CMF PRA Subgroup Meeting

Follow on Discussion

**Introduction**

The topics listed below were raised during a Contaminated Media Forum (CMF) Probabilistic Risk Assessment (PRA) subgroup meeting on June 23, 2014. During those discussions it was suggested that DEP provide some direction and context to better frame the issues. The narrative below is intended to provide that direction and context, and is presented as a basis for further discussion.

**Discussion Topics**

1. [Forward vs backward calculation of Probabilistic Risk Assessment (PRA)](#_Forward_vs_backward)
2. [PRAs based upon both variability and uncertainty distributions and 2D-PRA approaches](#_PRAs_based_upon)
3. [Distributions (uniform, triangular, etc.) based on professional judgment due to lack of available data](#_Distributions_(uniform,_triangular,)
4. [Distribution for toxicity values and toxic equivalency factors](#_Distribution_for_toxicity)
5. [Exposure start age and focus on protection of children for other exposure factors](#_Exposure_start_age)
6. [Adjusting the soil ingestion distribution](#_Adjusting_the_soil)
7. [Relative Bioavailability (literature-based variability distributions versus site-specific bioavailability studies)](#_Relative_Bioavailability_(literatur)

**Discussion**

### Forward vs backward calculation of Probabilistic Risk Assessment (PRA)

The method used to perform a PRA has an effect on the Cleanup Target Level selected as the risk management number. A forward iterative PRA uses the input distributions to iteratively calculate a CTL that meets the target risk level (1E-06) at the 90th percentile. A backward calculation results in a distribution of CTLs from which a target CTL is chosen at the 10th percentile. Under certain circumstances, the two methods can provide different answers even when using the same input distributions.

The magnitude of the difference between the two numbers cannot be predicted but is greater when the variables are correlated. Both methods could be used to determine a risk management number (at the appropriate risk level) but raises the questions of which number to use and how large of a difference is meaningful? Given the variability and uncertainty integral to the distributions used in a PRA, there is no empirical answer to the question of how large a difference is meaningful (or how small a difference is not meaningful). Given that there is no objective method for selecting which number to use, it may be administratively difficult to scientifically justify the selection of one number over the other.

The fundamental notion of a PRA is that you run the simulation to generate a range of risks and iteratively calculate a CTL that corresponds to the desired risk level (1x10-6 for Florida). The primary reason to perform a backward calculation is to avoid iterative calculations.

The standard of practice is to run a forward iterative calculation. DEP has used this approach in the two PRAs which it has undertaken (one of which was the development of new surface water numbers for Chapter 62-302)

Chapter 62-780 also implies a forward iterative PRA by referencing a 90th percentile risk. Although the rule also has a parenthetical reference to a 10th percentile CTL which would typically be generated using backward calculation.

Based on the current standard of practice, prior DEP practice and the implication in the rule, the results of a forward iterative PRA will be accepted (assuming there is consensus on the model inputs, design and methods). If a backward calculation is presented, a forward iterative calculation, using the same input parameters, distributions and assumptions, will be requested and the differences between the two results should be explained. If more than one number is calculated then the lowest of the calculated numbers should be selected as the risk management number, absent compelling justification to use some other value.

This topic will be further discussed at the next CMF PRA meeting and information shared during that discussion may lead to an amendment of the above.

### PRAs based upon both variability and uncertainty distributions and 2D-PRA approaches

The parameters used in a PRA may have distributions of both variability and uncertainty. Variability means that the variation in the parameter estimate is known or understood sufficiently that a distribution can be drawn that reasonably approximates that known variation. Uncertainty is when it is known that the parameter varies (or suspected that it varies) but there is insufficient information to construct a representative distribution or when the precision or accuracy of the parameter estimate is unknown. For example, body weight is an example of a parameter with reasonably well-understood variability and studies that include body weight provide distributions that reasonably approximate the observed variability. However, there is some uncertainty with the individual measurements of body weight based upon, for example, how the measurement was taken and the sensitivity of the scale used. Bioavailibility is an example of a parameter with a different relationship between variability and uncertainty. Bioavailability of a chemical from soil will vary among individuals (due to differences in GI physiology and related factors) and from site to site (due to differences in soil properties and contaminant-soil interactions). There will always be variability among individuals in bioavailability of a contaminant from soil at any given site that can be expressed in the form of a distribution. There is usually also considerable uncertainty with respect to: 1) how exactly bioavailability of a certain contaminant varies among individuals; and 2) how soil properties and contaminant interactions have affected bioavailability at a specific site. These uncertainties can be discussed qualitatively or their potential impact on calculated risks or CTLs can be estimated quantitatively.

Acknowledging that variability and uncertainty represent different aspects of the parameter distributions, they should be treated independently and not combined in a PRA. For the purposes of rule 62-780650(3), the best practice recommendation is to use a 1D PRA to address variability and to use a two-dimensional (2D) PRA for a full evaluation of uncertainty. Consequently, proposed ACTLs that depend upon a quantitative evaluation of uncertainty must be developed using a 2D PRA and may require additional review and justification. However, the rule only requires a discussion of uncertainty and not a formal quantitative evaluation. A conceptual evaluation of uncertainty could be supported by substituting selected point estimates from the uncertainty distribution in a series of one-dimensional (1D) PRAs with consistent treatment of the variability component. Alternatively, a 1D PRA could be run with variability alone and then separately with variability and uncertainty distributions. Comparison of the results of the two runs could also help support a conceptual evaluation of the effects of uncertainty.

### Distributions (uniform, triangular, etc.) based on professional judgment due to lack of data

For some parameters in a PRA, there may not be a consensus distribution because there is not enough data to construct one. In some of these cases there may be some information that suggests a particular distribution. In these cases, it is allowable to use a distribution based upon best professional judgment provided that a defensible technical rationale for the distribution is also provided. A point estimate can also be used in cases where there is no established distribution but doing so can compress the resultant risk distribution compared to using a distribution based on best professional judgment.

### Distribution for toxicity values and toxic equivalency factors

EPA specifically recommends against using distributions for toxicity values and recently re-affirmed this position in a comment on this PRA discussion. Distributions of toxicity values were also not used in the recent rulemaking for 62-302. Therefore, using a distribution for toxicity values is not recommended at this time for estimation of risk or calculating risk-based target levels. Using the “benchmark dose approach” to model uncertainty was discussed as a possible alternative to using a distribution of toxicity values, but further discussion on this topic is needed.

Distributions for dioxin toxic equivalency factors (TEFs) from Haws et al., 2006 are provided in the 2010 EPA document “Recommended Toxicity Equivalence Factors (TEFs) for Human Health Risk Assessments of 2,3,7,8-Tetrachlorodibenzo-p-dioxin and Dioxin-Like Compounds”. These TEF distributions represent uncertainty rather than variability regarding the toxic potency of dioxin congeners. The EPA considers these values suitable for sensitivity analyses, but not for conducting a quantitative uncertainty analysis or calculating confidence limits. Therefore, using TEF distributions is also not recommended at this time for estimation of risk or calculating risk-based target levels.

### Exposure start age and focus on protection of children for other exposure factors

The start age for exposure can be used as a variable in a PRA. Intuitively, exposure to a contaminant can begin at any age and so a random distribution of start age may seem appropriate. However, use of a random start age does not account for the fact that children are a more susceptible population and therefore tend to be represented in the upper tail of the risk distribution when using a random start age. Consequently, the number obtained from the PRA would not be considered protective of children. To ensure that children are adequately protected it is necessary to constrain start age so this sensitive population is protected.

Three options were discussed with regard to how to model start age all of which help improve the protectiveness to children, although to somewhat varying degrees. The first option is to use a random start age between 1 and 6. Alternatively, migration data for children in the 1-6 age range can be used to develop a more refined estimate (see the US EPA Exposure Factors Handbook, 2011 ed, section 16.5 “Population Mobility” for more information). The third option is to simply set the start age at 1. This latter method is current EPA practice for risk assessments and is the assumption the department made when originally developing the CTLs in Chapter 62-777 and so is acceptable for PRAs. However, use of the migration data distribution for selecting a start age between 1 and 6 would also be acceptable. The use of a random start age between 1 and 6 is not recommended because the existing migration data indicates that start age is not random.

It is important to note that the effects of start age on the estimate of risk varies by the type of contaminants under consideration and is of greater importance when considering those with cancer risk, especially mutagenic carcinogens (e.g., PAHs). n this latter case a start age of 1 is recommended.

This is scheduled to be discussed further at the next CMF PRA meeting.

### Adjusting the soil ingestion distribution

There are generally-accepted mean and upper-percentile incidental soil ingestion rates published by EPA in the Exposure Factors Handbook. These ingestion rates can be included in a distribution that represents the known population behavior for ingestion of soil and is generally divided into adult and child distributions, with children tending to ingest greater amounts of soil. In a PRA, it is possible to repeatedly sample this distribution for an individual to reflect the likelihood that incidental soil ingestion patterns vary with time. As part of the discussion, there seemed to be general consensus that these successive draws should be somewhat to highly correlated because soil ingestion patterns are not expected to dramatically change suddenly but rather could drift over time. No specific degree of correlation was agreed upon but values ranging from an R2 of .75 to .95 were discussed.

There are two different ways the correlation could be applied. The first is to always correlate with the original draw from the distribution and this method would tend to concentrate successive draws to be relatively close to the original draw. The second method is to apply the correlation sequentially, so that each draw is only correlated with the draw immediately prior. This latter method could lead to a greater degree of drift from the original draw than the first method.

Regardless of the method of correlation, iterative draws from the incidental sol ingestion rate distribution have the effect of collapsing the tails of the overall risk distribution because soil ingestion becomes represented by an approximation of the central tendency of the distribution rather than allowing it to vary across the entire range of values. The more draws that are made from the distribution, the more pronounced the collapsing effect becomes.

There seemed to be consensus that a draw for children and a second draw for an adult would be appropriate because these distributions are different, suggesting different soil ingestion behavior between children and adults. Suggestions were also made that multiple draws for children could also be made to reflect changing behavior of the child. Several suggestions were made with regard to resampling frequency including: six times per year, once per year, and once per age bin. However, there is no empirical data to suggest what the number of iterative draws should be and the collapse towards central tendency becomes more pronounced as more draws are made, thus potentially increasing the chance of underestimating a substantial portion of the child population.

It was noted that if the resampling frequency is high then the correlation between successive draws should also be high (e.g., 90-95%) because it is less likely for behavior to change over a short time period than a longer one.

Whichever approach is taken, it must be ensured that the distributions still have sufficient tails and the PRA should include a discussion of how that was achieved and any potential biases introduced through the modeling approach used. The discussion should include an explanation of how the resulting distribution is still representative of the original distribution and the sensitivity of the distribution to changing correlation and resampling frequency.

This topic is scheduled for further discussion at the next PRA CMF meeting.

### Relative Bioavailability (literature-based variability distributions vs. site-specific studies)

Some concerns were expressed over current agency practice to assume a relative bioavailability of 1 unless there is a site-specific study available that identifies a different number. At issue is the number of bioavailability studies that have been done for various contaminants and making sure we include the current state of knowledge in the PRA calculations. However, existing bioavailability studies highlight the variability in these estimates making it difficult to draw consistent conclusions among different studies. It would be useful to find a way to incorporate other information on bioavailability from non-site-specific studies. *In vitro* methods for bioavailability for lead have been accepted by EPA but such methods have not yet been well validated for other contaminants.

Examples of the range in bioavailability include the data for lead which shows a range of <10% to over 90% in different studies. Studies with the bioavailability of dioxin congeners in rat and swine give opposite results with regard to how the degree of chlorination affects bioavailability.

For some chemicals, perhaps most notably dioxin, there is sufficient data to establish that the bioavailability is something less than 1. However, there isn’t enough data to determine what the alternative value should be and so EPA has recommends the default bioavailability of dioxin remain at 1.

It may be possible to address bioavailability as uncertainty in the second dimension of a 2D PRA. In doing so, a justification must be provided for the inclusion or exclusion of particular bioavailability estimates.

This is scheduled for further discussion at the next CMF PRA meeting.