Use of Thyroid Disease Incidence and Dose-Response Analysis to Reduce Uncertainty in the Dioxin Oral Reference Dose

Goodrum, PE[1], Budinsky RA[2], Mendelsohn E[3], and Summers H[4]

[1]Integral Consulting Inc., Fayetteville, NY, USA, 13063, pgoodrum@integral-corp.com; [2]The Dow Chemical Company, Midland, MI, USA, 48674; [3]Integral Consulting, Inc., Annapolis, MD, USA, 21401; [4]Integral Consulting, Inc., New York, NY, USA, 10006.

Introduction

In 2012, the USEPA presented a reanalysis of the key issues related to dioxin toxicity in response to comments and recommendations from the National Academy of Sciences (NAS).[1] In response to the NAS and a Science Advisory Board, the USEPA conducted a quantitative uncertainty analysis that explored plausible ranges in the points of departure (POD) from two co-critical epidemiologic studies[2,3] by modeling average daily intakes, and the basis for uncertainty factors. Subsequently, in 2014, NAS provided further recommendations to USEPA for an improved framework for incorporating variability and uncertainty in the derivation of toxicity values. Specifically, NAS recommended 1) giving greater weight to Benchmark Dose (BMD) analysis over the use of NOAELs and LOAELs; 2) more directly linking an understanding of disease processes, mechanisms, and human heterogeneity in adverse outcomes; and 3) using Bayesian statistics to improve estimates of the parameters of dose-response models to integrate study-specific information with other lines of evidence.[4] In this presentation, we illustrate the application of the NAS recommendations to one of the co-critical studies identified by USEPA in the derivation of the RfD for TCDD. This analysis reduces uncertainty in the RfD and provides greater clarity regarding the magnitude and likelihood of adverse effects, as defined by potential changes in the incidence of congenital hypothyroidism (CH) in newborns.

Both co-critical studies in humans that USEPA relies upon for the derivation of the RfD provide information on indicators of potential disease among a population exposed to TCDD during and shortly after an explosion at a trichlorophenol manufacturing plant in July of 1976 in Seveso, Italy. The study by Baccarelli et al. (2008) reports on thyroid-stimulating hormone (TSH) levels in newborns from mothers in exposure areas and nearby reference areas.[2] Based on this study, USEPA used elevated neonatal TSH as a measurement endpoint for risk to define a POD, but without a clear rationale for defining a threshold as a NOAEL or LOAEL. We address this uncertainty by more directly linking TSH to incidence of disease. TSH is one of the more common serum measurements used in newborn screening programs throughout the world because early intervention is effective at reducing the risk of permanent CH, which can contribute to neurocognitive impairment in early childhood.[5] Permanent CH, associated with extremely elevated neonatal TSH, occurs in approximately one of 4,000 births in iodine-sufficient countries.[5] Extensive research supports current screening protocols by establishing ranges and thresholds of elevated neonatal TSH that are reliable predictors of permanent CH. Moderately elevated TSH levels are less predictive, in part because there are a wide range of conditions that can lead to transient hypothyroidism or transient hyperthyrotopinema (elevated TSH in the presence of normal T4 levels) in which thyroid hormone levels return to normal in a few months or years without intervention. Both the documented TSH screening protocols used by clinical practitioners as well as guidance from USEPA on dose-response analysis can be used to better characterize the relationship between threshold changes in TSH and CH incidence, which supports the selection of a POD for dioxin exposure that is grounded in the understanding of disease processes.[6]

Materials and methods

A literature search identified more than two dozen publications that summarize findings from neonatal screening programs. Incidence rates of permanent CH vary, in part, because of different TSH screening thresholds, typically ranging from 5 to 20 μU /mL. The positive predictive value (PPV), or fraction of neonates who both screen in and are diagnosed with permanent CH, tends to be the most reliable measure to compare statistics because there is more comprehensive documentation and follow-up of these individuals. For this analysis, the study by Saleh et al. (2016) of more than 444,000 births in Canada between 2006-2010 is used to calculate PPV estimates that can be binned into intervals of TSH (17-20, 20-30, 30-40, 50+ μU/mL).[7] PPV was calculated for

each bin as the ratio of positive CH diagnoses to the total count of individuals screened. Predictor values were set as the median TSH from each bin, and a value of 50 µU/mL was used for the highest TSH group, which was not bounded by a maximum concentration. Two parameter and three parameter log-logistic curves were fit to the data using USEPA Benchmark Dose Software (BMDS)[6] and 95 percent confidence intervals were generated with R version 3.4.1 (R Core Team 2017).

Numerous large multi-year screening programs report PPVs at various TSH thresholds between 5 and 20 µU/mL, but without corresponding individual level measurements of TSH, it is unclear how to determine the incremental change in PPV that can be attributed to a change in TSH. This is the fundamental concept from which risk thresholds and corresponding PODs are defined using dose-response analysis concepts. Saleh et al. (2016) provides a rare opportunity to apply the method because variability in incidence of disease can be associated with discrete ranges of neonatal TSH. As noted above, BMDS was used to define the initial relationship between neonatal TSH and PPV for CH in the absence of elevated dioxin exposures.[6] This relationship is needed first, so that a subsequent incremental change in PPV can be attributed to dioxin exposure. Uncertainty in the lower region of the curve (<17 µU/mL, the minimum from Saleh et al.) was addressed by incorporating additional reference population information from Baccarelli et al. (2008) to refine the low dose (TSH) region, and by determining the 95 percent lower confidence interval (BMDL) for TSH at standard benchmark response (BMR) thresholds of 1%, 5%, and 10% changes in risk of CH, as represented by the PPV.

Baccarelli et al. (2008) reports summary statistics of TSH values and CH by exposure areas for 1,014 neonates born to mothers who resided in Seveso in a Reference zone (no exposure associated with the explosion), Zone A (high exposures), and Zone B (low exposures).[2] Summary statistics and a graphical analysis support the assumption that TSH is lognormally distributed with parameters given in Table 1.

Table 1: Lognormal distributions of neonatal TSH (µU/mL) by Zone reported by Baccarelli et al. (2008).[2]

Statistic	Reference	Zone B	Ref and B	Zone A
N	533	425	958	56
GM	0.98	1.35	1.14	1.66
GSD	2.35	2.21	2.29	3.04
Percentiles				
5th	0.24	0.37	0.29	0.27
25th	0.55	0.79	0.65	0.78
75th	1.74	2.31	2.00	3.51
95th	3.99	4.98	4.46	10.34
99th	7.14	8.55	7.84	22.06

N = sample size; GM = geometric mean; GSD = geometric standard deviation; Ref and B is the weighted average of the combined datasets from Reference and Zone B areas.

Baccarelli et al. (2008) also reports the number of diagnosed CH cases by exposure zone for neonates that presented with TSH exceeding a single threshold of 10 µU/mL. Of the 1,014 infants evaluated, eight were reported to have TSH>10 µU/mL, and of these two were diagnosed with CH – one each from Zones A and B. The maximum reported TSH was 14 µU/mL, which means that all eight who screened in had TSH measurements between 10 and 14 µU/mL. Figure 1 illustrates the PPVs for this study together with the dose-response curve fit to data from Saleh et al.[7] PPVs for Reference and the combined Reference + Zone B are consistent with the reference population statistics reported by Saleh et al.[7]

Figure 2 illustrates the predicted CH incidence rates for the three Seveso zones based on alternative plausible TSH cutoff screening values of 0, 5, 10, 15, 20 µU/mL. A Monte Carlo simulation using 1,000,000 iterations was applied as follows: 1) select TSH at random from area-specific lognormal distribution for TSH (Table 1); 2) calculate the PPV using the dose-response curve for non-dioxin exposed populations (Figure 1), and 3) use PPV

as a binomial probability to determine if the individual is diagnosed with CH. PPVs for TSH values below screening values were set to 0 and, therefore, no CH diagnoses occurred below the screening values.

The PPV results for Reference and Zone B were used to inform the lower end of the TSH-PPV curve by incorporating Bayesian statistics, consistent with NAS (2014) recommendations. Bayesian updating of the dose-response slope and intercept parameters was conducted using R (version 3.4.1) and the 'bcrm' package[8], which calculates a posterior distribution of the slope and intercept of the dose-response curve based on a set of new TSH-CH observations. The posterior, updated dose-response model combining the Saleh et al. (2016) and Baccarelli et al. (2008) populations primarily refined the dose-response relationship within the lower TSH range (<17 µU/mL), providing greater confidence in estimating incremental changes in PPV associated with incremental changes in TSH in this region.

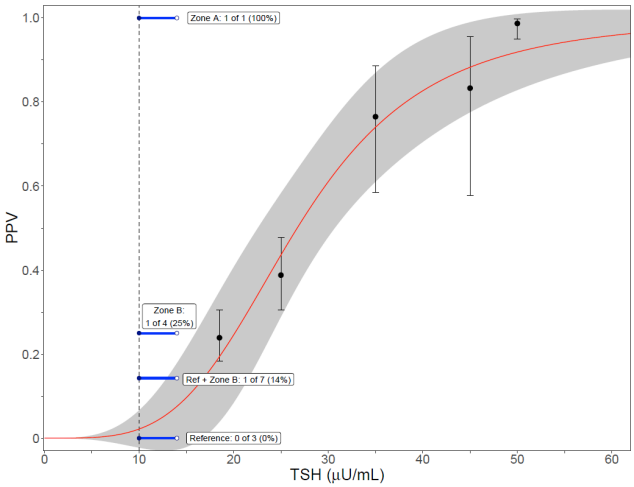

Figure 1: TSH-PPV for CH relationship showing the dose-response curve and 95% CI applied to neonatal screening by Saleh et al. (2016), and incidence data from Baccarelli et al. (2008).[2]

Results and discussion

Results are presented using the 2-parameter log-logistic model to solve for combinations of plausible risk thresholds (1%, 5%, and 10% change in PPV) and the corresponding TSH values. Because the log-logistic model has a non-zero intercept (i.e., PPV is 1.66% when TSH = 0), the BMR is defined by the "added risk" metric, meaning that this baseline risk is added to the incremental change to determine the final BMR. The 95% lower confidence interval for TSH (as represented by the BMDL statistic) at a given BMR provides an additional level of protectiveness. The range of results presents a plausible range of PODs in terms of neonatal TSH levels. To convert this to a corresponding RfD, the same methods and assumptions used by USEPA[1] to derive the current RfD were also applied here. Specifically, 1) a regression equation based on information from Baccarelli et al. (2008) is applied to estimate a maternal serum TCDD that corresponds with a neonatal TSH; 2) a kinetic model is applied to estimate the maternal administered dose that corresponds with the serum TCDD; and 3) a set of uncertainty factors is applied to account for extrapolation uncertainties in the calculation of the final RfD. Uncertainty in these factors has been explored in a separate assessment of additional factors that introduce uncertainties

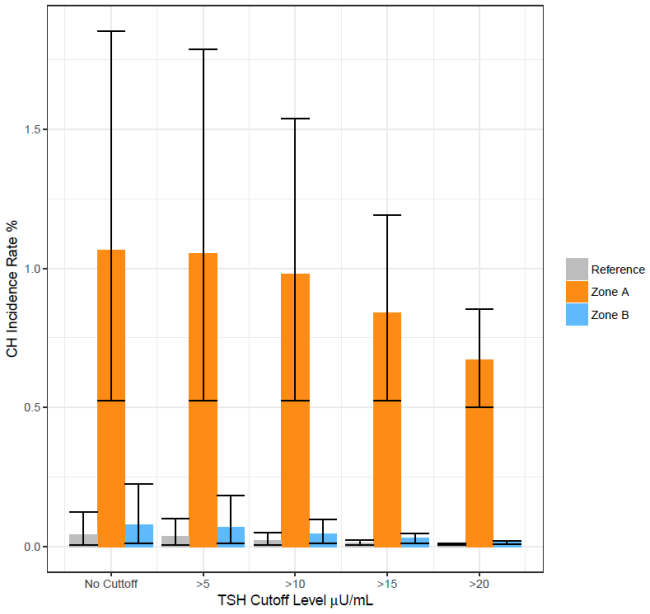

Figure 2: Predicted incidence rates of CH for Seveso populations[2] given lognormal distributions of TSH and the dose-response curve given in Figure 1. Error bars show 5th to 95th percentiles, and overlap with WHO's reported rate of CH for iodine replete populations (1:4000 or 0.03%)[5] for Reference and Zone B.

in the final RfD.

Traditionally, the NOAEL method is used as the POD in regulatory toxicity testing. The NOAEL is defined as the highest tested dose that does not give an effect which is statistically significant from that in the control group. For the current RfD for TCDD, USEPA interpreted a maternal TCDD concentration where the mother-neonatal TSH level exceeded 5 μU/mL to be a LOAEL, which triggered the use of an additional uncertainty factor (UF) of 10 for a LOAEL to NOAEL extrapolation. It may be argued that most national screening programs interpret an initial TSH screening level in the range of 5 μU/mL to be too low to place a neonate at increased risk of CH. Indeed, when 5 μU/mL is applied as a screening cutoff for the Seveso population, predicted incidence rates in both Reference and Zone B are no greater than the baseline rate of CH reported by WHO. The epidemiological literature on subclinical and transient hypothyroidism in neonates is mixed, with some studies suggesting that incremental changes in TSH levels less than 10 μU/mL may increase the risk of CH. The benchmark dose analysis presented here provides a supporting line of evidence that a TSH level of 5 μU/mL is among the lowest plausible BMDLs (i.e., incremental change of 1% in PPV, the benchmark response [BMR]), and that a TSH as high as 13 μU/mL corresponds with a standard predefined change in response (i.e., BMR of 10%) that may not yet be adverse. Using the USEPA's estimates of the relationships between neonatal TSH, maternal serum TCDD, and maternal dose, BMD modeling supports PODs for maternal serum TCDD of between 50 ppt to more than 3,000 ppt, and maternal TCDD doses of 0.002 to 1.59 ng/kg-day. Furthermore, by using human data and quantifying both variability in TSH and uncertainty in the dose associated with a specified benchmark response (i.e., incremental increase in risk of thyroid disease attributable to dioxin exposure), the need for UFs of 3 for interindividual variability and 10 for NOAEL to LOAEL extrapolation are likely no longer needed.

This analysis illustrates that the uncertainty in the RfD can be reduced by applying the NAS recommendations. By linking the POD more directly to empirical data on disease incidence, utilizing dose-response analysis instead of a LOAEL, and applying Bayesian statistics to reduce uncertainty in the dose-response parameters, the final dioxin RfD is likely to be protective at values as high as 530 pg/kg-day, depending on a determination of the incremental change in PPV for CH that may be considered adverse at the population level, and the need for a total UF of 30.

Acknowledgements:
This work was funded by The Dow Chemical Company.

References:

1. USEPA (2012). EPA's Reanalysis of Key Issues Related to Dioxin Toxicity and Response to NAS Comments, Volume 1. U.S. Environmental Protection Agency, Washington, DC.
2. Baccarelli A, Giacomini SM, Corbetta C, et al. (2008). *PLoS Med.* 5(7): 1133-1142
3. Mocarelli P, Gerthoux, PM, Patterson Jr. DG, et al. (2008). *Environ Health Perspect.* 116: 70-77
4. NAS (2014). Review of EPA's Integrated Risk Information System (IRIS) Process. National Research Council.
5. WHO (2007). Assessment of iodine deficiency disorders and monitoring their elimination. A guide for program managers, 3rd ed. World Health Organization, Geneva,
6. USEPA (2012). Benchmark Dose Technical Guidance. EPA/100/R-12/001. Risk Assessment Forum, U.S. Environmental Protection Agency, Washington, DC.
7. Saleh DS, Lawrence S, Geraghty MT, et al. (2016). *BMC Pediatrics*. 16:24
8. Sweeting M, Mander A, Sabin T (2013). *J Statistical Software*. 54(13): 1-26