www.epa.gov

Supplement: Uncertainty Analysis of In Vitro Metabolic Parameters and of In Vitro to In Vivo Extrapolation (IVIVE) Used in a Physiologically Based Pharmacokinetic (PBPK) Model for Chloroprene

July 2020

Office of Research and Development
U.S. Environmental Protection Agency
Washington, DC

Disclaimer

This document is a public comment draft for review purposes only. This information is distributed solely for the purpose of public comment. It has not been formally disseminated by EPA. It does not represent and should not be construed to represent any Agency determination or policy. Mention of trade names or commercial products does not constitute endorsement or recommendation for use.

Authors

Paul Schlosser, PhD	ORD/CPHEA/CPAD
Dustin Kapraun, PhD	ORD/CPHEA/CPAD

List of Figures

Ramboll (2020) presents analyses of the in vitro incubation data used to estimate metabolic parameters for chloroprene (Yang et al., 2012; Himmelstein et al., 2004). Their analyses identify mean or maximum likelihood values of those parameters, but also include confidence intervals for those estimates. The IVIVE calculations and PBPK model predictions presented by Ramboll (2020) are based on the mean parameter values and do not estimate the quantitative uncertainty in the PBPK model. In particular, the uncertainty resulting from uncertainty in the in vitro metabolic parameters and IVIVE calculations, was not presented. The U.S. Environmental Protection Agency (EPA) seeks to quantify uncertainty in predictions and estimates associated with the PBPK model.

This document describes initial results for one component of such an analysis and an approach for completing the full uncertainty analysis that would occur if the model is used by the EPA in an updated IRIS Toxicological Review for chloroprene. The uncertainty analysis of the in vitro data shown and proposed here is based on the premise that the experimental unit of observation is the incubation vial and that variation in results between one incubation vial and another that is not explained by concentration-dependence (which is captured by the in vitro model) represents uncertainty in the corresponding parameters. Since the data are obtained using pooled tissue samples, vial-to-vial variability does not represent inter-individual variability, but experimental variability, which is one source of parameter uncertainty.

The following material is divided into two main components: (1) uncertainty analysis of the metabolic parameter estimation from in vitro experimental data (including both the proposed analysis and some initial results); and (2) proposed uncertainty analysis of the IVIVE calculation and subsequent PBPK model predictions. For the uncertainty analysis of the in vitro metabolism data, there are two sub-components: (1a) analysis of uncertainty in the background system loss (including initial results); and (1b) analysis of uncertainty in the in vitro metabolic rate constants. Initial results from the uncertainty analysis indicate that the overall uncertainty range in some parameters may differ from the 95% confidence interval estimated for the *mean* value reported by Ramboll (2020).

(1) Uncertainty Analysis of Metabolic Parameter Estimation From In Vitro Experimental Data

(1a) Analysis of Uncertainty of Background Loss in the In Vitro Experimental System

Context

- EPA used data collected during in vitro experiments designed to assess rates of chloroprene metabolism in human liver and lung tissues (Yang et al., 2012; Himmelstein et al., 2004). For each "incubation," 1 mL of liquid media containing 1 mg of metabolic proteins was placed into a vial with a total volume of approximately 12 mL. Chloroprene was then placed into the headspace (air) of each closed vial, and the system inside the vial was allowed to come to equilibrium. Here, "equilibrium" means that the ratio of the concentrations of chloroprene in

media and air was stable, and it was assumed that this ratio was 0.69 (Himmelstein et al., 2001). Samples of air with volumes 0.2 mL (Yang et al., 2012) or 0.4 mL (Himmelstein et al., 2001) were removed at 0.0, 0.2, 0.4, 0.6, 0.8, and 1.0 hours after some initial time following establishment of equilibrium.

- The in vitro experiments of Yang et al. (2012) and Himmelstein et al. (2004) both included control experiments which used either: only phosphate buffer, heat-inactivated microsomes, or active microsomes but no $NADP^+$ co-factor (Yang et al., 2012; Himmelstein et al., 2004)[1]. The experiments provide information on the background rate of loss of chloroprene from the in vitro system. The instantaneous rate of background loss is assumed to be proportional to the concentration of chloroprene in the media (µmol/L) with constant of proportionality given by the parameter for the rate of background loss, RLOSS (L/h). Assessing uncertainty and variability in background rate of loss of chloroprene (as characterized by the parameter RLOSS) is important for understanding uncertainty and variability of metabolic rates as informed by the non-control experiments of Yang et al. (2012) and Himmelstein et al. (2004).

For all parameter estimation analyses described herein, we used the same two-compartment pharmacokinetic (PK) model of the in vitro experimental system as was used by Ramboll (2020). The initial derivation of the general model was described in detail by Kreuzer et al. (1991) for butadiene. The mass balance differential equation for the air phase in the current model is

$$\frac{dA_{\text{air}}}{dt} = K_{\text{gl}} \cdot \left(\frac{C_{\text{med}}}{P} - C_{\text{air}}\right) \quad (1)$$

and the mass balance differential equation for the incubation media (liquid) phase is

$$\frac{dA_{\text{med}}}{dt} = K_{\text{gl}} \cdot \left(C_{\text{air}} - \frac{C_{\text{med}}}{P}\right) - \frac{V_{\max} \cdot C_{\text{med}}}{K_{\text{m}} + C_{\text{med}}} - k_{\text{f}} \cdot C_{\text{med}} - R_{\text{loss}} \cdot C_{\text{med}}, \quad (2)$$

where

- A_{air} and A_{med} (µmol) are the respective amounts of chloroprene in the air and media;
- $C_{\text{air}} = A_{\text{air}}/V_{\text{air}}$ (µM) is the concentration in the air;
- $C_{\text{med}} = A_{\text{med}}/V_{\text{med}}$ (µM) is the concentration in the media;
- $V_{\text{air}} =$ V_air (L) and V_{med} = V_med (L) are the respective volumes of the air and media;
- K_{gl} (L/h) is the air-to-media mass transfer coefficient;
- P (no units) is the media-to-air equilibrium constant;
- $V_{\max} =$ VMAX (L/h) is the maximum velocity of enzyme-mediated metabolism in the media;
- $K_{\text{m}} =$ KM (µM) is the saturation constant for metabolism;
- $k_{\text{f}} =$ KF (L/h) is an alternate metabolic first-order rate constant for cases when saturation is not observed; and
- R_{loss} = RLOSS (L/h) is the rate constant for non-specific loss of chloroprene from the incubation vial.

[1] The methods section of Himmelstein et al. (2004) states, "Control incubations were performed without $NADP^+$ or with $NADP^+$ and heat-inactivated microsomes," however the legend of Figure 1 in that paper describes the data as "control loss in phosphate buffer." The study report describing the data used by Yang et al. (2012) states, "Control incubations were run in the absence of $NADP^+$ by addition of an equal volume of phosphate buffer," which the EPA has presumed to mean that phosphate buffer replaced the $NADP^+$ solution, but active microsomes were used.

- This model differs from the model used by Yang et al. (2012) and Himmelstein et al. (2004) in that there are separate mass balance differential equations for the air and media phases (Equations 1 and 2), while Yang et al. (2012) and Himmelstein et al. (2004) used a single differential equation for concentration in the media under the simplifying assumption that the air and media phases were continually at equilibrium; i.e., that $C_{\text{air}} = C_{\text{med}}/P$ at all times.
- The non-specific loss term appears in the media equation here since it appeared in the corresponding equation of Yang et al. (2012) and Himmelstein et al. (2004), making the value of RLOSS obtained here comparable to those previous analyses. However, it is recognized that some loss may be due to leakage around or through the vial cap or septum.
- Supplemental Material B of Ramboll (2020) provides additional details on the estimation of K_{gl}. The zip file that can be downloaded from the U.S. EPA HERO link for Ramboll (2020) contains three sub-folders, of which "Ramboll 2020 Chloroprene In Vitro Code.zip" contains the source code for the Ramboll (2020) version of the model. Source code for the preliminary analyses described here is not currently available due to restricted access to the U.S. EPA offices, where the computer containing the code is housed, in response to the COVID19 pandemic.

Estimation of RLOSS Using Yang et al. (2012) Control Data

- To assess uncertainty and variability of RLOSS, EPA analyzed data from the control experiments of Yang et al. (2012). In these experiments, 30 incubations were used to collect data. For each incubation, a sample injection volume, VINJ = 0.2 mL, of headspace air was removed from the vial by inserting a syringe into the membrane covering the vial, filling it, and then removing it at each of five times (0.0, 0.2, 0.4, 0.6, 0.8, and 1.0 h). This volume of removed air was injected into a gas-chromotography/mass-spectroscopy (GC/MS) system and the amount of chloroprene in the injection was assessed.
- The total volume of each vial (VVIAL) was assumed to be 11.6 mL, and the volume of liquid media was V_med = 1.0 mL. Thus, for most of the time for any given experiment, the volume of the headspace was V_air = VVIAL – V_med = 10.6 mL. However, during the instant just before the air-filled syringe was withdrawn, the volume drawn into the syringe, VINJ, was contiguous with the volume of the headspace, V_air. Thus, the total volume of the air in the system was at that moment V_air + VINJ. If we assume a homogeneous concentration in the air at the time when the syringe was removed, the total proportion of headspace chloroprene removed was equal to VINJ / (V_air + VINJ); i.e., the total volume in the syringe divided by the total volume in the headspace and the syringe just before it was withdrawn.
- EPA assumed that the concentration in the headspace volume V_air immediately following the removal of the sampling syringe was equal to the concentration in the sampling syringe, which was quantified and recorded by Yang et al. (2012). Thus, EPA sought to calibrate the model by determining a value of the parameter RLOSS that yielded optimal agreement between the observed sample concentrations and the model predicted headspace concentrations immediately following the removal of the sampling syringe.
- As a first-pass calibration, EPA used an ordinary least squares (OLS) methodology to obtain optimal values of RLOSS for each incubation. That is, EPA calculated values of RLOSS and A0 (initial amount of chloroprene in the vial) that cause the sum of squared residuals (SSR), where "residuals" are differences between observed and predicted concentrations, to be minimized.

The optimal model fit and the observed data for one of the Yang et al. (2012) control incubation data series is shown in Figure 1.

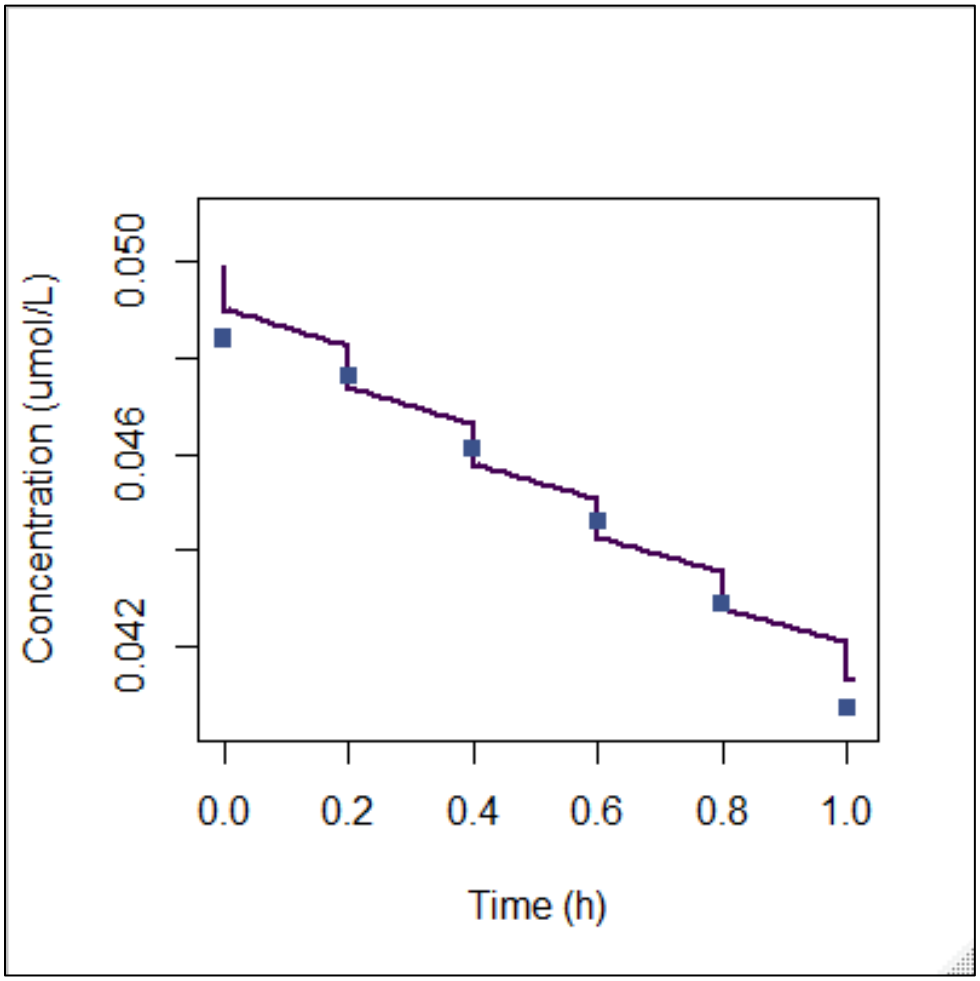

Figure 1. Optimal OLS model fit and the observed data for one of the Yang et al. (2012) control incubation data series.

- Summary statistics for the 30 OLS estimates of RLOSS were:
  - mean = 0.00146944 L/h
  - median = 0.001512171 L/h
  - standard deviation = 0.0005741165 L/h
  - minimum = 0.0003616011 L/h
  - maximum = 0.002397292 L/h
- Following the OLS calibration, the EPA generated a residual plot (residual vs. model-predicted concentrations). The residual plot is shown in Figure 2.

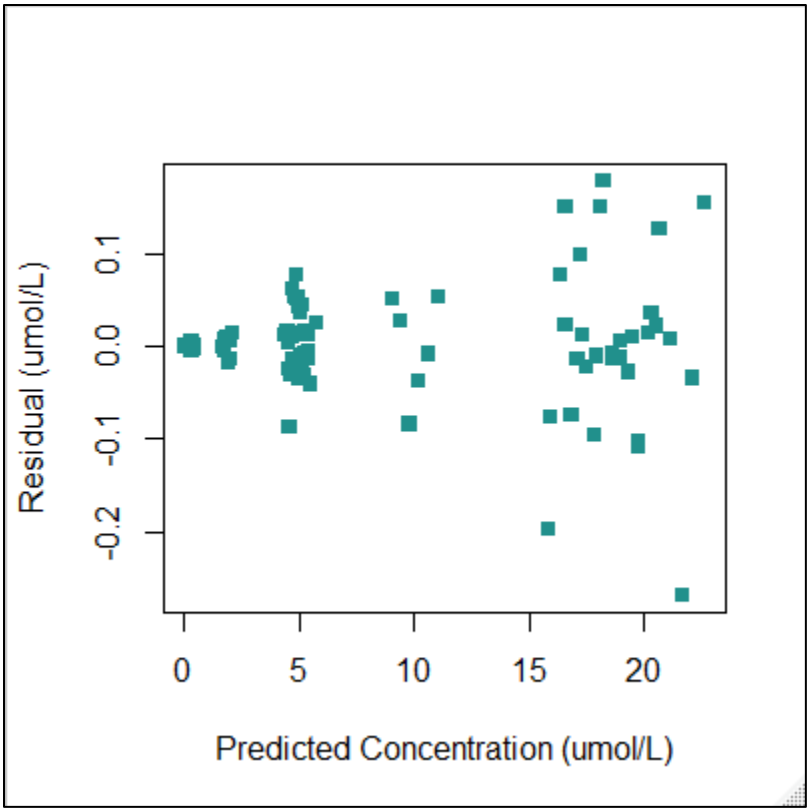

Figure 2. Residuals versus predicted concentrations for the control experiments of Yang et al. (2012).

- The residual plot shows that the magnitude of the residuals tends to be larger for larger model-predicted concentrations, suggesting that any statistical error model used to describe the relationship between observed concentration data and model-predicted concentrations should assume larger errors for larger model-predicted values. One such model would assume that the differences between the logarithms of the concentration data and the logarithms of the model-predicted concentrations are normally distributed with a mean of zero and some nonzero standard deviation. To test the plausibility of this statistical error model, EPA constructed a modified residual plot in which the modified residuals were computed as the differences of the logarithms of the observed and model-predicted concentrations. This modified residual plot is shown in Figure 3.

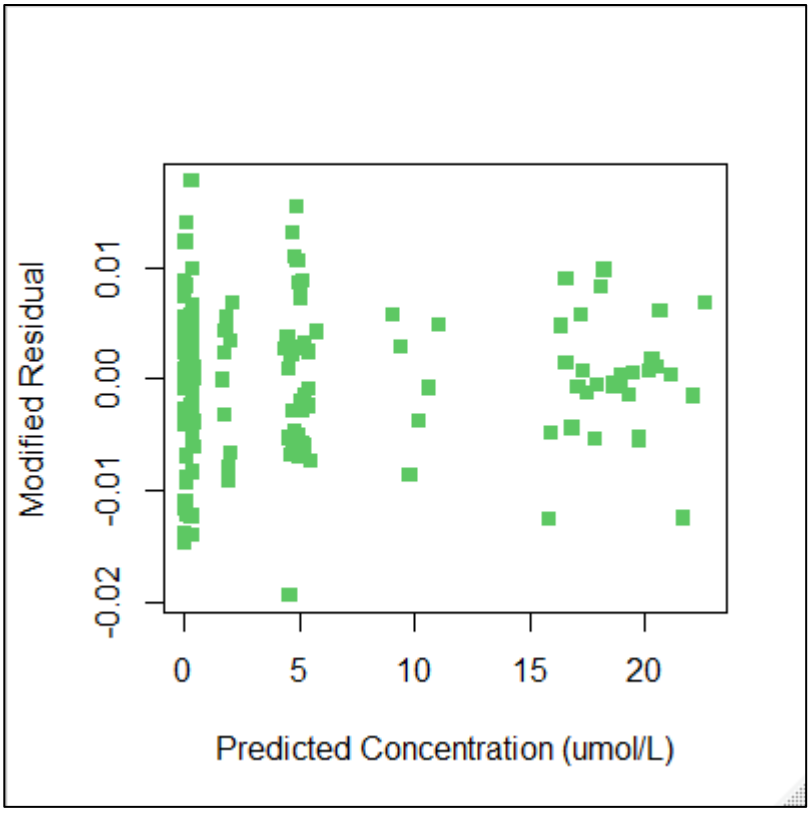

Figure 3. Modified residuals versus predicted concentrations for the control experiments of Yang et al. (2012).

- The modified residual plot indicates that the modified residuals do not vary substantially as model-predicted concentrations increase.
- The modified residuals for the Yang et al. (2012) control data and the OLS-calibrated model predictions have a mean near zero (-2.67×10^{-5}) and a standard deviation of 0.00626. Furthermore, based on the quantile-quantile plot shown in Figure 4, these modified residuals appear to be normally distributed.

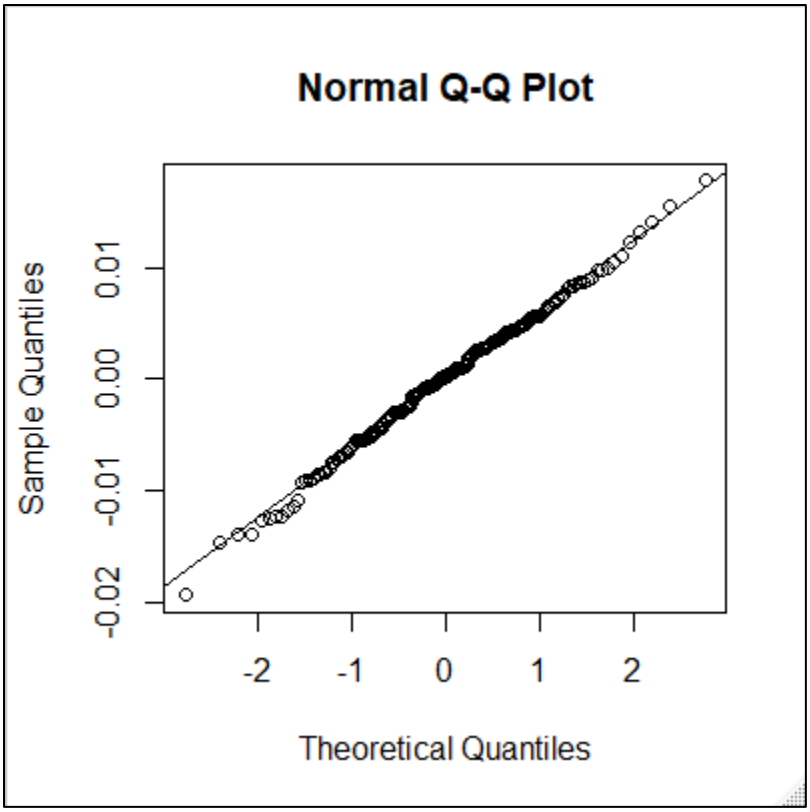

Figure 4. Quantile-quantile plot for the modified residuals for the Yang et al. (2012) control data and the OLS-calibrated model predictions.

- Working on the assumption that the modified residuals are samples of a normally distributed random variable with a mean of zero and some constant variance, EPA selected a statistical error model of the form

$$\log[C_{\text{dat}}(t)] = \log[C_{\text{mod}}(t;q)] + \varepsilon, \tag{3}$$

where $C_{\text{dat}}(t)$ is the observed concentration at time t, $C_{\text{mod}}(t;q)$ is the model-predicted concentration at time t assuming model parameters have values given in the vector q, and $\varepsilon \sim \mathcal{N}(0,\sigma)$; i.e., ε is a random variable that is normally distributed with mean zero and standard deviation σ.

- There are $N = 6$ data points in any given data series (one for each of the time points 0.0, 0.2, 0.4, 0.6, 0.8, 1.0). Thus, the likelihood for the parameters $\theta = (q,\sigma)$ given the data $\mathcal{D} = \left\{(t_j, C_{\text{dat}}(t_j)) : j \in \{1,\dots,N\}\right\}$ is given by

$$\mathcal{L}(\theta|\mathcal{D}) = \prod_{j=1}^{N} \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{\left(\log[C_{\text{dat}}(t_j)] - \log[C_{\text{mod}}(t_j;q)]\right)^2}{2\sigma^2}\right]. \tag{4}$$

- With this likelihood function, EPA calibrated the model using a maximum likelihood estimation (MLE) methodology to obtain optimal values of RLOSS for each incubation. That is, EPA calculated values of RLOSS and A0 that cause the value of the likelihood function to be

maximized (assuming a fixed value of 0.006 for σ). The optimal model fit and the observed data for one of the Yang et al. (2012) control incubation data series is shown in Figure 5.

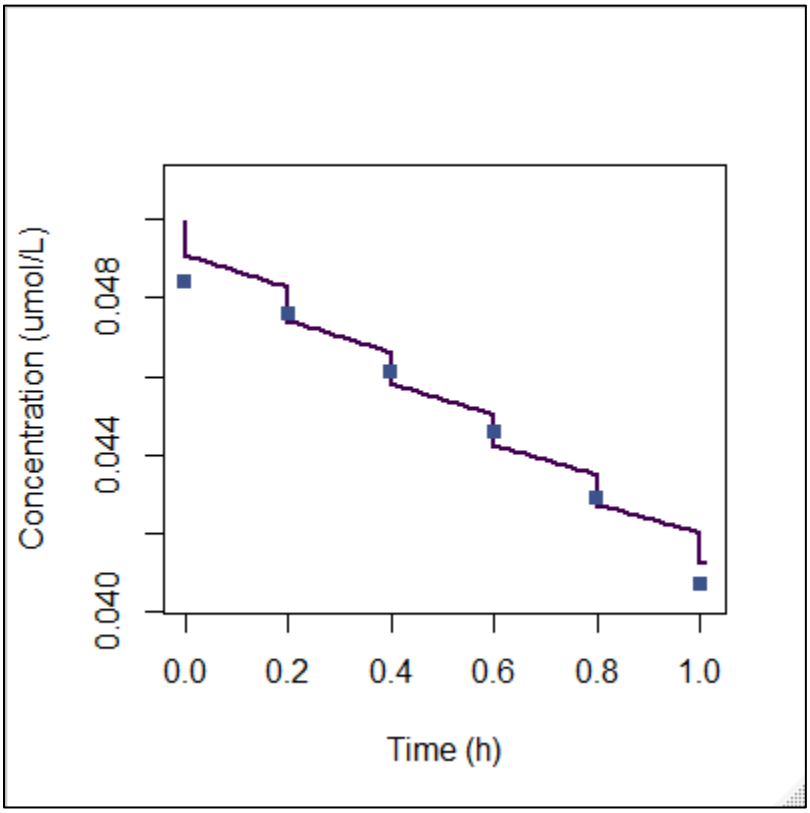

Figure 5. Optimal MLE model fit and the observed data for one of Yang et al. (2012) control incubation data series.

- Summary statistics for the 30 MLE estimates of RLOSS were:
  - mean = 0.00147567 L/h
  - median = 0.001536144 L/h
  - standard deviation = 0.0005733221 L/h
  - minimum = 0.0003642862 L/h
  - maximum = 0.002393498 L/h
- Note that these summary statistics are quite close to those calculated for the 30 OLS estimates of RLOSS and that they differ only in the third significant figures.
- For $\theta = (q, \sigma) = (\mathrm{A0}, \mathrm{RLOSS}, \sigma)$, EPA defined a vaguely informative prior that ensures A0, RLOSS, and σ are all non-negative and that assigns higher probability to smaller values of σ using a Jeffreys prior. In particular, the prior is given by

$$\pi_0(\theta) = \begin{cases} \frac{1}{\sigma^2} & \text{for } \mathrm{A0} \in (0,\infty), \mathrm{RLOSS} \in (0,\infty), \sigma \in (0,\infty), \\ 0 & \text{otherwise.} \end{cases} \quad (5)$$

- EPA used the likelihood function and prior defined in Equations 4 and 5, respectively, and the Yang et al. (2012) control data to generate samples from the Bayesian posterior distribution for θ via Markov chain Monte Carlo (MCMC). To do this, EPA used the Delayed Rejection Adaptive

Metropolis (DRAM) algorithm implemented in the R package "FME" to generate samples from the posterior distribution of $\theta = (\mathrm{A0}, \mathrm{RLOSS}, \sigma)$ for each of the 30 Yang et al. (2012) control incubation data sets. Note that the approach described here does not involve a Bayesian hierarchical model but accounts for the fact that the experiments involve serial observations of each incubation vial.

- EPA generated two distinct Markov chains based on two distinct starting points for each data series (i.e., for each incubation). EPA evaluated the convergence of the chains by comparing the Gelman-Rubin potential scale reduction factor (PSRF) for the RLOSS chains to an upper threshold value of 1.1. As of the time of writing of this report, chains for some of the data series had not converged based on this criterion (PSRF less than 1.1). These chains will be continued (using a larger number of iterations) until this criterion is met.
- The distributional estimates for RLOSS for the 30 Yang et al. (2012) control incubation data sets are illustrated as a series of box plots in Figure 6. In these box plots, the outer edges of each box show the 25th and 75th percentiles of the posterior distribution, the center line in each box shows the 50th percentile, and the outermost lines show the 2.5th and 97.5th percentiles.
- *Note: Not all chains had met the convergence criteria described above at the time the plots below were generated, but these plots are believed to be indicative of the variation between control incubations and parameter uncertainty for each.*

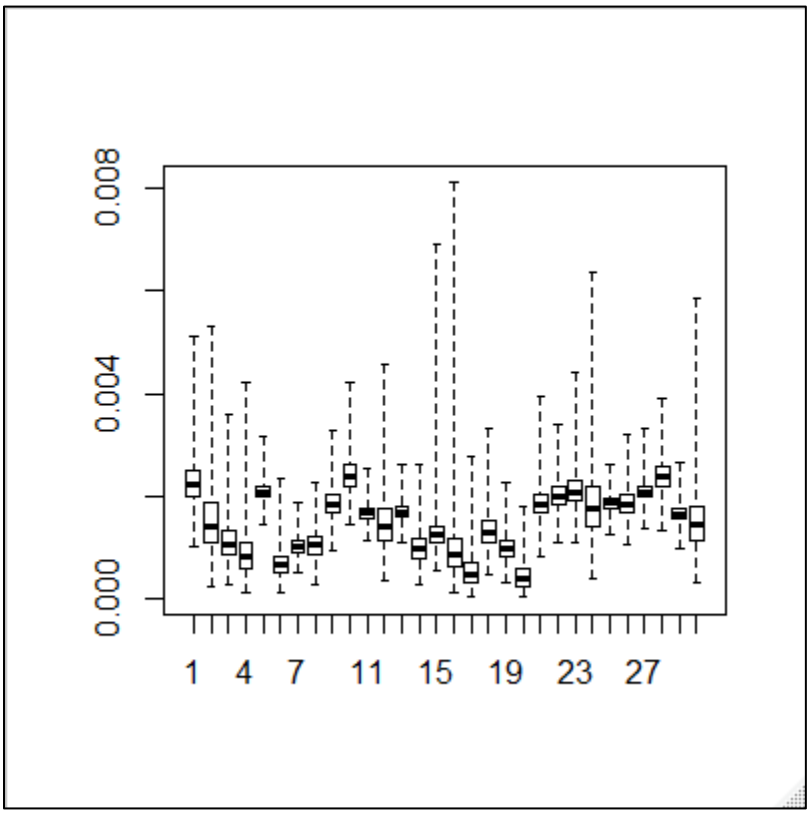

Figure 6. Box plots illustrating distributional estimates for RLOSS for the thirty Yang et al. (2012) control incubation data sets.

- EPA applied kernel density estimation (KDE) as implemented in the base R function "density" to generate probability density curves for RLOSS for the 30 Yang et al. (2012) control incubation data sets. These are shown in Figure 7 as the fine blue (lighter-colored) curves. EPA also pooled all the samples for all 30 data sets and applied KDE to generate an "overall" probability density curve for RLOSS. This is shown in Figure 7 as the thicker purple (darker-colored) curve.

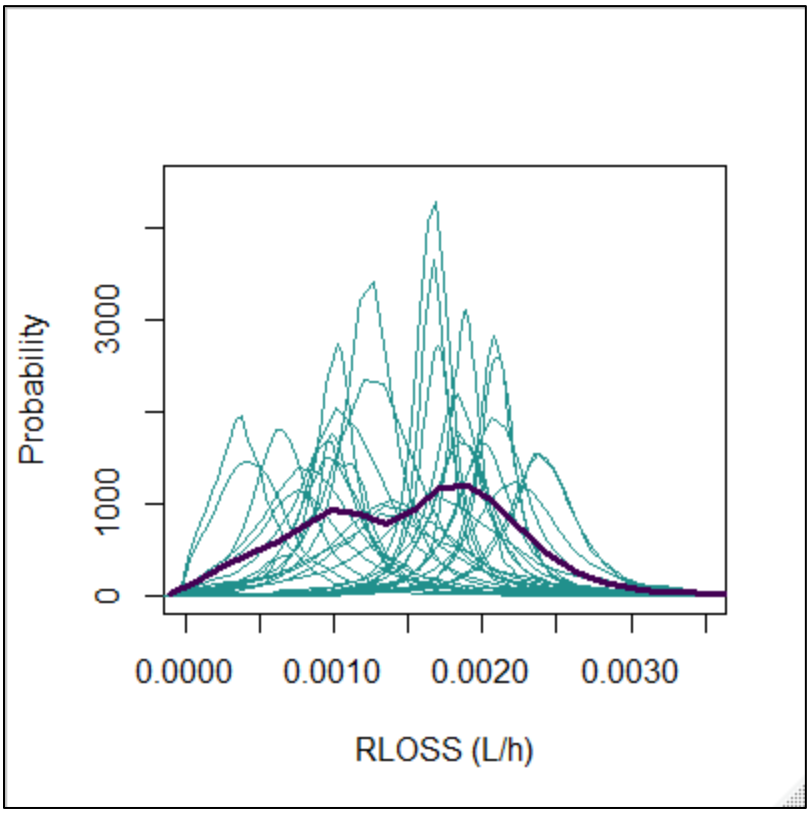

Figure 7. Probability density curves for RLOSS for the thirty Yang et al. (2012) control incubation data sets.

Estimation of RLOSS Using Himmelstein et al. (2004) Control Data

- To assess uncertainty and variability of RLOSS from the experiments of Himmelstein et al. (2004), EPA analyzed control data (for background rates of loss of chloroprene in phosphate buffer[2]) shown in their Figure 1A. In these experiments, five incubations were conducted with

[2] Figure 2A of Himmelstein et al. (2004) also shows results for incubations with mouse lung microsomes without NADP+ and with two inhibitors. Two of the control incubations show air space concentrations of chloroprene increasing 15–20% between the initial and second time point, which is unexpected because no new chloroprene material is being created or introduced during the in vitro experiments. While increases of a smaller magnitude (up to 7%) are seen in a few of the phosphate control experiments, these other cases appear due to sampling or analytic variability. An increase in concentration of this magnitude is not seen in any other incubations, and data for active human lung incubations and other active microsome incubations, including those with minimal metabolic activity, monotonically decline. Hence EPA concluded that an experimental artifact occurred with those particular control incubations and hence that the results should not be used in evaluating the background loss

different initial concentrations. For each incubation, samples of headspace air were removed from the vial by inserting a syringe through the septum covering the vial at each of five times (0.0, 0.2, 0.4, 0.6, 0.8, and 1.0 h) and withdrawing VINJ = 0.4 mL of air. This volume of air was injected into GC/MS system and the amount of chloroprene in the injection was assessed.

- For the Himmelstein et al. (2004) control experiments, the total volume of each vial was VVIAL = 12.0 mL, the volume removed at each sampling event was VINJ = 0.4 mL, and the volume of liquid media was V_med = 1.0 mL. As with the Yang et al. (2012) data, EPA assumed the total proportion of headspace chloroprene removed during each sampling event was equal to VINJ / (V_air + VINJ).
- EPA assumed that the concentration in the headspace volume V_air immediately following the removal of the sampling syringe was equal to the concentration in the sampling syringe, which was quantified and recorded by Himmelstein et al. (2004). Thus, EPA sought to calibrate the model by determining a value of the parameter RLOSS that yielded optimal agreement between the observed sample concentrations and the model predicted headspace concentrations immediately following the removal of the sampling syringe.
- With the likelihood function defined in Equation 4, EPA calibrated the model using a maximum likelihood estimation (MLE) methodology to obtain optimal values of RLOSS for each incubation. That is, EPA calculated values of RLOSS and A0 that cause the value of the likelihood function to be maximized (assuming a fixed value of 0.006 for σ based on the analysis of the Yang et al. (2012) data). The optimal model fit and the observed data for one of the Himmelstein et al. (2004) control incubation data series is shown in Figure 8.

rate. In addition, the individual points for each incubation of this set are difficult to discern and digitize from the figure, except for the phosphate buffer control. Therefore, EPA has chosen not to use those particular control experiments in its evaluation of the background rate of loss (RLOSS).

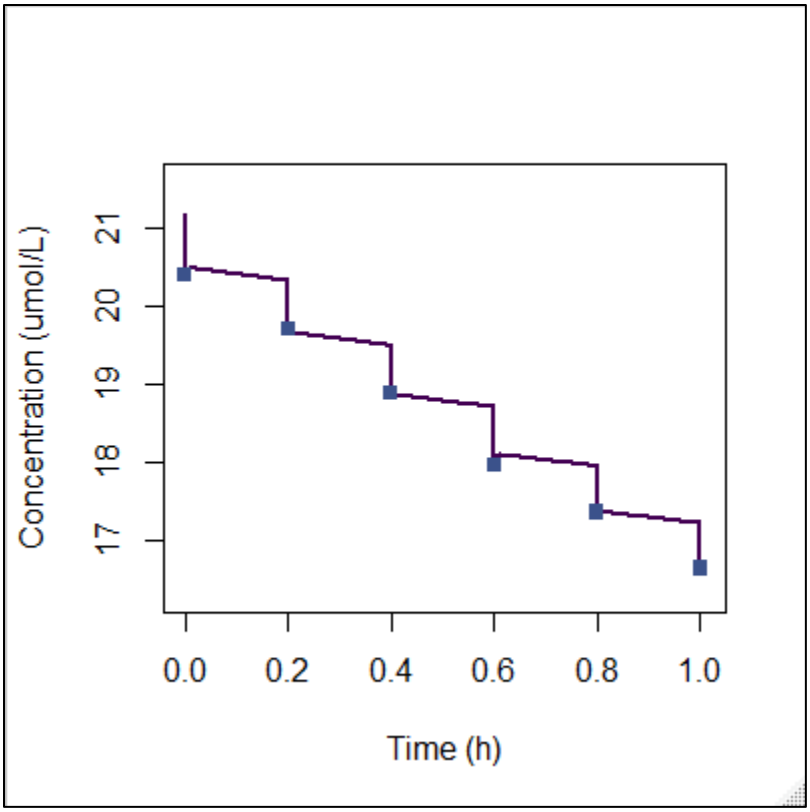

Figure 8. Optimal MLE model fit and the observed data for one of the Himmelstein et al. (2004) control incubation data series.

- Summary statistics for the five MLE estimates of RLOSS were:
  - mean = 0.0006659079 L/h
  - median = 0.0007406062 L/h
  - standard deviation = 0.0005938623 L/h
  - minimum = -0.00004382265 L/h
  - maximum = 0. 001235461 L/h
- ***Note that these summary statistics are considerably different from those calculated for the 30 MLE estimates of RLOSS using the*** Yang et al. (2012) ***control data; the standard deviation is comparable, but the values of the other statistics (e.g., the mean, median, and maximum) are about 50% less than those shown for the*** Yang et al. (2012) ***RLOSS results.*** This implies that the uncertainty and variability in the rates of background loss were comparable, but the magnitudes of these rates tended to be less for the Himmelstein et al. (2004) experiments than for the Yang et al. (2012) experiments.
- Thus, a question to be addressed is whether the analysis of human liver and lung metabolism data from Himmelstein et al. (2004) should be conducted using the distributional estimate of RLOSS based only on the Himmelstein et al. (2004) control data, for which there are five experiments, the distributional estimate of RLOSS based only on the {Yang, 2012, 3854472@@author-year control data, for which there are 30 experiments, or if the results for two sets of control experiments should be combined in a KDE.

- Also, it should be noted that while the minimum MLE estimate of RLOSS indicates that RLOSS could be negative, this implies a creation of chloroprene mass over the course of the experiment, which is physically impossible.
- EPA used the likelihood function and prior defined in Equations 4 and 5, respectively, and the {Himmelstein, 2004, 625152@@author-year} control data to generate samples from the Bayesian posterior distribution of $\theta = (\mathrm{A0}, \mathrm{RLOSS}, \sigma)$ for each of the 5 Himmelstein et al. (2004) control incubation data sets via Markov chain Monte Carlo (MCMC). Note that the approach described here does not involve a Bayesian hierarchical model but accounts for the fact that the experiments involve serial observations of each incubation vial.
- The distributional estimates for RLOSS for the five Himmelstein et al. (2004) control incubation data sets are illustrated as a series of box plots in Figure 9. In these box plots, the outer edges of each box show the 25th and 75th percentiles of the posterior distribution, the center line in each box shows the 50th percentile, and the outermost lines show the 2.5th and 97.5th percentiles.
- *Note: as with the results for Yang et al. (2012), the RLOSS chains for Himmelstein et al. (2004) had not converged at the time the following plots were generated, but the results are considered sufficient to illustrate key features of the data and apparent differences between the two studies.*

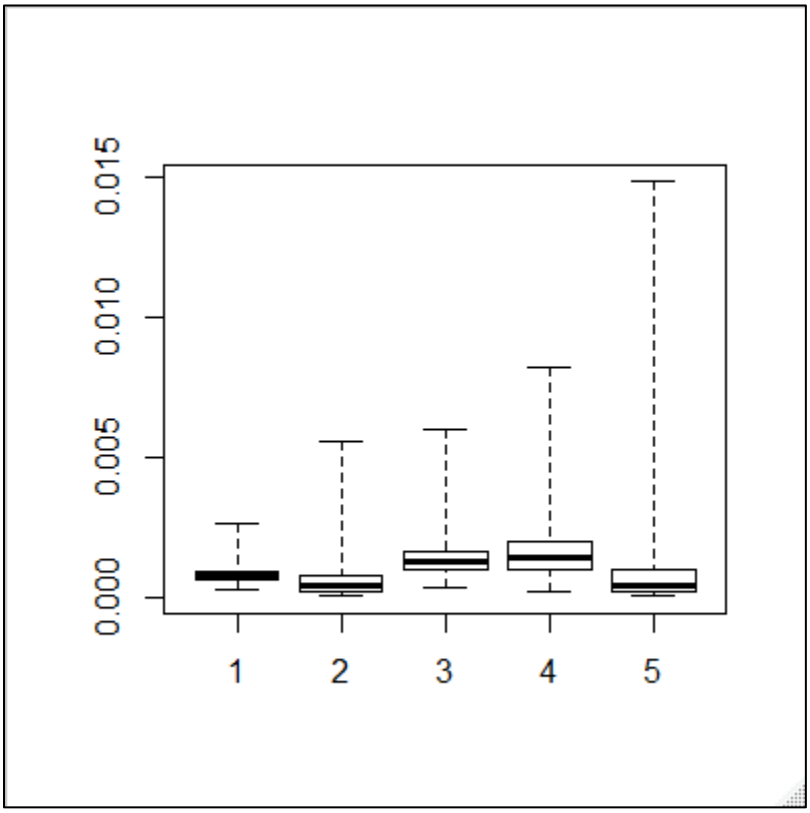

Figure 9. Box plots illustrating distributional estimates for RLOSS for the five Himmelstein et al. (2004) control incubation data sets.

- EPA applied KDE as implemented in the base R function "density" to generate probability density curves for RLOSS for the five Himmelstein et al. (2004) control incubation data sets. These are shown in Figure 10 as the fine blue (lighter-colored) curves.
- EPA also pooled all the samples for all five data sets and applied KDE to generate an "overall" probability density curve for RLOSS. This is shown in Figure 10 as the thicker purple (darker-colored) curve.

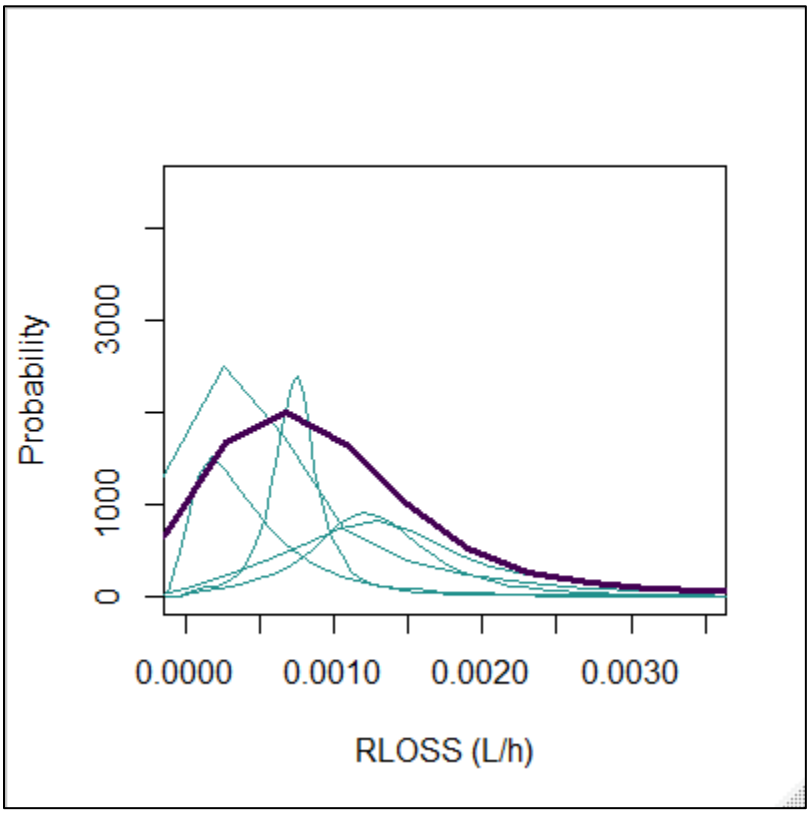

Figure 10. Probability density curves for RLOSS for the five Himmelstein et al. (2004) control incubation data sets.

(1b) Assessment of Uncertainty in Metabolic Parameters

Given the probability density curves for RLOSS described above, in particular those representing values of RLOSS that were obtained using the control data of Himmelstein et al. (2004), the uncertainty in the metabolic parameters obtained using active microsomal incubations in this study will be evaluated. The approach planned to evaluate this uncertainty is summarized to allow effective review of the intended methods for data analysis. Finalized methods will be developed based on review comments received.

These are key assumptions regarding the incubations.

a) KM is an intrinsic property of the enzymes involved and is independent of the amount of enzyme (or microsomes) in an incubation vial and therefore between incubation vials for the

same incubation type (species, sex, and tissue). Hence the KM for a given incubation type will be assumed to come from a common distribution across all such incubation vials.

b) Because enzyme activity for the human lung is low and the data do not show clear evidence of saturation/nonlinearity, it is not expected that independent values of KM can be identified for this tissue (unlike the mouse lung). Therefore, the first-order parameter, KF, will be estimated for the human, with VMAX = 0.

c) Since the exact amount of enzyme added to each incubation vial and the activity of enzyme in the sample may vary between vials, the VMAX or KF for a given incubation type is assumed to be different between vials, hence a different distribution for VMAX or KF will be fitted to each vial.

d) Because the exact amount of chloroprene added at the beginning of each incubation may vary, the estimate of the initial amount (A0) for each vial will be assumed to come from an independent distribution, though a common prior distribution will be used for each set of incubations with the same target initial concentration.

e) The uncertainty in VMAX and KM or KF for a given vial exists in addition to the measurement uncertainty for any individual data point.

f) The data for any incubation vial is expected to strongly indicate the total rate of loss in the vial, while the portion of that loss due metabolism vs. non-specific losses is expected to be more uncertain. In particular, the mean value of KF estimated by Yang et al. (2012) using the same human lung data was $9x10^{-4}$ L/h/mg microsomal protein. In an incubation with 1 mg protein, the effective rate is then $9x10^{-4}$ L/h, which is only slightly greater than the mode of the kernel density estimate shown in Figure 10.

Because of the issue noted in item (f) convergence of MCMC for a direct estimation of RLOSS and KF may be slow. An alternative analysis to aid with convergence of the MCMC algorithm will be evaluated, specifically to use the following change of variables:

$$TLOSS = KF + RLOSS. \quad (6)$$

Given the results in Figure 10 and that Yang et al. (2012) reported SD/mean = 0.8 for their estimate of KF, a value of 0.0035 L/h is 3.6 standard deviations above the previously estimated mean KF *and* is a reasonable upper bound on RLOSS. Therefore, an uninformative prior for TLOSS will be set as uniform distribution between 0 and 0.007 L/h (i.e., between zero and the sum of those upper bounds). This prior is based partly on the same data as will be used for the analysis, since no other data are available on which it might be based, but only to set an upper bound. This upper bound will be increased if initial MCMC chains are found to have significant density near the bound.

Since we require that KF be non-negative, the prior for TLOSS and RLOSS will be set such that the probability of RLOSS > TLOSS is zero (and the probability of TLOSS < 0 or RLOSS < 0 is zero).

Further details of the experimental data set and MCMC analysis are then as follows.

- Himmelstein et al. (2004) conducted in vitro experiments (as described previously in this document) with activated microsomes from human liver and lung tissues. The experiments used pooled samples of microsomes taken from 15 individuals in the case of liver microsomes and 5 individuals in the case of lung microsomes.
- EPA proposes to use information about uncertainty and variability in RLOSS (determined as described previously in this document) in order to support distributional estimates of the model parameters VMAX and KM or KF that quantify rates of metabolism of chloroprene in human liver and lung.

Estimation of In Vitro Kinetic Parameters for the Human Liver

- For the liver, the parameter TLOSS is defined as TLOSS = RLOSS + VMAX/KM, which is the approximate linear first-order total loss rate constant; i.e., as the concentration of chloroprene approaches zero.
- The data to be analyzed are fifteen non-control chloroprene oxidation incubations, using human liver microsomes, for which the data series are shown in Figure 1D of Himmelstein et al. (2004) ; for example, the data for the five incubation vials run using the middle concentration level with human liver microsomes are indicated with different symbols Figure 11.

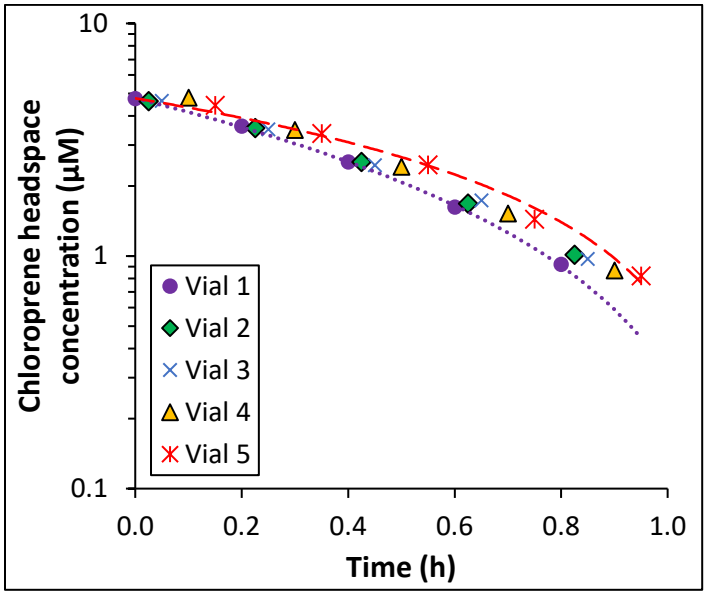

Figure 11. In vitro chloroprene incubation data using human microsomes, identified for each incubation vial. Data are from Figure 1D of Himmelstein et al. (2004). Simple trend-lines (dotted and dashed curves) are included for the first and last incubation vial to indicate the extent of variation among the samples.

- To estimate VMAX and KM for the human liver EPA will simultaneously generate distributional estimates (via MCMC) for the following parameters:
  - a common KM parameter, which is assumed to apply to all fifteen incubations;
  - fifteen TLOSS parameters, TLOSS_i for $i \in \{1, \ldots, 15\}$, one for each incubation vial;

  - fifteen RLOSS parameters, RLOSS_i for $i \in \{1, \ldots, 15\}$, one for each incubation;
  - fifteen A0 parameters, $\mathrm{A0}_i$ for $i \in \{1, \ldots, 15\}$, one for each incubation, for the amount of chloroprene in each incubation vial at the start of the experiment; and
  - fifteen observational noise parameters, σ_i for $i \in \{1, \ldots, 15\}$, one for each incubation.
  - Since TLOSS is defined as TLOSS = RLOSS + VMAX/KM, VMAX will be calculated as KM*(TLOSS – RLOSS) for use in the saturable metabolism equation for the liver. Model simulations to calculate the likelihood will then use the nonlinear saturable metabolic rate. Hence the combined rate of elimination will be slower than the first-order approximation, to the extent that saturation occurs.
  - The prior distribution for TLOSS in liver will be a uniform distribution from zero to $\mathrm{TLOSS}_{\max}$, which will be calculated as the sum of an upper bound for the RLOSS distribution (see below) and an upper bound for VMAX/KM obtained from a previous analysis, similar to the prior distribution for TLOSS in lung described above
  - Then, for the parameter vector $\theta = (\mathrm{KM}, \mathrm{A0}_1, \ldots, \mathrm{A0}_{15}, \mathrm{RLOSS}_1, \ldots, \mathrm{RLOSS}_{15}, \mathrm{TLOSS}_1, \ldots, \mathrm{TLOSS}_{15}, \sigma_1, \ldots, \sigma_{15})$, EPA will define a prior that ensures the components of θ are all non-negative, $\mathrm{TLOSS}_i \leq \mathrm{TLOSS}_{\max}$ and $\mathrm{RLOSS}_i \leq \mathrm{TLOSS}_i$ for $i \in \{1, \ldots, 15\}$.
  - The prior also assigns higher probability to smaller values of each σ_i using a Jeffreys prior as in Equation 5 and assigns relative probabilities to each RLOSS_i according to the combined distributional estimate of RLOSS shown in Figure 10.
  - The prior for KM will be based on the central estimate of this parameter provided in the Ramboll (2020) report with suitably wide bounds.
  - A0 for each incubation will be sampled from a uniform distribution selected to encompass the corresponding target initial concentration.
- These 61 distributional parameter estimates will be generated simultaneously, as opposed to separately (independently) because KM will be assumed to be the same for any given set of liver VMAX values and because the background loss rate that occurred in the active metabolism experiments may be somewhat different from the distribution estimated from the control experiments. Note that the approach described here does not involve a Bayesian hierarchical model but accounts for the fact that the experiments involve serial observations of each incubation vial.

Estimation of In Vitro Kinetic Parameters for the Human Lung

- A similar approach to the analysis for human liver described above will be used with data for five chloroprene oxidative metabolism experiments shown in Figure 2D of Himmelstein et al. (2004), conducted with human lung microsomes:
  - five TLOSS parameters, one for each of the five human-lung microsome incubations;
  - five RLOSS parameters, one for each lung microsome incubation;
  - five A0 parameters, one for each incubation; and
  - five σ parameters, one for each incubation
  - KF will be calculated as TLOSS – RLOSS for use in the in vitro kinetic model.
  - For each incubation, EPA will estimate a parameter vector $\theta = (\mathrm{A0}, \mathrm{RLOSS}, \mathrm{TLOSS}, \sigma)$ using Bayesian parameter estimation. To do this, EPA will define a prior that ensures the components of θ are all non-negative, $\mathrm{TLOSS} \leq 0.007\ L/h$, and $\mathrm{RLOSS} \leq \mathrm{TLOSS}$. (The value of 0.007 L/h was selected specifically for the human lung as described previously.)
  - A0 for each incubation will be sampled from a uniform distribution selected to encompass the corresponding target initial concentration.

    - The prior also assigns higher probability to smaller values of σ using a Jeffreys prior as in Equation 5 and assigns relative probabilities to RLOSS according to the combined distributional estimate of RLOSS shown in Figure 10.
- These 20 distributional parameter estimates can be generated as five separate sets of four (TLOSS, RLOSS, A0, and σ), because the total loss (TLOSS) and background loss rate (RLOSS) that occurred in the active metabolism experiments may be somewhat different from the distribution estimated from the control experiments. Again, the approach described here does not involve a Bayesian hierarchical model but accounts for the fact that the experiments involve serial observations of each incubation vial.

The prior for the RLOSS parameter will be based on distributional estimates of RLOSS generated (as described previously in this document) from the Himmelstein et al. (2004) control data and/or the Yang et al. (2012) control data. For example, we may collect all 35 distributional estimates of RLOSS, which are samples of the 35 posterior distributions generated via MCMC, and combine them (e.g., via KDE), to obtain a single prior distribution for RLOSS. However, preliminary results shown in Figures 7 and 10 above indicate that the loss rate distribution that occurred during the experiments of Himmelstein et al. (2004) was somewhat lower than that which occurred during the experiments of Yang et al. (2012). The incubation vials used in the two sets of experiments were slightly different in volume, indicating a difference in the manufacturing process, which could lead to differences in losses through and around the septum. And based on the previous analyses, use of the non-concurrent control data from Yang et al. (2012) could then lead to a significant under-estimation of human lung metabolism (relative to the total activity observed in these experiments). Therefore, use of only the Himmelstein et al. (2004) control data to define the prior distribution for RLOSS seems preferable.

(2) Assessment of Uncertainty in PBPK Model Prediction of Metabolic Rates and Venous Blood Concentrations

This is an overview of quantitative uncertainty analysis that the U.S. EPA is proposing for chloroprene PBPK model predictions of oxidative chloroprene metabolism in human lung, liver, and kidney. In responding to the corresponding charge questions, EPA anticipates that reviewers will evaluate both the specific modeling assumptions and techniques indicated here as well as the extent to which this analysis (or alternatives that might be recommended) would be expected to accurately indicate bottom line uncertainty in the predicted metabolism of chloroprene from actual human exposures.

The chloroprene PBPK model uses a typical structure for such models and the values of the physiological parameters represent average adult mice, rats, and humans and are attributed to appropriate sources. Also, the objective of cancer risk assessment is to identify the average cancer risk in the population. Therefore, the focus of the uncertainty analysis described here is on the uncertainty in the IVIVE calculations for humans, rather than estimating inter-individual variability that derives from variation in all physiological parameters among the population. Hence, only the impact of uncertainty in the physiological parameters directly involved in the IVIVE calculations will be evaluated. The parameters involved in the IVIVE calculations and important to model predictions are:

- The in vitro estimates of Vmax for liver and KF for lung, which are rates per mass (mg) of microsomal protein.
  - Alternately, the parameter A1 for lung-to-liver relative activity instead of the in vitro Vmax for the lung.
  - An additional parameter, A2, for kidney-to-liver relative activity will be evaluated to estimate metabolic activity in that tissue.
- The concentration of microsomal protein, CMP (mg protein/kg tissue), in liver and lung.
- The fractions of total body weight (BW) which are liver and lung (VLIC and VLUC), and possibly kidney (VKC).

In addition to uncertainty contributed by these identified parameters, an otherwise hidden uncertainty occurs because the IVIVE calculation assumes that the rate of oxidative metabolism per mg microsomal protein observed in vitro equals the rate of oxidative metabolism per mg protein in the endoplasmic reticulum that occurs in vivo. Additionally, only single pooled samples of tissue/species specific microsomes were utilized. For example, single samples of commercially available pooled human microsomes were procured. The degree to which these samples represent population average metabolic rates is an additional uncertainty. The U.S. EPA has not yet identified appropriate information sources to quantify this uncertainty.

Uncertainty in the PBPK model predictions for a given exposure scenario will be quantified by conducting model simulations using Monte Carlo (MC) sampling of each of the aforementioned metabolism parameters, treating them as independent random variables. To be more specific, we will sample KM from the posterior distribution for that parameter generated as described in Section 2b; we will sample liver VMAX from the combined pool of samples of the 15 posterior distributions for this parameter that were generated via MCMC as described in Section 2b; and we will sample lung KF from the combined pool of samples of the 5 posterior distributions for this parameter that were generated via MCMC as described in Section 2b. A distribution representing the uncertainty in the microsomal content will be obtained by surveying the literature identified in the Ramboll (2020) report to determine the range of values reported among various studies for the human liver. Uncertainty in the estimated microsomal content for the human lung may also depend on the ratio of lung-to-liver content obtained from measurements in rodents. To the extent possible, uncertainty in the tissue fractions of body weight will be based on uncertainty in the population mean, but it may not be possible to completely separate uncertainty in the mean from inter-individual variability among subjects from whom empirical data were collected.

Using a fixed exposure scenario, a sufficient number of MC samples (10–20k expected) will be generated to obtain a good estimate of the 1st, 5th, 95th, and 99th percentiles of the resulting distribution for the rate of production of oxidative metabolites in liver and lung (and kidney) and the average venous blood concentration. The width of these distributions will be presumed to represent in the level of uncertainty in the PBPK model predictions associated with the IVIVE calculations.

References:

Himmelstein, MW; Carpenter, SC; Hinderliter, PM. (2004). Kinetic modeling of beta-chloroprene metabolism: I. In vitro rates in liver and lung tissue fractions from mice, rats, hamsters, and humans. Toxicol Sci 79: 18-27. http://dx.doi.org/10.1093/toxsci/kfh092

Himmelstein, MW; Carpenter, SC; Hinderliter, PM; Snow, TA; Valentine, R. (2001). The metabolism of beta-chloroprene: preliminary in-vitro studies using liver microsomes. Chem Biol Interact 135-136: 267-284. http://dx.doi.org/10.1016/S0009-2797(01)00214-9

Kreuzer, PE; Kessler, W; Welter, HF; Baur, C; Filser, JG. (1991). Enzyme specific kinetics of 1,2-epoxy-3-butene in microsomes and cytosol from livers of mouse, rat, and man. Arch Toxicol 65: 59-67. http://dx.doi.org/10.1007/BF01973504

Ramboll. (2020). Incorporation of in vitro metabolism data in a physiologically based pharmacokinetic (PBPK) model for chloroprene.

Yang, Y; Himmelstein, MW; Clewell, HJ. (2012). Kinetic modeling of β-chloroprene metabolism: Probabilistic in vitro-in vivo extrapolation of metabolism in the lung, liver and kidneys of mice, rats and humans. Toxicol In Vitro 26: 1047-1055. http://dx.doi.org/10.1016/j.tiv.2012.04.004