose employed • Small sample size of rat offspring/dose evaluated (n = 4–7) • TCDD used for dosing was of unknown purity and origin • Only two treatment levels • Neurobehavioral effects induced by TCDD at earlier or later gestational dosing dates are unknown because of single gavage administration on GD 18	One of a few neurobehavioral toxicity studies; somewhat limited study size

Table 4-6. Qualitative analysis of the strengths and limitations/uncertainties associated with animal bioassays possessing candidate points-of-departure for the TCDD RfD (continued)

Study	Strengths	Limitations	Remarks
NTP (1982b)	• Large sample size of mice and rats/dose employed • Comprehensive 2-year bioassay that assessed body weights, clinical signs, and pathological changes in multiple tissues and organs	• Elevated background levels of hepatocellular tumors in untreated male mice • Gavage exposure was only 2 days/week • Only two treatment levels	Comprehensive chronic toxicity evaluations of TCDD in rodents; HED highly uncertain using mouse PBPK model
NTP (2006a)	• Chronic exposure duration with several interim sacrifices • Large number of dose groups with close spacing • Large number of animals per dose group • Comprehensive suite of endpoints evaluated • Comprehensive biochemical, clinical, and histopathological tests and measures • Detailed reporting of results, with individual animal data presented as well as group summaries	• Single species, strain, and sex • Lowest dose tested too high for establishing NOAEL	Study is the most comprehensive chronic TCDD toxicity evaluation in rats to date
Shi et al. (2007)	• Study design evaluated TCDD effects on aging female reproductive system (i.e., exposure began in utero and spanned across reproductive lifespan) • Several female reproductive endpoints were evaluated, including cyclicity, endocrinology, serum hormone levels, and follicular reserves	• Relatively small sample size of rats/dose employed (n = 10)	Endpoint similar to effects observed at higher exposure levels in humans
Smialowicz et al. (2008)	• Sheep red blood cell (SRBC) plaque forming cell assay is highly sensitive and reproducible across laboratories when examining TCDD	• Small sample size of animals/dose (n = 8) • Only female mice were tested • Thymus and spleen weights were only other immune response-related endpoints tested	Limited immunotoxicity study
Toth et al. (1979)	• Large sample size of mice/dose employed • Chronic exposure duration	• Reporting of findings is terse and lacks sufficient detail (e.g., materials and methods, thorough description of pathological findings, etc.) • Limited number of endpoints examined • Only male mice were tested	Limited chronic study; HED highly uncertain using mouse PBPK model

Table 4-6. Qualitative analysis of the strengths and limitations/uncertainties associated with animal bioassays possessing candidate points-of-departure for the TCDD RfD (continued)

Study	Strengths	Limitations	Remarks
Vos et al. (1973)	• Three different animal species tested (guinea pigs, mice, and rats) • Effects of TCDD tested on both cell-mediated and humoral immunity	• Small sample size of animals/dose employed in each experiment (n = 5−10) • Only female guinea pigs and rats were tested, and only male mice were tested • Only one experimental assay was utilized to assess cell-mediated or humoral immunity; humoral immunity was only investigated in guinea pigs • TCDD used for dosing was of unknown purity	Endpoints relevant to humans but study size limited; PBPK model not available for estimation of HED
White et al. (1986)	• Total hemolytic complement (CH50) is representative functional assay of the complete complement sequence	• Small sample size of rats/dose employed (n = 6−8) • Individual complement factors may be significantly depleted without affecting CH50 activity (only C3 is measured) • TCDD used for dosing was of unknown purity	Endpoint similar to effects observed at higher exposure levels in humans; HED highly uncertain using mouse PBPK model

Table 4-7. Basis and derivation of the TCDD reference dose

Principal study detail		
Study	**POD (ng/kg-day)**	**Critical effects**
Mocarelli et al. (2008)	0.020 (LOAEL)	Decreased sperm count (20%) and motility (11%) in men exposed to TCDD during childhood
Baccarelli et al. (2008)	0.020 (LOAEL)	Elevated TSH (>5 µU/mL) in neonates

RfD derivation	
POD	0.020 ng/kg-day (2.0E−8 mg/kg-day)
UF	30 (UF_L = 10, UF_H = 3)
RfD	7×10^{-10} (7E-10) mg/kg-day (2.0E−8 ÷ 30)

Uncertainty factors		
LOAEL-to-NOAEL (UF_L)	10	No NOAEL established; cannot quantify lower exposure group in Baccarelli et al. (2008); magnitude of effects at LOAEL sufficient to require a 10-fold factor.
Human interindividual variability (UF_H)	3	A factor of 3 ($10^{0.5}$) is used because the effects were elicited in sensitive lifestages. A further reduction to 1 was not made because the sample sizes were relatively small, which, combined with uncertainty in exposure estimation, may not fully capture the range of interindividual variability. In addition, chronic effects are levels are not fully elucidated for humans and could possibly be more sensitive.
Interspecies extrapolation (UF_A)	1	Human study.
Subchronic-to-chronic (UF_S)	1	Chronic effect levels are not well defined for humans; however, animal bioassays indicate that duration of exposure does not seem to be a determining factor in toxicological outcomes. Developmental effects and other short-term effects occur at doses similar to effects noted in chronic studies. Considering that exposure in the principal studies encompasses the critical window of susceptibility associated with development, a UF to account for exposure duration is not used.
Database sufficiency (UF_D)	1	The database for TCDD contains an extensive range of human and animal studies that examine a comprehensive set of endpoints. There is no evidence to suggest that additional data would result in a lower RfD.

Table 4-8. Alternative PODs for the impact of TCDD exposure during gestation and nursing on semen quality of male offspring (Mocarelli et al., 2011)

POD type	Age-at-conception scenario	Averaging protocol[a]	Maternal intake (ng/kg-day)	
			TCDD only	TCDD + DLC[b]
NOAEL	30 years	Cont. avg.	2.9×10^{-4}	2.90×10^{-3}
LOAEL			1.64×10^{-3}	5.36×10^{-3}
NOAEL	45 years	Cont. avg.	1.9×10^{-4}	1.90×10^{-3}
LOAEL			1.10×10^{-3}	4.80×10^{-3}

[a]Cont. avg. = average continuous exposure over the specified duration.
[b]Added DLC = 9 × TCDD intake for NOAEL (in background range), 3.51×10^{-3} ng/kg-day for LOAEL (above background).

Table 4-9. Alternative PODs for developmental endpoints other than increased neonatal TSH and semen quality

Population, endpoint (cite)	POD type	Averaging protocol[a]	TCDD only (ng/kg-day)		TCDD + DLC (ng/kg-day)	
			Needham	Eskenazi	Needham[b]	Eskenazi[c]
Girls, duration of menstrual cycle as women (Eskenazi et al., 2002b)	NOAEL	Cont. avg.	9.52×10^{-3}	2.90×10^{-3}	0.0130	0.0120
	LOAEL	Peak	3.13	2.94	3.13	2.95
		Window	0.122	0.126	0.126	0.135
		P/W avg.	1.64	1.53	1.64	1.54
Girls and boys, developmental dental effects (Alaluusua et al., 2004)	NOAEL	Peak	0.0655	0.0437	0.0688	0.0528
		Window	0.0157	0.0175	0.0190	0.0266
		P/W avg.	0.0406	0.0306	0.0439	0.0397
	LOAEL	Peak	1.65	1.51	1.65	1.52
		Window	0.149	0.151	0.152	0.160
		P/W avg.	0.897	0.841	0.900	0.841
Girls, age at menarche (Warner et al., 2004)	NOAEL	Peak	0.604	0.517	0.607	0.526
		Window	0.0394	0.0424	0.0427	0.0515
		P/W avg.	0.322	0.280	0.325	0.289

[a]Cont. avg. = average continuous daily intake over the specified duration; Peak = average intake for peak exposure; Window = average intake for critical-window exposure; P/W avg. = average of "Peak" and "Window" intakes.
[b]Added DLC = 3.51×10^{-3} ng/kg-day for girls, 3.33×10^{-3} ng/kg-day for boy/girl average.
[c]Added DLC = 9.10×10^{-3} ng/kg-day for all.

Table 4-10. Alternative PODs for adult endpoints for which critical exposure windows are undefined

Population, endpoint (cite)	POD type	Averaging protocol[a]	TCDD only (ng/kg-day)	TCDD + DLC[b] (ng/kg-day)
Men, sex ratio of offspring (Mocarelli et al., 2000)	NOAEL	Peak	0.0341	0.0373
		Window	1.58×10^{-3}	4.73×10^{-3}
		P/W avg.	0.0178	0.0210
	LOAEL	Peak	0.162	0.165
		Window	4.69×10^{-3}	7.84×10^{-3}
		P/W avg.	0.0831	0.0863
Women, age at menopause (Eskenazi et al., 2005)	NOAEL	Peak	1.6×10^{-4}–3.4×10^{-3}	1.6×10^{-3}–6.9×10^{-3}
		Window	1.6×10^{-4}–1.0×10^{-3}	1.6×10^{-3}–4.5×10^{-3}
		P/W avg.	1.6×10^{-4}–2.2×10^{-3}	1.6×10^{-3}–5.7×10^{-3}
	LOAEL	Peak	0.013–0.052	0.016–0.055
		Window	1.7×10^{-3}–3.4×10^{-3}	5.2×10^{-3}–7.0×10^{-3}
		P/W avg.	7.3×10^{-3}–0.028	0.011–0.031
Women, ovarian function, progesterone (Warner et al., 2007)	NOAEL	Peak	0.204	0.208
		Window	3.00×10^{-3}	6.51×10^{-3}
		P/W avg.	0.104	0.108

[a]Cont. avg. = average continuous daily intake over the specified duration; Peak = average intake for peak exposure; Window = average intake for critical-window exposure; P/W avg. = average of "Peak" and "Window" intakes.
[b]Added DLC = 3.15×10^{-3} ng/kg-day for males, 3.51×10^{-3} ng/kg-day for females, 3.33×10^{-3} ng/kg-day for male/female average.

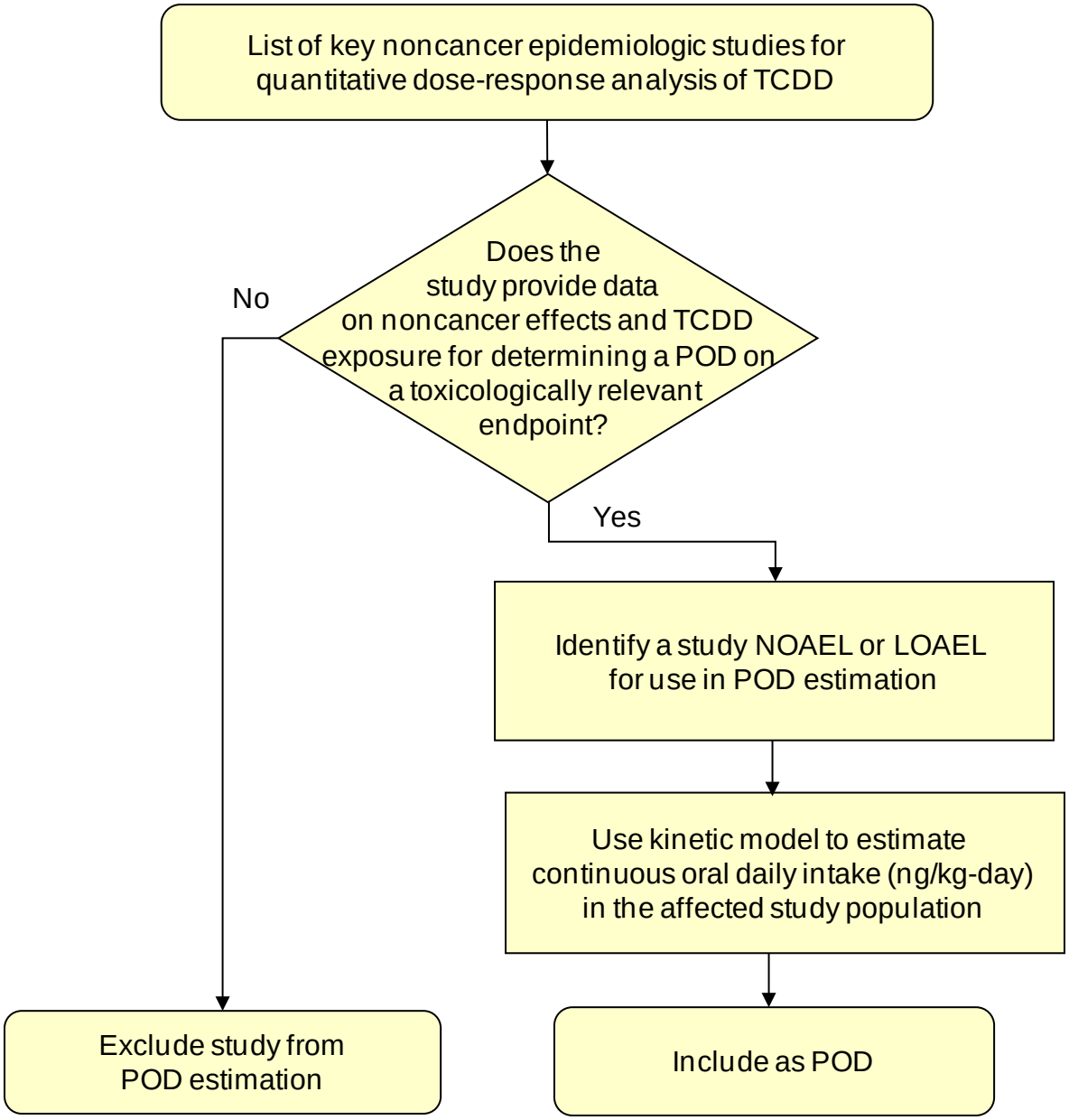

Figure 4-1. EPA's process to identify and estimate PODs from key epidemiologic studies for use in noncancer dose-response analysis of TCDD.
For each noncancer study that qualified using the study inclusion criteria, EPA evaluated the dose-response information developed by the study authors for whether the study provided noncancer effects and TCDD dose data for a toxicologically relevant endpoint. If such data were available, EPA identified a NOAEL or LOAEL as a POD. Then, EPA used a human kinetic model to estimate the continuous oral daily intake (ng/kg-day) for the POD that could be used in the derivation of a candidate RfD based on the study data. If all of this information was available, then the result was included as a POD.

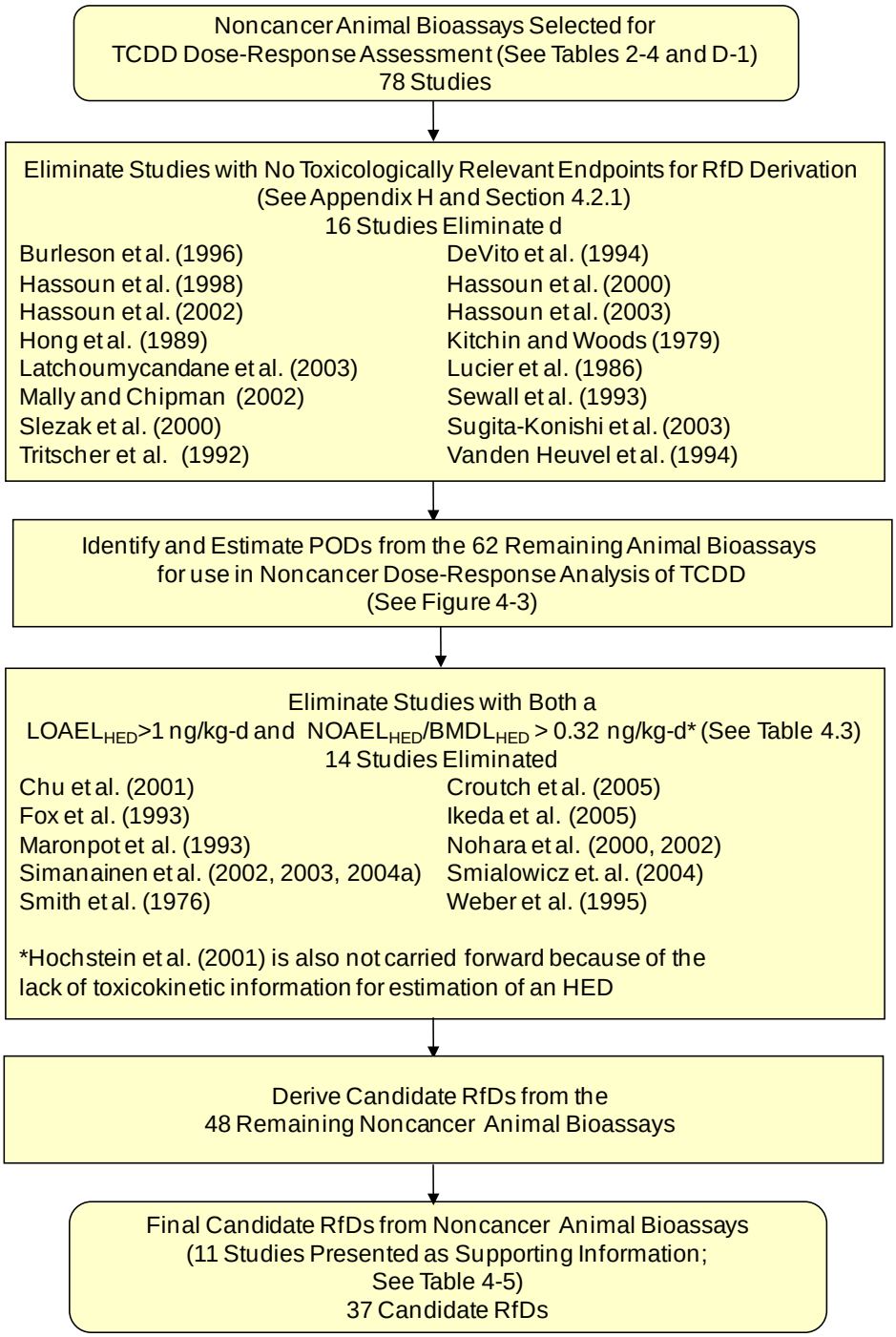

Figure 4-2. Disposition of noncancer animal bioassays selected for TCDD dose-response analysis.

EPA evaluated each noncancer endpoint found in the 78 studies that passed the study inclusion criteria. From this evaluation, EPA eliminated 16 studies that contained no toxicologically relevant endpoints for RfD derivation. Then, as detailed in Figure 4-3, EPA selected and identified PODs for use in deriving candidate RfDs. EPA then eliminated 13 studies based on dose limits for the PODs' HEDs; one study was also not carried forward because of the lack of toxicokinetic information for estimation of an HED. Of the remaining 48 studies, EPA derived 37 RfD candidates, with 11 studies presented as supporting information.

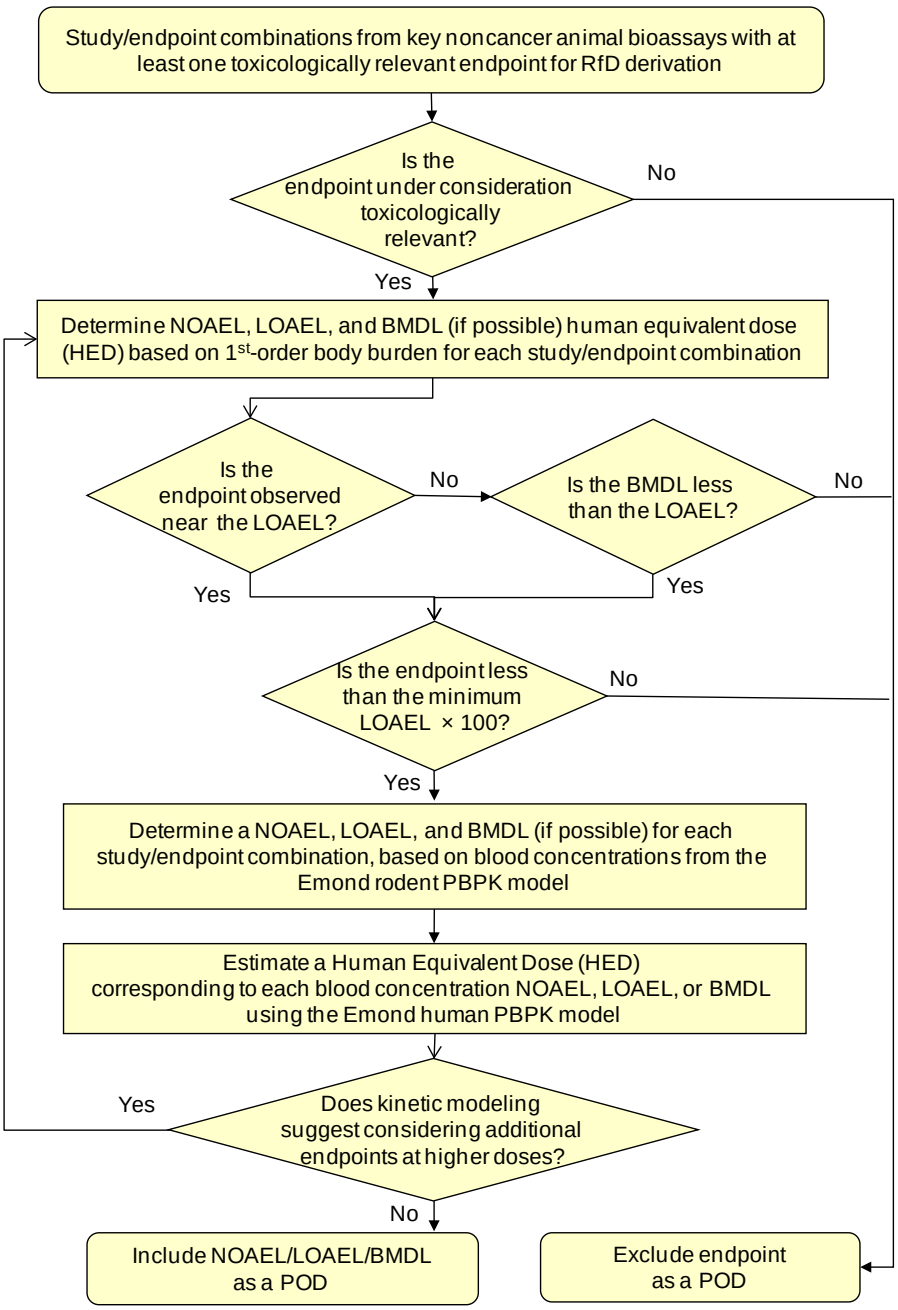

Figure 4-3. EPA's process to identify and estimate PODs from key animal bioassays for use in noncancer dose-response analysis of TCDD.

For the studies with at least one toxicologically relevant endpoint, EPA first determined if each endpoint was toxicologically relevant. If so, EPA determined the NOAEL, LOAEL, and BMDL Human Equivalent Dose (HED) based on 1st-order body burdens for each endpoint. Within each study, these potential PODs were included when the endpoint was observed near the LOAEL and if the BMDL was less than the LOAEL. Then, if the endpoint was less than the minimum LOAEL ×100 across all studies, EPA calculated PODs based on blood concentrations from the Emond rodent PBPK model and, for all of the PODs, HEDs were estimated using the Emond human PBPK model. If the kinetic modeling results suggested considering additional endpoints at higher doses, the process was repeated. Finally, the lowest group of the toxicologically relevant PODs was selected for final use in derivation of candidate RfDs.

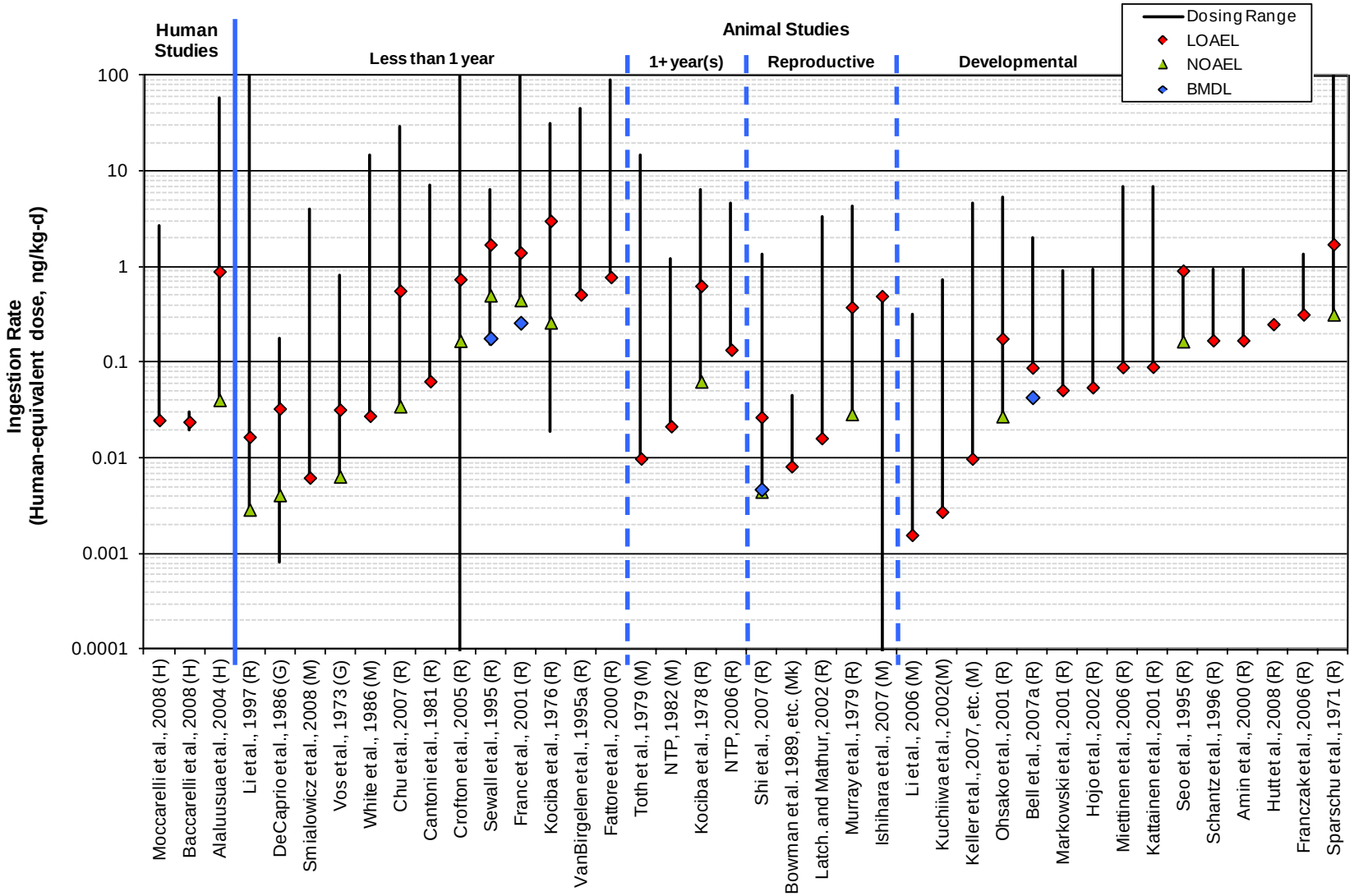

Figure 4-4. Exposure-response array for ingestion exposures to TCDD.

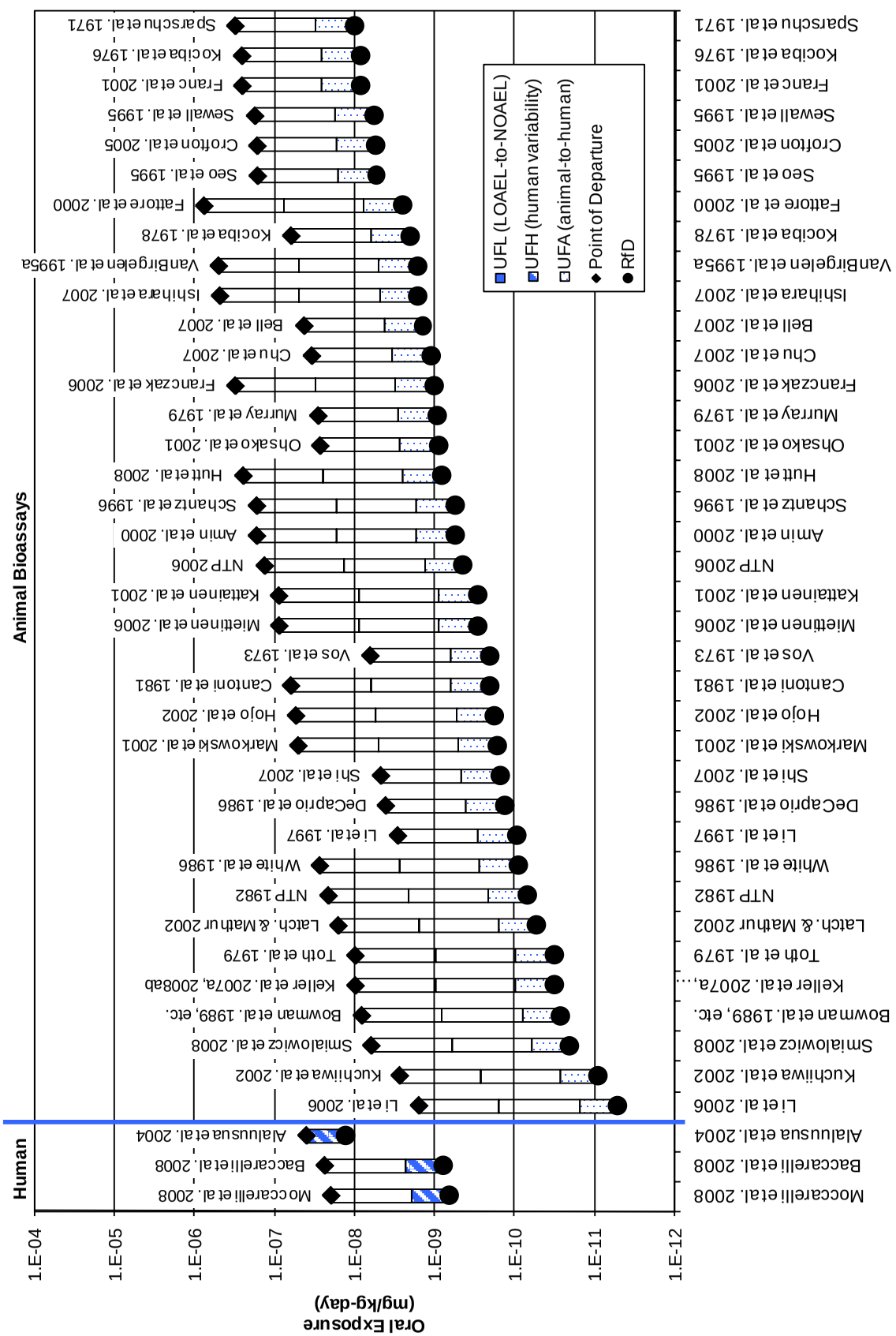

Figure 4-5. Candidate RfD array.

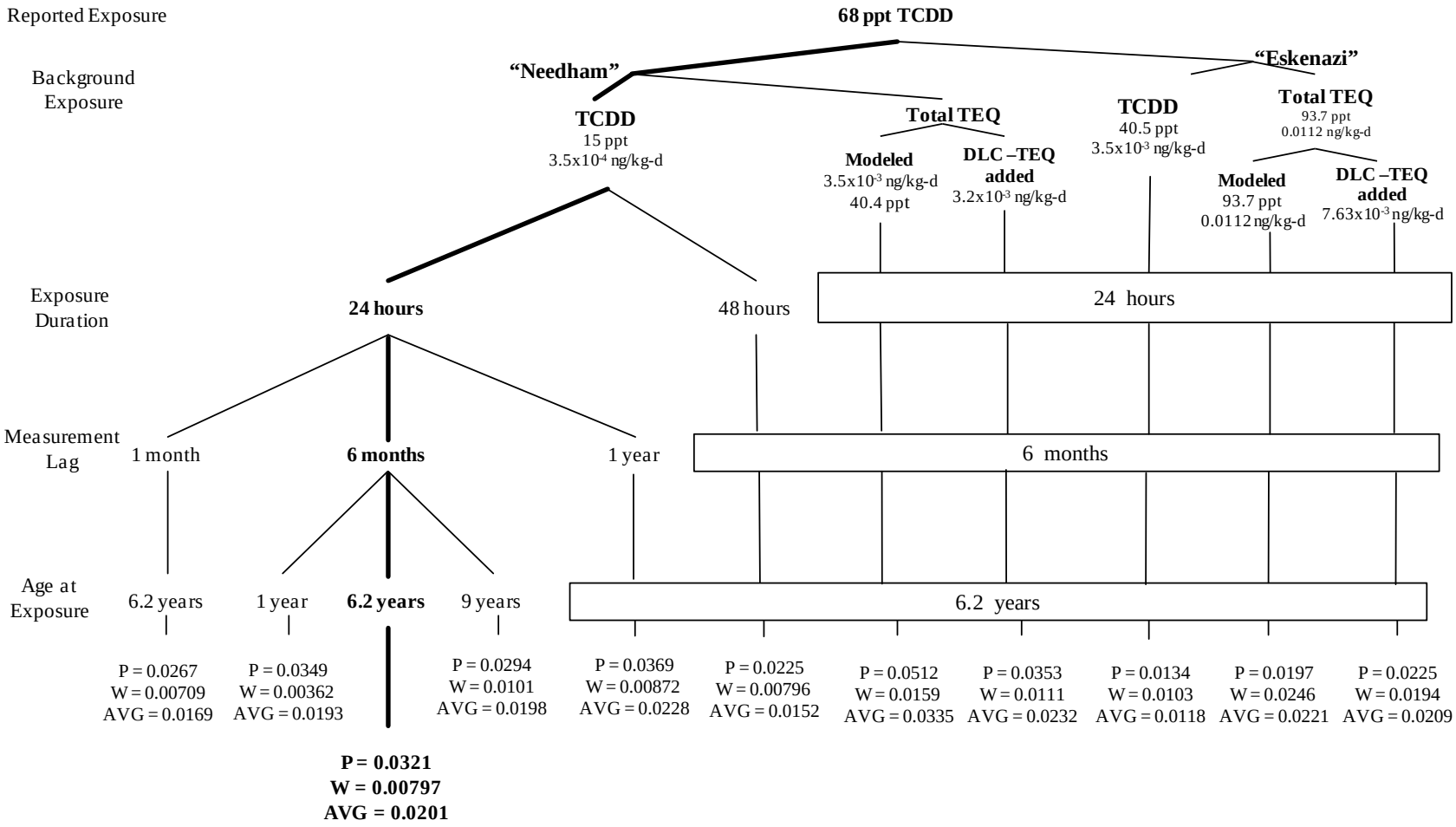

Figure 4-6. Sensitivity tree showing TCDD exposure-variable uncertainty for Mocarelli et al. (2008).

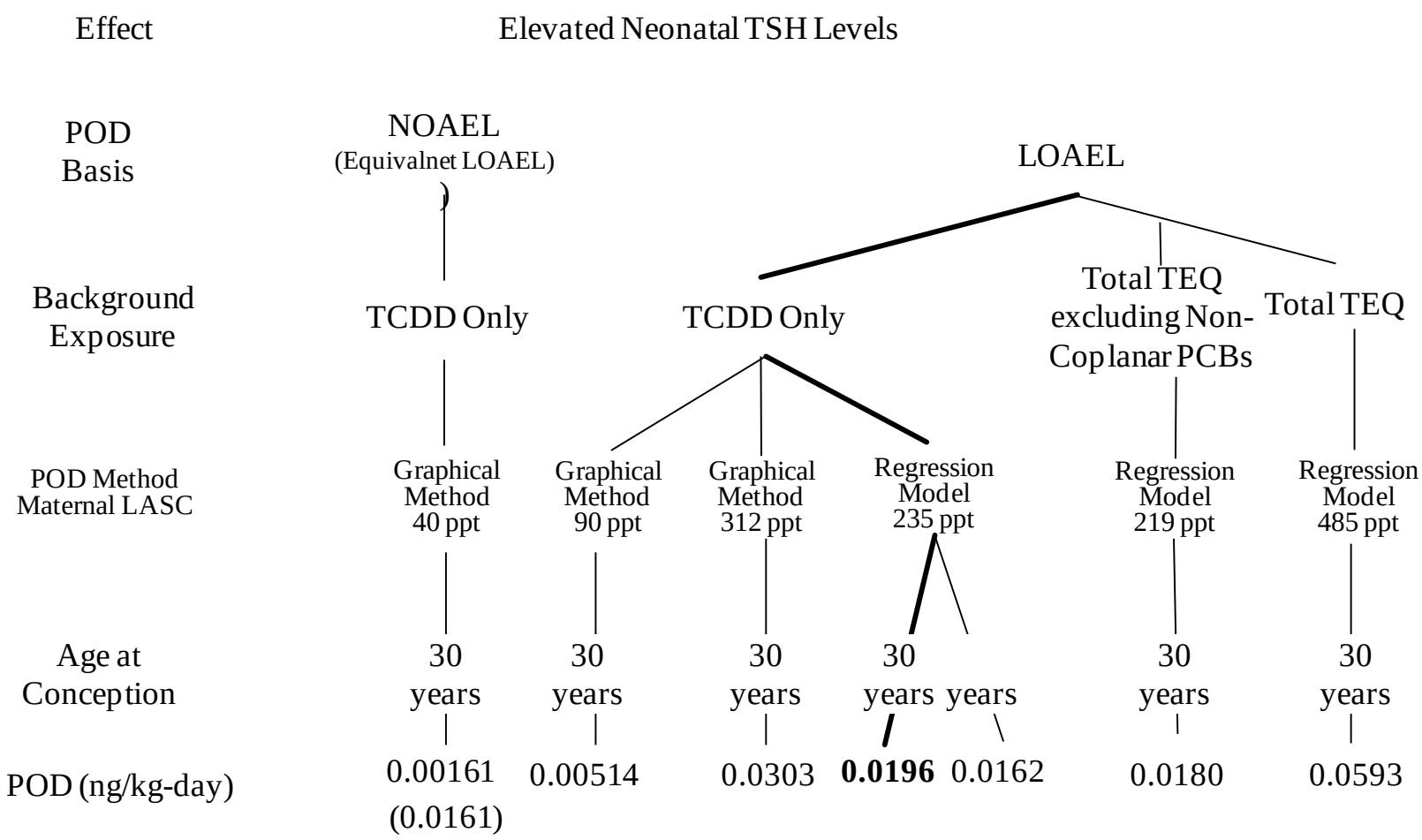

Figure 4-7. Sensitivity tree showing TCDD exposure-variable uncertainty for Baccarelli et al. (2008).

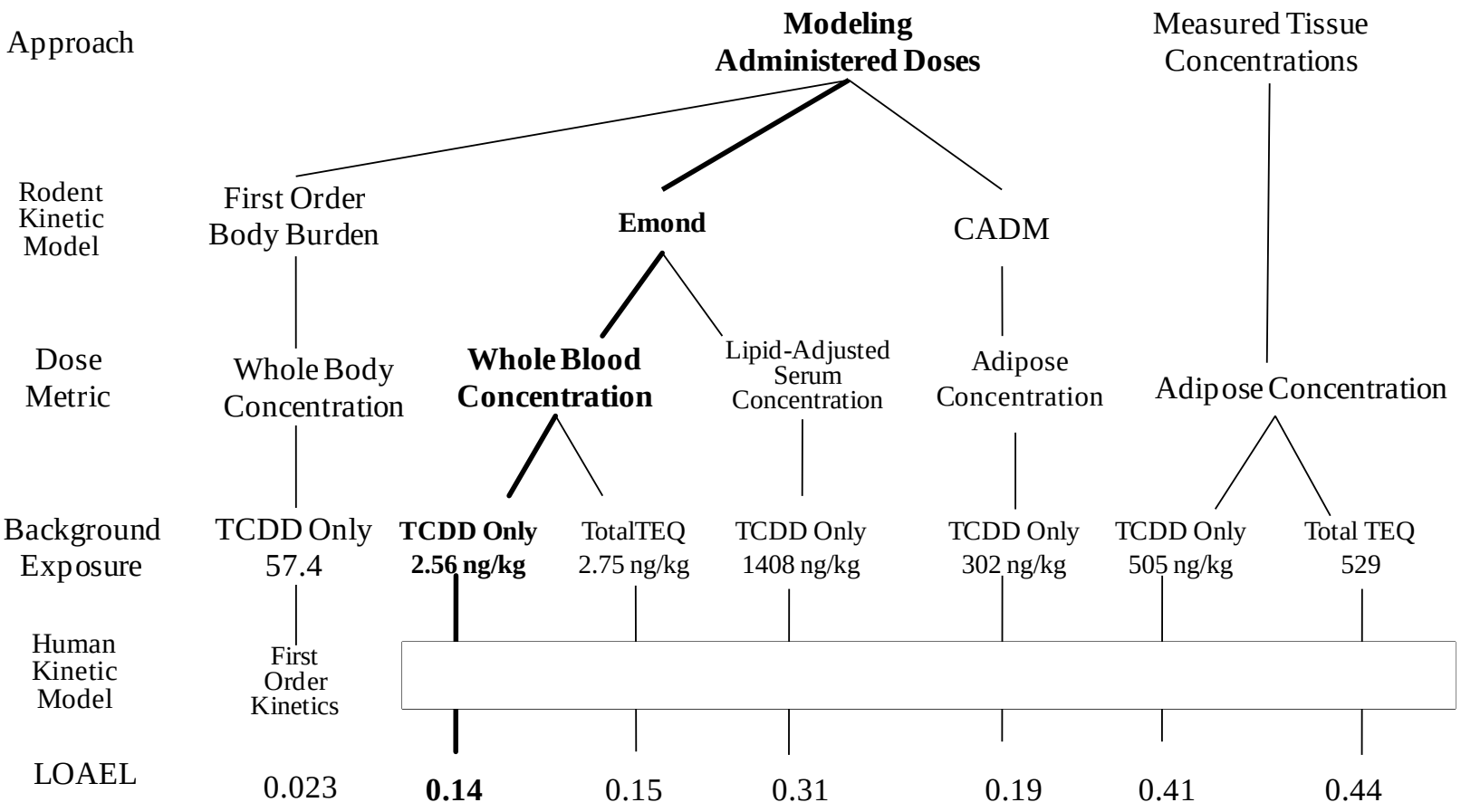

Figure 4-8. Sensitivity tree showing TCDD exposure-variable uncertainty for NTP (2006a).

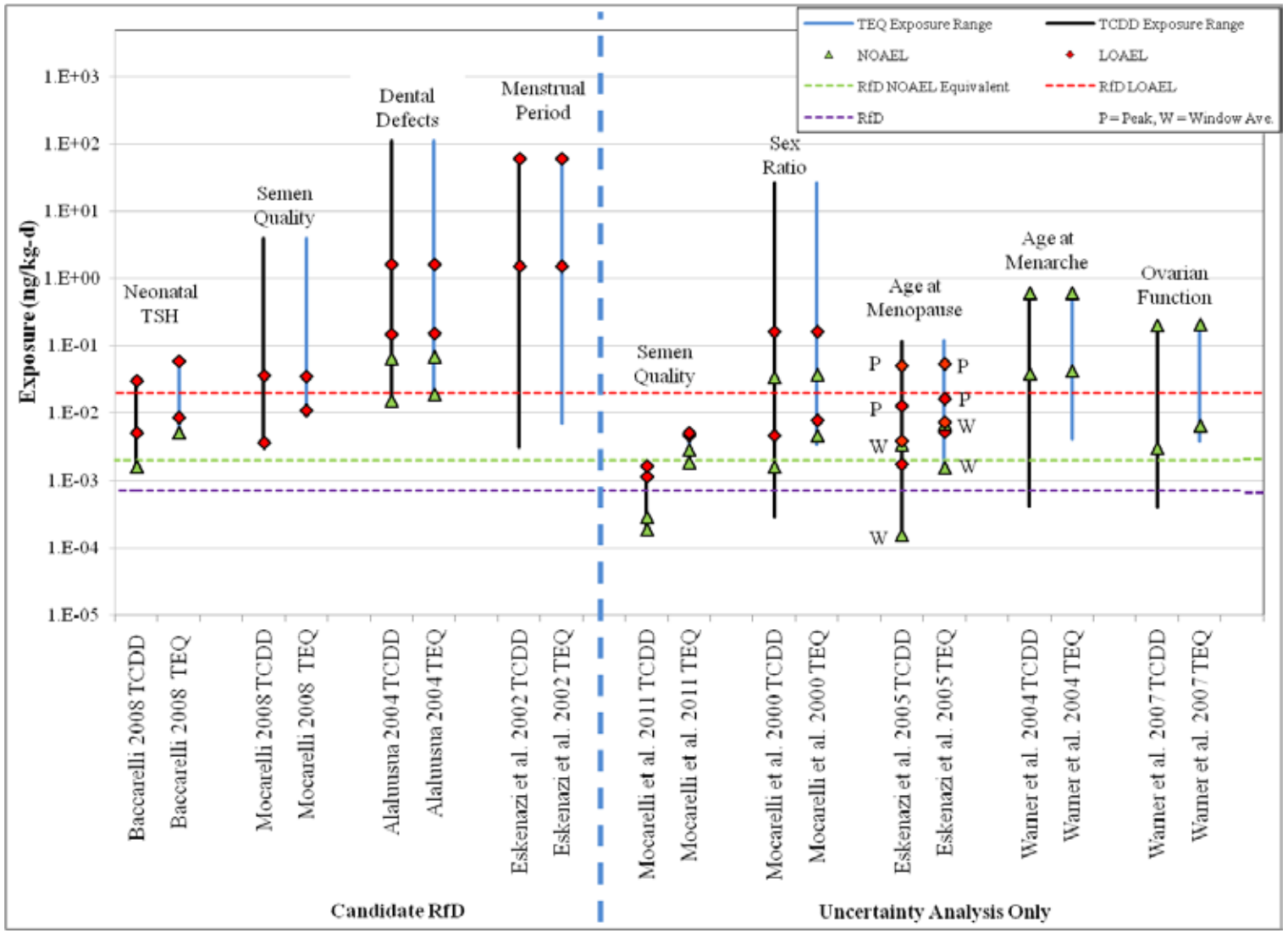

Figure 4-9. Alternative POD exposure-response array.
W = critical window average, P = peak exposure.