Journal of Exposure Analysis and Environmental Epidemiology (2004) 14, 204–213

www.nature.com/jea

Reconstructing population exposures from dose biomarkers: inhalation of trichloroethylene (TCE) as a case study

MICHAEL D. SOHN,[a] THOMAS E. MCKONE[a,b] AND JERRY N. BLANCATO[c]

[a]*Lawrence Berkeley National Laboratory, Berkeley, California 94720, USA*
[b]*University of California, Berkeley, California 94720, USA*
[c]*National Exposure Research Laboratory, United State Environmental Protection Agency, Las Vegas, Nevada 89193, USA*

Physiologically based pharmacokinetic (PBPK) modeling is a well-established toxicological tool designed to relate exposure to a target tissue dose. The emergence of federal and state programs for environmental health tracking and the availability of exposure monitoring through biomarkers creates the opportunity to apply PBPK models to estimate exposures to environmental contaminants from urine, blood, and tissue samples. However, reconstructing exposures for large populations is complicated by often having too few biomarker samples, large uncertainties about exposures, and large interindividual variability. In this paper, we use an illustrative case study to identify some of these difficulties, and for a process for confronting them by reconstructing population-scale exposures using Bayesian inference. The application consists of interpreting biomarker data from eight adult males with controlled exposures to trichloroethylene (TCE) as if the biomarkers were random samples from a large population with unknown exposure conditions. The TCE concentrations in blood from the individuals fell into two distinctly different groups even though the individuals were simultaneously in a single exposure chamber. We successfully reconstructed the exposure scenarios for both subgroups — although the reconstruction of one subgroup is different than what is believed to be the true experimental conditions. We were however unable to predict with high certainty the concentration of TCE in air.
Journal of Exposure Analysis and Environmental Epidemiology (2004) **14,** 204–213. doi:10.1038/sj.jea.7500314

Keywords: *exposure assessment, biomarkers, PBPK modeling, Bayesian inference.*

Introduction

Physiologically based pharmacokinetic (PBPK) modeling is a well-established toxicological tool designed to transform an exposure into a target tissue dose (Ramsey and Andersen, 1984; Klaassen, 1996). In reviewing the published literature in journals such as *Toxicology and Applied Pharmacology, Risk Analysis, Environmental Health Perspectives, and the Journal of Exposure Analysis and Environmental Epidemiology,* we observe that PBPK modeling is poised to move beyond the phase of model development to the phase of model application. The motivation for such a move is driven in large part by the emergence of federal and state programs for environmental health tracking. There is certainly consensus in the community on how to build PBPK models — basing them on known physiological processes that are easily generalized (blood flow rates, tissues volumes, breathing rates, etc.), on chemical-specific processes that are predicted from existing data and regression models (partition coefficients, chemical density, molecular weight, etc.), and on processes such as metabolism that are highly variable among species and individuals (Gargas et al., 1989). The US EPA National Exposure Research Laboratory in Las Vegas has had a generalized PBPK model called the Exposure-Related Dose Estimating Model (ERDEM) for some 5 years.

In spite of the consensus on how to build PBPK models, there is much less consensus or discussion on how to use PBPK models to find environmental determinants of chronic disease from biomonitoring — such as measured pollutant levels in blood and urine samples for a cross-section of the population. Yet such applications of PBPK models are needed to build hypotheses about possible relationships between exposures, dose, and disease; to monitor trends in environmental quality and disease; and to provide public health professionals with reliable information for early detection and prevention of diseases. Rather than new models, such issues demand a new framework for applying PBPK models.

1. Address all correspondence to: Dr. Michael D. Sohn, Lawrence Berkeley National Laboratory, One Cyclotron Road, MS: 90R3058; Berkeley, CA 94720, USA. Tel.: +1-510-486-7610. Fax: +1-510-486-6658. E-mail: mdsohn@lbl.gov

The United States Environmental Protection Agency through its Office of Research and Development funded and collaborated in the research described here under Interagency Assistance Agreement to Lawrence Berkeley National Laboratory (IAG #89938190). It has been subjected to Agency review and approved for publication.
Received 13 January 2003; accepted 13 August 2003

One application for such a new model framework is improving the process of relating the production, use, or release of a chemical to corresponding doses in an exposed population. Most existing PBPK applications were applied to well-understood or characterized studies of an individual or a small cohort. The new challenge is how to apply PBPK models to larger and more poorly characterized human populations that have highly variable exposures, activities, physiology, and pharmacokinetics (Bois, 2001). An important research question here is whether PBPK models are broadly applicable as tools for relating dose biomarkers to measures of population exposure and health risk. If feasible such an application offers the opportunity to better relate biomarkers to specific sources of exposure, for example, household pesticide use *versus* food residues, and volatile organic compound (VOC) emissions from consumer products *versus* those from automobiles or from stationary sources.

Although limited to date, there have been recent advancements on this front. The EPA dioxin reassessment used PBPK models to evaluate the reasonableness of their earlier estimated cumulative dietary intake of dioxin compounds (Pinsky and Lorber, 1998; USEPA, 2001). Wallace and Pellezzari (1995) and Wallace (1997) assessed the utility of using exhaled breath for estimating exposure and body burden for VOC based on PBPK models. Chinnery and Gleason (1993) and McKone (1993) used PBPK models of chloroform applied to breath samples reported by Jo et al. (1990) to determine the relative contribution of inhalation and dermal exposure routes for adults showering with water containing residual chloroform from disinfection. By developing methods to treat PBPK model parameters as random variables within the constraints of empirically observed distributions, Bois et al. (1996a, b) and Gelman et al. (1996) illustrated the feasibility of using population-based models for interpreting exposures to tetrachloroethylene and benzene.

Resources and opportunities to produce population-scale source-to-dose reconstructions may come from new and ongoing national and regional health surveys. The Center for Disease Control (CDC) National Health and Nutrition Examination Survey (NHANES) (CDC, 2001) employs a home interview with health tests to collect information about the health and diet of people in the United States. It includes data on blood levels of cadmium, lead, mercury, pesticides, and combustion products. Through the National Human Exposure Assessment Survey (NHEXAS) (Sexton et al., 1995), the Children's Total Exposure to Persistent Pesticides and other Persistent Organic Pollutants (CTEPP) (USEPA, 2002), and other programs, the US EPA is developing databases on exposures of human populations to a wide range of pollutants in air, water, food, soil, and indoor/residential environments, and over a wide range of space and time scales. The University of California, Berkeley Center for the Health Assessment of Mothers and Children of Salinas (CHAMACOS) is collecting biomarkers in farming communities for pesticides and other important pollutants from mothers and their newborn children from conception through early childhood (Castorina et al., 2002).

However, biomarkers obtained from these surveys are inherently variable as a result of the inter- and intra-individual variability in both the exposures to the population and the physiology of the individuals in the population (Bois, 2001). The key questions are whether and how well we can quantify the variation in source-to-dose relationships against the noise contributed by these other variable and uncertain factors. Examining the input and output information obtained in these surveys will be critical. What exposure information is currently available, and what additional information is likely or practical to be obtained? How do we make PBPK models compatible with the available information instead of using the models in the form originally developed — based on controlled laboratory conditions? Answering these questions is an important direction for ongoing development and application of PBPK models. It will require more use and development of statistical and other quantitative methods for integrating uncertainties and variability in both model predictions and biomarker data.

In this paper, we use an illustrative example to highlight some of the challenges facing exposure analysts and lay out an approach for reconstructing population-scale exposures using commonly known Bayesian inference techniques. A Bayesian approach provides flexibility for evaluating multiple exposure scenarios and alternative data sets, which will be critical for systematically reconstructing exposures using biomarkers collected from a large and diverse population. The approach can also help practitioners apply PBPK models, set priorities for exposure and health monitoring programs, and decide what exposure information to gather in the near term.

In the illustrative application, we treat data gathered from a controlled experiment as if they were samples from individuals in a large population with unknown exposure conditions. We reconstruct the exposure conditions as if we had limited information about the individuals in the population.

We also discuss methods to determine what exposure assessment information is important in the dose reconstructions, what information other than that commonly gathered in exposure assessments could improve the dose reconstruction, and how quality and quantity of data affects the reconstructions, thereby assisting future exposure and epidemiological studies. We do not demonstrate the feasibility of using PBPK models for dose reconstruction. This has been done, so our goal is to study some of the limitations and capabilities of using PBPK models and Bayesian inference for difficult cases, such as the volatile

chemical trichloroethylene (TCE), which has a short residence time within the body and multiple sources of exposure.

Exposure classification using Bayesian inference

In the most general sense, an exposure assessment involves quantifying a link between a source of contamination, its transport and transformation among a set of environmental media, human contact with exposure media, and the route of application or entry (USEPA, 1989; McKone and Daniels, 1991; USEPA, 1992; Zartarian et al., 1997). Each component in the source-to-dose link includes some level of uncertainty or bias. For example, conceptual and process models in PBPK models are developed from infrequently sampled yet highly variable and uncertain data. Ignoring the variability and uncertainty in the models can imply overconfidence in the PBPK model, and can cause erroneous estimates of exposure-to-dose relationships.

We confront these uncertainties using Bayesian inference methods - methods increasingly applied in environmental and health analyses. Brand and Small (1995), Bois et al. (1996a), and Sohn et al. (2000) have described these methods in detail. Relevant applications in multi-pathway, multi-parameter, environmental systems include assessing environmental health risk (Taylor et al., 1993; Spear and Bois, 1994; Brand and Small, 1995; Wakefield, 1996; Gelman et al., 1996; Bois et al., 1996a; Pinsky and Lorber, 1998; Vicinni et al., 1999) and conducting environmental value-of-information analyses (Finkel and Evans, 1987; Dakins et al., 1996). Because of the wide availability of publications on Bayesian inference, we only describe here briefly the relevant details of this technique as it pertains to population-scale source-to-dose analyses.

As a starting point for Bayesian inference, the practitioner develops mechanistic, statistical, and/or empirical models that predict the source-to-dose relationship. Any unknown, uncertain, or variable model input is probabilistically described using parametric or nonparametric uncertainty distributions. Examples of unknowns include the alternative exposure scenarios, variability in the pharmacokinetics, alternative conceptual models such as two-compartment or five-compartment PBPK models, first-order or second-order environmental degradation, well-mixed or multicompartment indoor air models, and various model parameter uncertainties. Field data, epidemiological studies, best engineering judgment, and any quantitative or subjective information are possible sources for developing value ranges of the probabilistic distributions. Generally, the practitioner will assign wide uncertainty distributions due to the limited information.

The practitioner next predicts model end points that can be compared to specific biomarker data. A Monte Carlo or Latin Hypercube sampling technique may be applied to generate a library consisting of several thousand realizations of exposure scenarios and biomarker predictions. The values and the uncertainties of model predictions, model input parameters, and conceptual models employed are often referred to as the ''prior'' since they collectively define the model formulation prior to being compared to data.

The practitioner next assesses the agreement between biomarker data and each model prediction in the library of model simulations using Bayes' rule.

$$p(Y_k|O) = \frac{L(O|Y_k)p(Y_k)}{\sum_{i=1}^{K} L(O|Y_i)p(Y_i)} \quad (1)$$

where $p(Y_k|O)$ is the probability of the kth Monte Carlo simulation making prediction Y_k given the biomarker data O, $L(O|Y_k)$ is the likelihood of observing measurements O given model prediction Y_k, $p(Y_k)$ is the prior probability of the kth Monte Carlo simulation and K is the number of Monte Carlo simulations. Before data comparison, each of the model realizations is equally likely (i.e., $p(Y_k) = 1/K$.

The probability $p(Y_k|O)$ is often referred to as the posterior probability of the kth realization since it describes the probability after the kth realization is compared to data. In this case, the posterior probability describes the degree that the kth dose prediction — and the associated model input parameters, conceptual models, and exposure scenario used to generate that prediction — accurately describes the observed biomarker concentrations. The posterior probability thus replaces all of the uncertain prior probabilities (e.g., unknown model input parameters and exposure scenarios) and model predictions (e.g., dose estimates). Brand and Small (1995) and Sohn et al. (2000) provide the relevant equations for estimating posterior means, variances, and correlation coefficients.

The likelihood function, $L(O|Y_k)$, in Eq. (1) quantifies the error structure of the data, that is, the differences between the data and the model predictions resulting from measurement error, spatial and temporal averaging or correlations, and imperfect model representation. If many independent measurements are considered, for example following random samples in a large epidemiological survey, the likelihood of observing all of the measurements is the product of all of the individual likelihoods:

$$L(O|Y_k) = \prod_{s=1}^{S} L(O_s|Y_{s,k}) \quad (2)$$

where S is the number of independent measurements.

For independent unbiased measurements, a Gaussian likelihood function is often assumed (Taylor et al., 1993; Brand and Small, 1995; Dakins et al., 1996; and Sohn et al., 2000). It is important for the practitioner to estimate properly the error variance term(s) in the developed or assumed likelihood function. Failure to do so can lead to overly confident posterior probabilities, and can cause erroneous estimates of exposure-to-dose relationships (Sohn et al., 2000; Small and Fischbeck, 1999).

We further note that the Bayesian updating procedure presented here is executed in several sequential steps including: (1) develop models and predict dose, (2) compare predictions to data, and (3) update uncertainties using the likelihoods. Alternatively, one could predict model end points and compare them to biomarkers using one of several other variations of Bayesian updating such as Markov Chain Monte Carlo or Gibbs sampling (Gelman et al., 1995, 1996; Bois et al., 1996a; Roy and Georgopoulos, 1998). These approaches apply Eq. (1) and (2) iteratively and are capable of quickly searching the model parameter uncertainties. However, we did not find these more complex methods necessary for the illustrative demonstration that follows. Since the operation of the PBPK models was very fast on a standard desktop personal computer, we did not have difficulty in adequately sampling the uncertainties and variability in the model inputs using standard approaches.

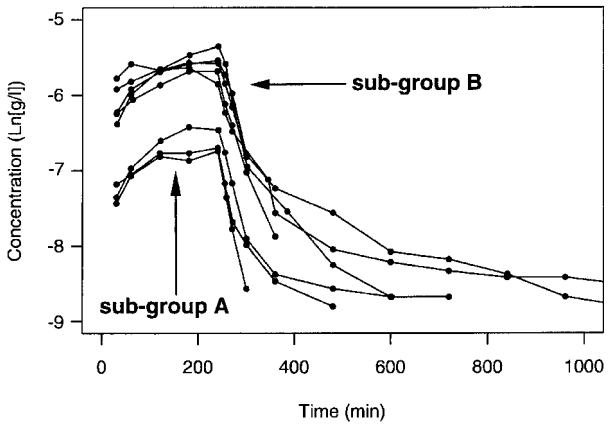

Figure 1. TCE concentrations in venous blood. Eight adult males were exposed in a chamber to TCE in air at a concentration of 100 ppm for 240 min. The TCE concentration in air was zero thereafter ($t > 240$ min).

Application to a TCE-exposed cohort study

We analyze biomarker concentrations in samples gathered from a controlled experiment as if they were samples from individuals in a large population with unknown exposure conditions. It is important to note that the purpose of the application is to evaluate the feasibility of reconstructing exposures to populations and highlight some of the difficulties facing the exposure analyst in this process. They are not to demonstrate that we are able to reconstruct the exposures to these specific individuals. Those types of analyses have been demonstrated by Gelman et al. (1996), Vicini et al. (1999) and others. For our purposes, we recognize that the variability and uncertainties in a large health tracking study will be larger than those found in this controlled study. Nevertheless, the errors and variability found in the data are sufficient to demonstrate the approach and highlight the difficulties that arise in reconstructing exposures to populations.

The biomarker data come from experiments conducted by Fisher and colleagues (Fisher, 1998). Eight adult males were exposed to TCE vapors in air for 240 min in an enclosed chamber. The details of the experiment and the laboratory setup are described in Fisher (1998). Figure 1 provides plots of the times-series of TCE concentrations in the venous blood of the eight exposed subjects. All eight individuals were contained in the same chamber, although we note that three individuals, who we call subgroup A, have significantly lower concentrations than the other five individuals, who we call subgroup B. There are no obvious experimental or equipment variations that can explain the differences in the blood concentrations.

We developed a PBPK model to predict the concentration of TCE in venous blood from exposure by inhalation. Ramsey and Andersen (1984) provide details of a typical first-order multicompartment PBPK model. Researchers at the US EPA successfully developed a multicompartment PBPK model using the ERDEM model (formerly referred to as DEEM) for an individual exposed to TCE by inhalation (Blancato et al., 2002). We could have developed a simple PBPK model in ERDEM. However since ERDM requires information and detail not usually available in population-scale exposure studies, we used a simpler PC-based PBPK model for the Bayesian-inference process. Our model is a first-order five-compartment model consisting of four well-mixed tissue groups — fat, liver, slowly and rapidly perfused tissue — and a pulmonary compartment to represent the blood/air transfer in the lungs. Metabolism in the liver is described by Michaelis–Menton kinetics. The model is coded in FORTRAN and solved using the VODE solver (Brown et al., 1989).

Table 1 summarizes the parameters needed to characterize the uncertainty and variability in the exposure to TCE and the pharmacokinetics of this population. These parameters include wide value ranges for both exposure-scenario descriptors and human pharmacokinetic parameters. Much of this information was obtained from the original study. However, we intentionally excluded information, such as individual body weight and exposure duration, which was collected from these individuals but would not be available in actual population-scale exposure analyses. We predicted TCE concentration in blood by sampling the probability distributions in Table 1 using Latin Hypercube sampling and then simulating various exposure scenarios using the five-compartment PBPK model. A total of 20,000 samples and model simulations sufficiently sampled the parameter value ranges summarized in Table 1. Figure 2 shows the resulting range of predicted TCE concentrations in blood. The wide uncertainty bounds in Figure 2 result from the combined uncertainties in the parameters describing the exposure scenario and pharmacokinetics.

Table 1 Exposure and pharmacokinetic uncertainty or variability

Model parameter	Range	Distribution
Exposure		
TCE Conc. in air (ppm)	25–300	Uniform
Exposure duration (h)	1–6	Uniform
Onset of exposure (h)	−3–3	Uniform
Metabolism		
V_{max} (g/s)	GM: 2.84e-5, GSD: 2.38	Lognormal
K_m (g/l)	GM: 8e-4, GSD: 2.8	Lognormal
Pulmonary flows		
Q_{air} (l/s)	GM: 0.108, GSD: 2.16	Lognormal
Q_{fat} (l/min)	GM: 0.3, GSD: 2.3	Lognormal
Q slowly perfuse tissue (l/min)	GM: 0.81, GSD: 2.22	Lognormal
Q rapidly perfuse tissue (l/min)	GM: 3.98, GSD: 2.22	Lognormal
Q liver (l/min)	GM: 1.12, GSD: 2.3	Lognormal
Volume		
V fat (1)	GM: 12.87, GSD: 2.1	Lognormal
V slowly perfuse tissue (1)	GM: 42.3, GSD: 2.06	Lognormal
V rapidly perfuse tissue (1)	GM: 12.94, GSD: 2.18	Lognormal
V liver (1)	GM: 2.18, GSD: 2.08	Lognormal
Partition coefficient		
P blood/air	GM: 18, GSD: 2.18	Lognormal
P fat	GM: 50.9, GSD: 2.3	Lognormal
P slowly perfuse tissue	GM: 1.5, GSD: 2.28	Lognormal
P rapidly perfuse tissue	GM: 3.67, GSD: 2.22	Lognormal
P liver	GM: 5.81, GSD: 2.3	Lognormal

GM and GSD are abbreviations for geometric mean and geometric standard deviation, respectively.

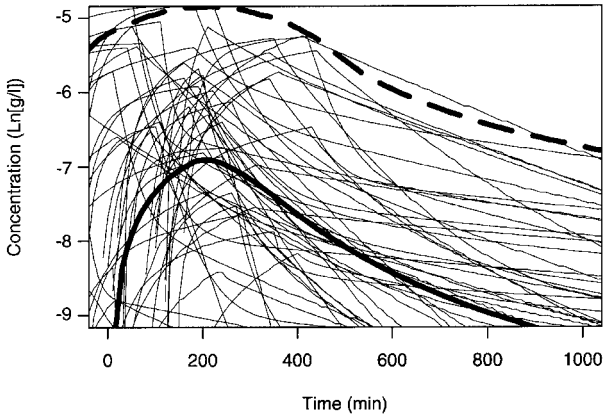

Figure 2. Predictions of TCE concentration in venous blood before comparison to biomarker data. Each thin line represents a concentration profile predicted from a sample of the exposure and pharmacokinetic unknowns in Table 1. Of the 20,000 simulations, 50 are plotted here. The thick line is the median of the 20,000 simulations and the thick dotted line is the upper bound of the two-sided 95% confidence interval. The lower bound is below the limits of the *y*-axis, hence is not visible in this figure.

We next applied Bayes' rule to establish posterior parameter ranges for all parameters in Table 1. We focus here on the posterior range of the three exposure parameters in Table 1, since we use these parameter to reconstruct the exposure to the population. The posterior estimates of the kinetic parameters (metabolism, flow, volume, and partition parameters) did not have the large posterior ranges seen for the exposure parameters so, for brevity, they are not considered further. In actual exposure surveys, it is rare to have both temporal exposure data for each individual and cross-sectional exposure data within the population. Therefore, we elected not to aggregate the individual profiles to estimate a population-equivalent profile, but to retain a profile for each individual and consider posterior parameter value ranges among individuals. We also chose not to use a hierarchical Bayesian approach (see e.g., Gelman et al., 1996) because it was not needed for the purposes of this paper.

In reconstructing the exposure to the population based on Eq. (1), we assumed a lognormal population error structure and assumed that the errors were uncorrelated. We estimated the error variance term in the lognormal distribution empirically by pooling the log-transformed data at time intervals, estimating the mean and errors, and pooling the errors to estimate the variance (Weisberg, 1985, p. 90), which we estimated as 0.6 [ln(g/l)]. Alternative methods for estimating the errors or alternative population error structures could be applied to this data set. However, more sophisticated methods are unlikely to be possible in actual analyses given the infrequent and limited data typically available.

Figure 3 shows the reconstructed exposure scenarios and Figure 4 shows the reconstructed TCE concentration in the blood. The exposure reconstruction for subgroup A is consistent with the experimental conditions. The uncertainty bounds narrow to the values reported by Fisher (1998) for the duration of the exposure (Figure 3), the onset of the exposure (Figure 3), and the predicted concentration in blood (Figure 4). Although the median TCE concentration is close to the reported air concentration, the predicted range of the TCE concentration in air (Figure 3) still contains considerable uncertainty. This is despite the high temporal resolution of the biomarkers.

In contrast to subgroup A, the exposure reconstructions for subgroup B are not consistent with the values reported by Fisher (1998). The Bayesian updating reduces uncertainty, but the predicted exposure duration and onset are longer than what is reported to have occurred in the experiment.

The differences in reconstructions for the two data sets are in large part due to the unexplained differences in the TCE concentrations. Our discussions with the authors of the experimental study did not yield an obvious explanation. One explanation may be that eight individuals were too few individuals to describe the pharmacokinetic behavior of a population. A second explanation, although highly speculative, may be that the experimental chamber was not uniformly well mixed. A poorly mixed chamber could

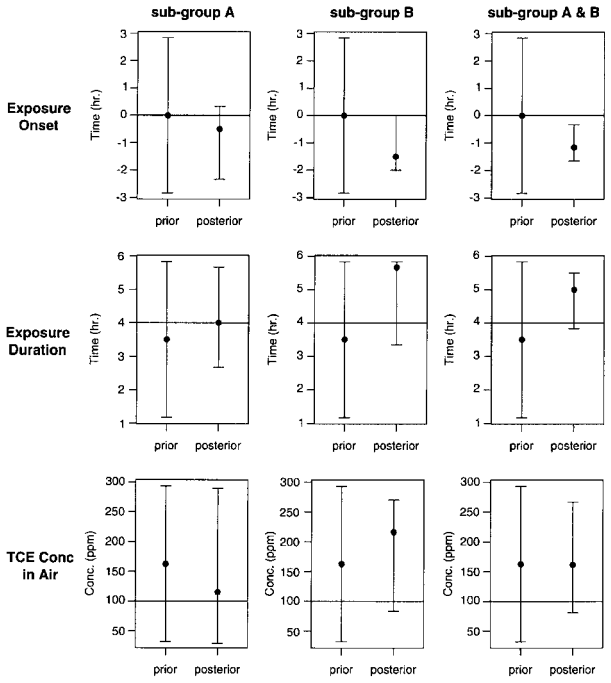

Figure 3. Reconstructed exposure and concentration of TCE in air for the data sets identified in Figure 1. The prior and posterior bars present the uncertainty before and after comparison to biomarker data, respectively. The whiskers of the bars are the two-sided 95% confidence interval and the circle is the median. The horizontal line represents the value reported by Fisher (1998).

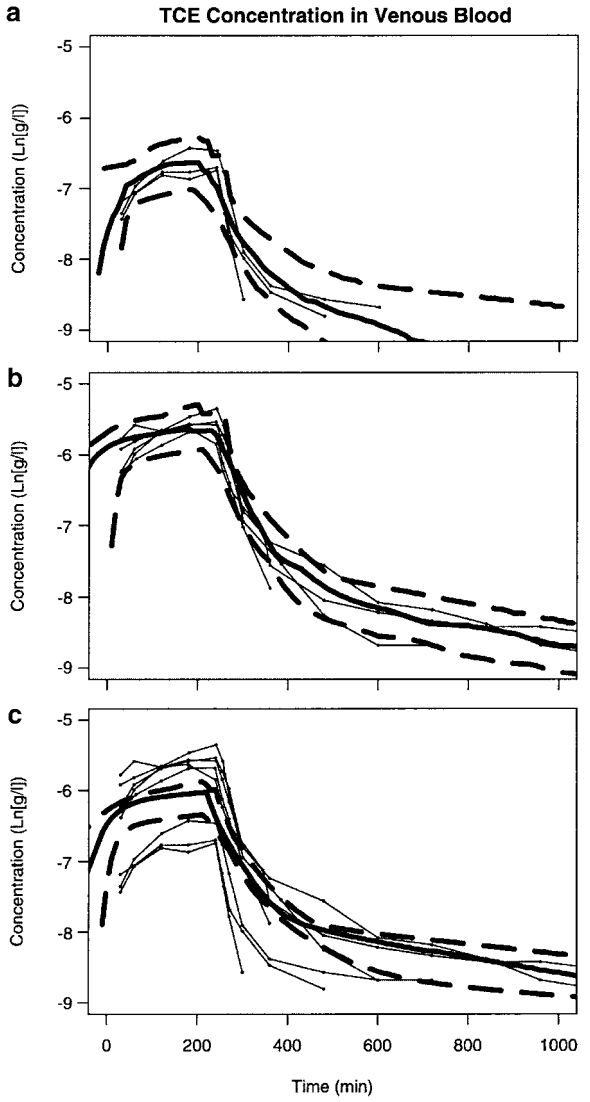

Figure 4. Updated predictions of TCE in venous blood using biomarker data from (a) subgroup A, (b) subgroup B, and (c) subgroups A and B combined (see Figure 1). The thick dotted lines define the two-sided 95% confidence interval. The thick solid line is the median.

produce turbulent eddies of high concentration that some of the individuals (i.e., those defined in subgroup B) contacted. A third explanation may be that the two subgroups reflect different metabolism profiles or breathing rates among the population of possible subjects. However, this does not explain the presence of two unique groups and not a simple continuum. If metabolism or breathing rates are random effects, and thus probabilistically distributed over some quantifiable range, we would expect the blood concentrations to be somewhat evenly spread throughout the range of concentrations. Instead, we find two unique groups. Finally, another explanation may be the presence of a genetic polymorphism among subgroup B that results in the presence or absence of a metabolism pathway and the resulting bimodal distribution of exposure to biomarker ratios.

In the final stage of the analysis, we reconstructed the exposure of the whole population. Figures 3 and 4 show the reconstructed exposure conditions and the predicted TCE concentrations in blood. Based on the reported conditions of the experiment, the predicted TCE concentration in blood (Figure 4c) appears to have erroneous uncertainty bars. They are so narrow that a large percentage of the data fall outside of the two-sided 95% confidence interval. If the biomarker data represented a large population, then the empirically derived error variance term is too narrow. If, however, the population is not a single-mode population, but a two-mode (or multiple subgroup) population, then the assumed likelihood function is incorrect. We show this result to emphasize the importance of applying correct modeling assumptions when developing source-to-dose links, irrespective of the exposure reconstruction method employed.

In actual health-tracking, we cannot expect to have sufficient data to dissaggregate the population into subgroups, or to even realize that subgroups exist. We may have more sampled individuals, but it is unlikely that we would have such high temporal resolution for very many individuals. In Figure 5, we construct an example of the biomarker data that might come from a more realistic health-tracking study. We construct this figure by randomly selecting data at various times from the study population. The time series in

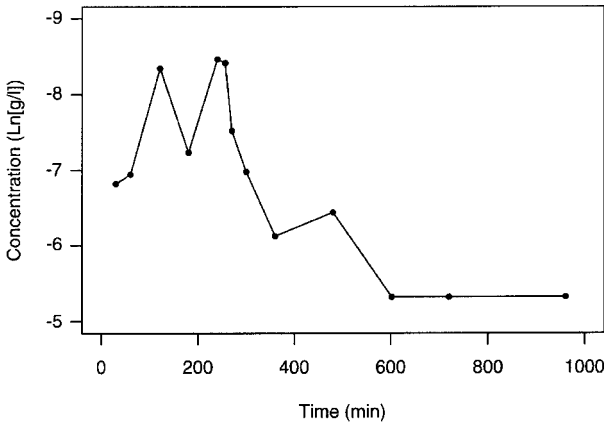

Figure 5. TCE concentration in venous blood generated by randomly sampling the data in Figure 1.

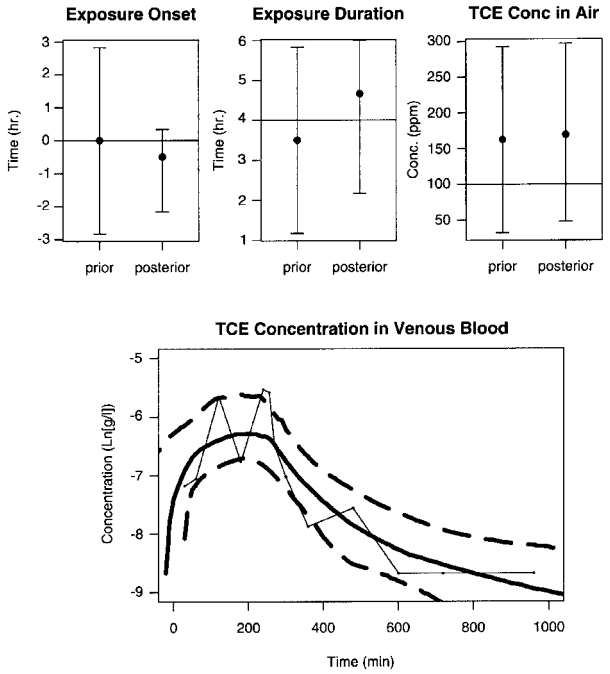

Figure 6. Reconstructed exposure profile and TCE concentration using the data set in Figure 5. The prior and posterior bars in the barplot represents the uncertainty before and after comparison to biomarker data, respectively. The whiskers in the barplot represent the two-sided 95% confidence interval, the circle is the median, and the horizontal line represents the value reported by Fisher (1998). The thick dotted lines in the time-series plot define the two-sided 95% confidence interval. The thick solid line is the median.

Figure 5 appears to represent a single-mode population with error due to random inter-individual variability. Figure 6 shows the exposure reconstruction. The uncertainty around the predicted TCE concentrations, though wide, appears to bound the data correctly. The predicted exposure conditions are also wide but correctly bound what is believed to be the exposure conditions.

Discussion

In the preceding paragraphs, we applied Bayesian inference as a tool to reconstruct exposure conditions. The process is not straightforward and can be confounded by heterogeneity and variability in exposure conditions and metabolism. Although the results suggest the existence of two subgroups that were exposed to different air concentrations, considerable uncertainty remains in the exposure estimates. For example, we were unable to predict with high certainty the concentration of TCE in air from any of the data sets. An important limitation of using PBPK models to interpret biomarker data may be the nonuniqueness of inverse solutions due to the combined effects of variability in exposures and human pharmacokinetics — particularly metabolism.

In spite of the difficulties noted above, the Bayesian exposure assessment approach offers some key advantages for reconstructing exposures. An important attribute of the approach is its flexibility for analyzing and comparing the utility of various types or quantities of data without excessive computational or numerical burdens. For example, we recalculated exposure reconstructions using data from subgroup A only, B only, A and B combined, and a randomly sampled set without re-executing Monte Carlo simulations of the PBPK model.

We can also test whether more data or other types of data could improve the exposure reconstructions. For example, could better information about the environmental, exposure, and pharmacokinetic properties improve reliability of exposure reconstruction, and at what cost? As a demonstration of such an exploratory search, we recalculated each of the exposure reconstructions with the added assumption that (1) the exposure duration is known or (2) the time of exposure onset is known. Including the additional information did not reduce the uncertainty in the predicted concentration of TCE in air in any of the reconstructions; we therefore did not plot these latter results. However, this additional exercise suggests that the population-scale variability of the pharmacokinetics alone or in combination with other unknowns dominate the uncertainty of the predicted TCE concentration in air. From this we learn that one should not expend excessive resources obtaining exposure onset and duration data if their primary objective is reducing uncertainty in the predicted concentration of TCE in air.

Recommendations

Given the potential problems we observe for exposure reconstruction, it is important to consider what properties of a biomarker or other types of information can improve the reliability of the exposure reconstruction process. As the dose delivered to an exposed individual depends on (a) the time

scales of the pharmacokinetics of the agent, (b) the route of entry, and (c) the rate of intake or uptake at the human/environment boundaries, we must first recognize the importance of selecting the most appropriate time scale for collecting information. For example, resolving the temporal variability of exposure events requires differentiating between (i) a recent relatively mild peak exposure or a long-term relatively high exposure, or (ii) many common exposures occurring simultaneously. How persistent must the biomarker (body burden) be relative to exposure duration for this task? That is, if we want to infer exposure to a pollutant over a 1-week period, we must consider the minimum biological half-life of the biomarker required to reconstruct doses reliably.

Of similar importance is the time characteristics of PBPK model input parameters. For example, breathing automatically averages air concentrations over the duration of a breath (at least at the lung level), and drinking and eating are discontinuous. These conditions define both limiting processes and the time-averaging periods for health-relevant doses. PBPK models establish the characteristic time of pollutant within the human body. This time is needed to classify exposures as intermittent (going to zero or negligible levels periodically or randomly) *versus* continuous (with various degrees of stability/uncertainty), and as sequential, additive, or cumulative. Exposures to carcinogens at low rates of uptake (when the cumulative damage rate is proportional to the uptake) require a dose assessment with a time resolution that need only reflect the cumulative uptake or intake of the agent into the body. In contrast, an agent such as an acid gas (where short-term nonlinear effects with large variations in respiratory susceptibility are important) requires much more detailed specification of the time, population, and even spatial resolution of exposure. PBPK modeling may help in the design, timing, and placement of measurements that are necessary for developing such a technique.

Based on the results presented here, we propose that the persistence of the biomarker should be long relative to exposure duration for estimating long-term, or population scale, exposure effects. That is, if one wants to infer exposure to a pesticide over a 1-week period, it is useful to have a biomarker that persists for more than a week. Perhaps the best marker is one that is truly cumulative, that is, a chromosome aberration that is heritable from one cell generation to another. But how do we establish a lower bound on biomarker persistence? In the TCE data set, high temporal resolution throughout the exposure event allowed us to reconstruct exposures for each of the subgroups with reasonable success. Could even better resolution or a more persistent biomarker improve the predictions of the TCE concentration in air?

Better understanding of how variability in metabolism impacts the reliability of PBPK models to determine unique links between exposure and tissue dose is also critical when designing exposure classification studies. Are PBPK models so limited by metabolic pathway uncertainty/variability, that higher compartment resolution — and thus the precision of the model predictions or exposure estimates — is of no value? This could explain some of the uncertainty in the exposure reconstructions for the various combinations of the TCE data set. Or will age-specific variation of physiological parameters, particularly for persistent chemicals such as DDT, PCBs, TCDD that can accumulate in tissues (fat in particular) over decades make population-wide exposure-to-dose impact assessments too specific to certain subpopulations, in this case according to age breakdown?

Such exploratory studies and classification studies should be carried out before and during health-tracking studies to ensure that the practitioner obtains the most informative data, whether they are environmental, biological, or chemical (i.e., the properties of the pollutant). However, these data for a wide number of chemicals and exposure routes are not adequately described in the current literature on PBPK models or epidemiological surveys.

As a last point, we note that the reconstruction of the combined data set (Figures 3 and 4) stresses the importance of appropriately developing and applying PBPK models that are consistent with the data. For example, reconstructing an apparent two-mode population using a singe-model PBPK model resulted in incorrect uncertainty estimates. Either the subgroups should be analyzed separately or a two-mode model should have been developed. When the dataset is sparse (Figure 6), the practitioner may satisfactorily describe the data with a single-mode PBPK model.

Concluding remarks

Heath assessments require public-health tracking, and health tracking requires exposure tracking, a process for linking body burdens to a distribution of population doses as reflected in biomarkers or some other longer-term exposure indicator. Of particular importance is the ability to classify and link health outcomes with exposures to harmful substances — pesticides, industrial chemicals, combustion products, consumer products, etc. This requires both sufficient and reliable information about population exposures and doses to those pollutant sources that most significantly contribute to observed markers of potential health detriment. To make better use of body burden/biomarker data in the process of public health tracking, two essential scientific research tools, models and measurements, must be better integrated. Models provide the means to integrate and interpret measurements, design hypothesis-driven experiments, and predict the effectiveness of risk management strategies. Measurements, in turn, provide tests of the models and ''ground truth''.

We presented an integrated approach for improving the communication between the needs and capabilities of modeling with the needs and capabilities of health surveys. The Bayesian statistical approach allows for easily quantifying the value of biomarkers, or other environmental, chemical, biological parameters, for exposure classification. The analysis of the TCE data set demonstrated some of the attributes of this type of approach and, perhaps more importantly, highlighted the difficulties with back-estimating exposures to a large and diverse population. We provide in the discussion and recommendation sections a list of factors that contribute to these difficulties and recommend several quantifiable measures that should be studied before and during health and exposure surveys to identify what types of biomarkers to gather, when to gather them, and how much data is needed.

Acknowledgments

This work was supported in part by the US Environmental Protection Agency National Exposure Reasearch Laboratory through Interagency Agreement # DW-988-38190-01-0 and carried out at Lawrence Berkley National Laboratory through the US Department of Energy under contract Grant No. DE-AC03-76SF00098. This work was also supported in part by the Health Tracking Center of Excellence at the University of California Berkley, and funded by interagency agreement No. U50CCUT922409-01 with US Centers for Disease Control. We thank M. Rigas and F.Power from the US EPA for their help in developing the PBPK model. We also thank M.Macleod, W.Riley, and the reviewers for their helpful comments on an earlier version of this manuscript.

References

Blancato J., Power F., Fisher J., and Kantor E. Application and use of dose estimating exposure model (DEEM) for dose comparisons after exposure to trichloroethylene (TCE). 39th Annual Meeting, Society of Toxicology, Philadelphia, PA, March 19–23, 2002.

Bogen K.T., and Spear R.C. Integrating uncertainty and interindividual variability in environmental risk assessment. *Risk Anal* 1987: 7(4): 427–436.

Box G.E.P., and Tiao G.C. *Bayesian Inference in Statistical Analysis*. Wiley Classics, New York, 1973.

Bois F.Y., Gelman A., Jiang J., Maszle D.R., Zeise L., and Alexeef G. 'Population toxicokinetics of tetrachloroethylene. *Arch Toxicol* 1996a: 70: 347–355.

Bois F.Y., Jackson E.T., Pekari K., and Smith M.T. Population toxicokinetics of benzene. *Environ Health Persp* 1996b: 104: 1405–1411.

Bois F.Y. Applications of population approaches in toxicology. *Toxicol Lett* 2001: 120: 385–394.

Brand K.P., and Small M.J. Updating uncertainty in an integrated risk assessment — conceptual framework and methods. *Risk Anal* 1995: 15(6): 719–731.

Brown P.N., Byrne G.D., and Hindmarsh A.C. VODE, a variable-coefficient ODE solver. *SIAM J Sci Stat Comput* 1989: 10(5): 1038–1051.

Castorina R., Bradman A., McKone T.E., Barr D.B., Harnly M.E., and Eskenazi B. Assessing cumulative organophosphate pesticide exposure and risk among pregnant women living in an agricultural community. *Environ Health Persp* 2002; submitted.

Centers for Disease Control and Prevention (CDC). National report on human exposure to environmental chemicals. Published by and available from the Centers for Disease Control and Prevention, National Center for Environmental Health, division of Laboratory Sciences, Atlanta, GA, 2001, www.cdc.gov/nceh/dls/report.

Chinnery R., and Gleason K. A compartment model for the prediction of breath concentration and absorbed dose of chloroform after exposure while showering. *Risk Anal* 1993: 13: 51–62.

Dakins M.E., Toll J.E., Small M.J., and Brand K.P. Risk-based environmental remediation — Bayesian Monte Carlo analysis and the expected value of sample information. *Risk Anal* 1996: 16(1): 67–79.

Finkel A.M., and Evans J.S. Evaluating the benefits of uncertainty reduction in environmental health risk management. *JAPCA* 1987: 37: 1164–1171.

Fisher J. Physiologically based pharmacokinetic models for trichloroethylene and its oxidative metabolites. An EPA report draft 2, April 7, 1998 from the Operational Toxicology branch, Human Effectiveness directorate, Wright–Patterson AFB, OH 45433, 1998.

Gargas M.L., Burgess R.J., Voisard D.E., Cason G.H., and Andersen M.E. Partition coefficients of low-molecular weight volatile chemicals in various liquids and tissues. *Toxicol Appl Pharmacol* 1989: 98: 87–99.

Gelman A., Bois F.Y., and Jiang J. Physiological pharmacokinetic analysis using population modeling and informative prior distributions. *J. Am Stat Associ* 1996: 91: 1400–1412.

Gelman A., Carlin J.B., Stern H.S., and Rubin D.B. *Bayesian Data Analysis*. Chapman & Hall/CRC, Boca Raton, FL, 1995.

Jo W.K., Weisel C.P., and Lioy P.J. Chloroform exposure and body burden from showering with chlorinated tap water. *Risk Anal* 1990: 10: 575–580.

Klaassen C.D. Distribution, excretion and absorption of toxicants. In: Klaassen C. D., Amdur M. O., and Doul J. (Eds.). *Cassarett and Doul's Toxicology, the Basic Science of Poisons*. Pergamon Press, New York, NY, 1996, pp. 33–63.

McKone T.E. Linking a PBPK model for chloroform with measured breath concentrations in showers: implications for dermal exposure models. *J Expos Anal Environ Epidemiol* 1993: 3: 339–365.

McKone T.E., and Daniels J.I. Estimating human exposure through multiple pathways from air, water, and soil. *Regul Toxicol Pharmacol* 1991: 13: 36–61.

Morgan M.G., and Henrion M. *Uncertainty: A Guide to Dealing with Uncertainty in Quantitative Risk and Policy Analysis*. Cambridge University Press, New York, NY, 1990.

Pinsky P.F., and Lorber M.N. A model to evaluate past exposure to 2,3,7,8-TCDD. *J Expos Anal Environ Epidemiol* 1998: 8(2): 187–206.

Ramsey J.C., and Andersen M.E. A physiologically-based description of the inhalation pharmacokinetics of styrene in rats and humans. *Toxicol Appl Pharmacol* 1984: 73: 159–175.

Roy A., and Georgopoulos P.G. Reconstructing week-long exposures to volatile organic compounds using physiologically based pharmacokinetic models. *J Expos Anal Environ Epidemiol* 1998: 8(3): 407–422.

Sexton K., Kleffman D.E., and Callahan M.A. An introduction to the National Human Exposure Assesment Survey (NHEXAS) and

related Phase I Field Studies. *J Expos Anal Environ Epidemiol* 1995: 5: 3229–3232.

Sohn M.D., Small M.J., and Pantazidou M. Reducing uncertainty in site characterization using Bayes Monte Carlo methods. *J Environ Engi — ASCE* 2000: 126(10): 893–902.

Spear R.C., and Bois F.Y. Parameter variability and the interpretation of physiologically based pharmacokinetic modeling results. *Environ Health Persp* 1994: 102(11): 61–66.

Small M.J., and Fischbeck P.S. False precision in Bayesian updatng with incomplete models. *Hum Ecol Risk Assess* 1999: 5(2): 291–301.

Taylor A.C., Evans J.S., and McKone T.E. The value of animal test information in environmental control decisions. *Risk Anal* 1993: 13(4): 403–412.

US Environmental Protection Agency (USEPA). Exposure and Human Health Reassessment of 2,3,7,8-Tetrachlorodibenzo-p-Dioxin (TCDD) and Related Compounds. (DRAFT). *Rpt No. EPA/600/P-001*, US Environmental Protection Agency Office of Research and Development, National Center for Environmental Assessments, Washington, DC, 2001. Report available at: http://www.epa.gov/ncea/dioxin.htm.

US Environmental Protection Agency (USEPA). Risk Assessment Guidance for Superfund. Volume I. Human Health Evaluation Manual (Part A). US Environmental Protection Agency, Office of Emergency and Remedial Response. Washington, DC, 1989. EPA/540/1-89/002.

US Environmental Protection Agency (USEPA). Environmental Protection Agency, Guidelines for Exposure Assessment: Notice. *Federal Register* 1992: 57(104): 22888–22938.

US Environmental Protection Agency (USEPA). Children's total exposure to persistent pesticides and other persistent organic pollutants (CTEPP). US Environmental Protection Agency, National Exposure Research Laboratory, Research Triangle Park, NC, 2002.

Vicini P., Pierce C.H., Dills R.L., Morgan M.S., and Kalman D.A. Individual prior information in a physiological model of 2H8-toluene kinetics: an empirical Bayes estimation strategy. *Risk Anal* 1999: 19(6): 1127–1134.

Wakefield J. The Bayesian analysis of population pharmacokinetic models. *J Am Stat Assoc* 1996: 91(433): 62–75.

Wallace L.A. Human exposure and body burden for chloroform and other trihalomethanes. *Crit Rev Environ Sci Technol* 1997: 27(2): 113–194.

Wallace L.A., and Pellizzari E.D. Recent advances in measuring exhaled breath and estimating exposure and body burden for volatile organic compounds (VOCs). *Environ Health Persp* 1995: 103: 95–98.

Weisberg S., *Applied Linear Regression*, 2nd ed. John Wiley & Sons, NY.

Zartarian V.G., Ott W.R., and Duan N.H. A quantitative definition of exposure and related concepts. *J Expos Anal Environ Epidemiol* 1997: 7(4): 411–437.