

Physiological Pharmacokinetic Analysis Using Population Modeling and Informative Prior Distributions

Andrew GELMAN, Frederic BOIS, and Jiming JIANG

We describe a general approach using Bayesian analysis for the estimation of parameters in physiological pharmacokinetic models. The chief statistical difficulty in estimation with these models is that any physiological model that is even approximately realistic will have a large number of parameters, often comparable to the number of observations in a typical pharmacokinetic experiment (e.g., 28 measurements and 15 parameters for each subject). In addition, the parameters are generally poorly identified, akin to the well-known ill-conditioned problem of estimating a mixture of declining exponentials. Our modeling includes (a) hierarchical population modeling, which allows partial pooling of information among different experimental subjects; (b) a pharmacokinetic model including compartments for well-perfused tissues, poorly perfused tissues, fat, and the liver; and (c) informative prior distributions for population parameters, which is possible because the parameters represent real physiological variables. We discuss how to estimate the models using Bayesian posterior simulation, a method that automatically includes the uncertainty inherent in estimating such a large number of parameters. We also discuss how to check model fit and sensitivity to the prior distribution using posterior predictive simulation. We illustrate the application to the toxicokinetics of tetrachloroethylene (perchloroethylene [PERC]), the problem that motivated this work.

KEY WORDS: Bayesian methods; Hierarchical models; Informative prior distributions; Markov chain simulation; Pharmacokinetics; Posterior predictive checks; Sensitivity analysis; Tetrachloroethylene; Toxicokinetics.

1. INTRODUCTION

We discuss statistical estimation of models in *pharmacokinetics*—the study of the mechanisms and kinetics of chemicals' absorption, distribution, metabolism, and elimination by the body in animals or humans (e.g., Rowland and Tozer 1989). Our purposes are twofold: to estimate the parameters of the model and, more important, to use the model to estimate particular quantities of interest, such as the rate at which the compound is metabolized under specified conditions (e.g., breathing air with a known concentration of the compound). Inferences about such estimands can then be applied to risk analyses. For this and other purposes, it is useful to have quantitative assessments of uncertainty.

For the purposes of public health, it is desirable to estimate the distribution of individual characteristics, such as parameters in a pharmacokinetic model, and the associated quantities of interest over the population. Even better is to determine how these quantities vary as functions of individual characteristics such as sex, age, and body mass. In any case, however, the goal is to measure the distribution of population characteristics, rather than a single set of parameters representing the "average person." Hierarchical models have a long history in pharmacokinetics (see, e.g., Racine-Poon and Smith 1990, Sheiner 1984, Sheiner, Rosenberg, and Melmon 1972, and Yuh et al. 1994 for a comprehensive bibliography).

We discuss the estimation of pharmacokinetic models using indirect measurements on individuals. A known dose of the compound is given to a subject, and then the concentration of the compound is measured in the subject's blood and exhaled air. Repeated measurements are taken over several hours. A *pharmacokinetic model* is set up that relates the concentration of the chemical in different *compartments* (e.g., blood, fatty tissues, other tissues) within the subject's body, and the parameters of the model for that individual are estimated based on the time series of measurements.

A characteristic difficulty of estimating pharmacokinetic models is that they predict a pattern of concentration over time that is close to a mixture of declining exponential functions, with the amplitudes and decay times of the different components corresponding to functions of the model parameters. It is well known (see, e.g., Acton 1970, p. 253) that estimation of the decay times of a mixture of exponentials is an ill-conditioned problem; that is, the parameters in such a model are hard to estimate simultaneously.

Because of the difficulty of estimation, it is common practice to reduce the number of parameters to be estimated in a pharmacokinetic model, either by fixing all but a few parameters to guessed values and estimating the remaining parameters or by setting up a model with very few parameters—a one- or two-compartment model. The first approach has the serious problem of producing inaccurate estimates and underestimating uncertainty when the parameters to be fixed are not known accurately, as is typical in experiments with live human subjects (see the discussion in Woodruff and Bois 1993). The second approach has been used with some success recently (e.g., Wakefield, Smith, Racine-Poon, and Gelfand 1994), but has the limitation of not allowing realistic multicompartment models to be fit.

Andrew Gelman is Associate Professor, Department of Statistics, Columbia University, New York, NY 10027. Frederic Bois is Staff Scientist, Lawrence Berkeley National Laboratory, Berkeley, CA 94720. Jiming Jiang is Assistant Professor, Department of Statistics, Case Western Reserve University, Cleveland, OH 44106. The authors thank Xiao-Li Meng, Phillip Price, and several referees for helpful comments; Don Maszle for computing assistance; Dr. A. C. Monster for providing experimental data; the American Lung Association for financial support; the National Science Foundation for grant DMS-9404305 and Young Investigator award DMS-9457824; and Springer-Verlag for permission to reprint data from one table and two figures from *Archives of Toxicology*.

© 1996 American Statistical Association
Journal of the American Statistical Association
December 1996, Vol. 91, No. 436, Applications and Case Studies

The complexity required of a model will depend on the aims of the analysis and on the data available.

The problem of estimating metabolism from indirect data can be addressed using a *physiological* pharmacokinetic model; that is, one in which the individual and population parameters have direct physical interpretations (e.g., blood flow through the fatty tissue, tissue/blood partition coefficients). Such models allow the simulation and prediction of a variety of doses in specific target organs, while including complex nonlinear behavior of metabolic pathways (Balan and Gex-Fabry 1990; Gerlowski and Jain 1983; Leung 1991; Menzel 1987; Ritschel and Banerjee 1986). These models are also rich in structure and permit the identification of many of their parameter values through prior (e.g., published) physiological data. Following our own work (Bois, Zeise, and Tozer 1990) and that of others (Bogen and MacKone 1988; Guberman and Fernandez 1974; Koizumi 1989; Ward, Travis, Hetrick, Andersen, and Gargas 1988), we decided to use such models. Because the parameters of these models are essentially impossible to estimate from the data alone, it is crucial that they have physical meaning and can be assigned informative prior distributions.

Recent developments in Monte Carlo Bayesian statistical computing have removed some of the obstacles that could hamper the alliance of statistics and physiological pharmacokinetic modeling. We describe the application to our model of Markov chain simulation, which is a particularly simple and powerful tool. A similar approach to pharmacokinetic modeling was presented by Wakefield et al. (1994) and Wakefield (1996), for classical one-compartment models. By using a physiological model, we increase the number of parameters but can take advantage of a large body of prior information on the parameter values.

We present our general approach in the context of the application to the metabolism of tetrachloroethylene from exposure in air. As results we report predictions and interval estimates on the fraction of tetrachloroethylene (perchloroethylene [PERC]) metabolized at low and high doses in humans. We discuss how this information can improve our use of toxicokinetic modeling for exposure to toxic substances in the air. (*Pharmacokinetic* refers to drugs, and *toxicokinetic* refers to toxic compounds; for convenience, we use the former term to refer to both.)

1.1 Background and Experimental Data for the PERC Example

Detailed results of our analysis of PERC are given by Bois et al. (1996). We briefly give the background here. The greatest occupational exposures to PERC are in dry cleaners, where 50 parts per million (ppm) is a typical concentration in air. (In comparison, the normal concentration of PERC in air is about .001 ppm.) PERC is carcinogenic in animals and is likely carcinogenic in humans as well. Although the actual carcinogenic metabolites have not yet been identified, the fraction metabolized is likely to be a better measure of dose than is PERC exposure itself. Inference is needed because the total amount of metabolites formed has never been directly measured; other endpoints

have been observed, such as PERC blood and exhaled air concentrations or the blood concentration and urinary excretion of particular metabolites. Extrapolation is needed to infer low-dose metabolism from high-dose exposures; after a high exposure, the fraction metabolized amounts to a few percent, but this may not hold at low doses because of the possible saturability of PERC metabolism (Ikeda, Ohtsuki, Inamura, and Komoike 1972; Ohtsuki, Sato, Koizume, Kumai, and Ikeda 1983).

We used previously collected data on the concentrations of PERC in exhaled air and venous blood for six healthy young adult Caucasian male volunteers exposed to PERC in an inhalation chamber for 4 hours (Monster, personal communication; Monster, Boersma, and Steenweg 1979), with concentrations of PERC measured over the week following exposure. Two exposure levels were used: 72 ppm (204 $\mu\text{g/L}$) and 144 ppm (409 $\mu\text{g/L}$). A subset of the experimental data is displayed in Figure 1; for clarity, measurements are shown for only two of the six subjects.

2. SETTING UP A MODEL

We set up a model in several stages. A *pharmacokinetic model* describes the flow of the compound in the body in terms of several parameters for each individual. A *population model* describes the distribution of the parameters in the population as a function of several population parameters, which in turn have a *prior distribution* based on scientific knowledge. Finally, the *measurement model* describes the distribution of deviations of the data from their expected values predicted from the pharmacokinetic model. Once all parts of the model have been specified, they are combined into a posterior distribution for Bayesian analysis.

2.1 Individual Pharmacokinetic Model

We use a previously developed model (Bois et al. 1990) in which the human body is divided into four compartments: poorly perfused tissues, well-perfused tissues, fat, and liver. Given a known concentration of the compound in the air, the concentration of the compound in each compartment over time is governed by a first-order differential equation, with parameters for the volume, blood flow, and partition coefficient (i.e., equilibrium concentration relative to the blood) of each compartment. Compartments are assumed to be homogeneous, and distribution is flow limited. Pulmonary exchanges are modeled by assuming instantaneous equilibrium between alveolar air, venous blood, and arterial blood, according to Andersen (1981). In addition, differential equations of the form $dC_s/dt = (C_{\text{art}} - C_s/P_s)F_s/V_s$ describe the time dependence of the concentration C_s of the compound in each compartment s as a function of blood flow F_s , volume V_s , arterial blood concentration C_{art} , and partition coefficient P_s . These equations are linear except for the liver compartment, in which a Michaelis-Menten term describes the kinetics of saturable enzymatic reactions (Michaelis and Menten 1913) for the metabolic clearance of the compound. For this compartment, $dC_s/dt = (C_{\text{art}} - C_s/P_s)F_s/V_s - V_{\text{max}}C_s/[V_s(K_m + C_s)]$, where V_{max} is the maximum rate of metabolism and K_m is the

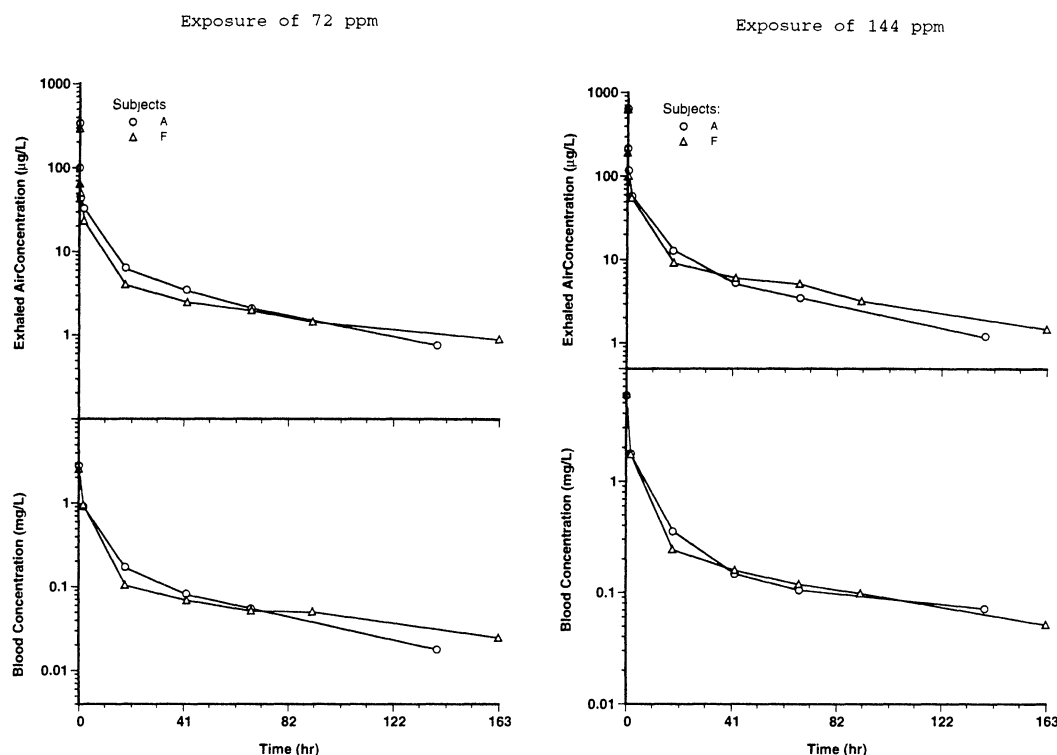


Figure 1. Individual Exhaled Air and Blood Concentrations Postexposure Time Profiles for Two Volunteers (\circ , Subject A; \triangle , Subject F) Exposed to PERC for 4 Hours (Experimental Data of Monster et al. 1979). Measurements began immediately after PERC exposure ended. The left and right halves of the figure correspond to exposed concentrations of 72 and 144 ppm. For clarity, measurements for only two of the six experimental subjects are shown.

Michaelis constant. We use the notation $\theta_k = (\theta_{k1}, \dots, \theta_{kL})$ as the vector of L parameters associated with person k . In our particular setup, $L = 15$ (see Table 1 for a list of the parameters).

This sort of model is standard in pharmacokinetics; see the list of references for more details. Given the values of the physiological parameters and initial exposure conditions, we solve the differential equation to obtain concentrations of the compound and rate of metabolism as a function of time, using Gear's routine for stiff systems (Gear 1971a,b) as implemented in our own software, MC-Sim (available from STATLIB), to compute expected concentrations in blood and exhaled air and quantity metabolized for a given period.

2.2 Population Model

One goal of this work is to estimate the distribution of the individual pharmacokinetic parameters, in particular the distribution of predicted values such as fraction metabolized, in the general population or in subsets of it. In an experiment with K individuals, we set up a hierarchical model on the K vectors of parameters, $\theta_1, \dots, \theta_K$. The hierarchical model allows us to estimate the population variability and also improves the estimates for the individual subjects. Both these points have been discussed by Yuh et al. (1994) and, in a Bayesian context, by Wakefield et al. (1994).

A skewed, lognormal-like distribution is generally observed for biological parameters, whereas most, if not all, parameters also have physiological bounds. For each

individual, k , the individual pharmacokinetic parameters, θ_{kl} , after log transformation and appropriate scaling (see later), are modeled as normal with population mean μ_l and variance Σ_l^2 , truncated to lie within a specified number (typically, 3) of standard deviations of the mean, where $l = 1, \dots, L$ indexes the pharmacokinetic parameters in the model. The distributions are truncated to restrict the model parameters to scientifically reasonable values (see Sec. 2.3). In addition, the truncations serve a useful role when we monitor the simulations of the parameters from their posterior distribution. If the simulations for a parameter are stuck near truncation points, this indicates that the data and the pharmacokinetic model strongly contradict the prior distribution, and some part of the model should be reexamined. Because of the ill-conditioned nature of the problem, we fear that if we were to fit a completely untruncated model, we could possibly end up with parameter estimates that are scientifically unreasonable. If the simulations for a parameter are stuck near a truncation point, this would indicate a flaw in the model. An alternative approach would be to use untruncated normal distributions and then check whether simulations happened to be more than 3 standard deviations from the prior mean; using truncation is a way to ensure that the parameter simulations are in a reasonable range.

A minor modeling difficulty is that some of the parameters are constrained by definition. For each individual k , the parameters $\theta_{k2}, \theta_{k3}, \theta_{k4}$, and θ_{k5} , the fractions of blood flow to each compartment, are constrained to sum to 1. Also, the parameters θ_{k6}, θ_{k7} , and θ_{k8} , corresponding to the scaling coefficients of the organ volumes, are constrained to

Table 1. Prior and Posterior (Fitted) Distributions of the Physiological Parameters, for Each Individual and for the Population, From Bois et al. (1996).

Parameter	Population prior	Posterior distributions for individuals						Population posterior
		A	B	C	D	E	F	
Ventilation/perfusion ratio (VPR)	1.6($\times \div 1.3$) $\times \div 1.3$	1.16 $\times \div 1.15$	1.26 $\times \div 1.15$	1.19 $\times \div 1.14$	1.33 $\times \div 1.15$	1.22 $\times \div 1.15$.961 $\times \div 1.15$	1.19 $\times \div 1.13$
Blood flow, well-perfused tissues (Fwp)	.47($\times \div 1.17$) $\times \div 1.17$.653 $\times \div 1.06$.658 $\times \div 1.07$.647 $\times \div 1.07$.660 $\times \div 1.06$.626 $\times \div 1.08$.606 $\times \div 1.08$.637 $\times \div 1.06$
Blood flow, poorly perfused tissues (Fpp)	.20($\times \div 1.22$) $\times \div 1.22$.121 $\times \div 1.12$.123 $\times \div 1.13$.127 $\times \div 1.13$.123 $\times \div 1.12$.132 $\times \div 1.13$.134 $\times \div 1.13$.129 $\times \div 1.11$
Blood flow, fat (Ff)	.07($\times \div 1.27$) $\times \div 1.27$.048 $\times \div 1.13$.0442 $\times \div 1.13$.0462 $\times \div 1.14$.0437 $\times \div 1.13$.0507 $\times \div 1.14$.0582 $\times \div 1.14$.0488 $\times \div 1.12$
Blood flow, liver (Fl)	.25($\times \div 1.15$) $\times \div 1.15$.173 $\times \div 1.15$.170 $\times \div 1.16$.175 $\times \div 1.15$.168 $\times \div 1.15$.185 $\times \div 1.16$.195 $\times \div 1.15$.179 $\times \div 1.11$
Volume, well-perfused tissues (Vwp)	.27($\times \div 1.36$) $\times \div 1.36$.189 $\times \div 1.14$.201 $\times \div 1.15$.202 $\times \div 1.15$.201 $\times \div 1.15$.183 $\times \div 1.15$.188 $\times \div 1.14$.196 $\times \div 1.09$
Volume, poorly perfused tissues (Vpp)	.55($\times \div 1.17$) $\times \div 1.17$.649 $\times \div 1.04$.636 $\times \div 1.05$.636 $\times \div 1.05$.636 $\times \div 1.05$.655 $\times \div 1.04$.65 $\times \div 1.04$.641 $\times \div 1.03$
Volume, liver (Vl)	.033($\times \div 1.1$) $\times \div 1.1$.032 $\times \div 1.1$.033 $\times \div 1.1$.033 $\times \div 1.1$.033 $\times \div 1.1$.033 $\times \div 1.1$.032 $\times \div 1.1$.033 $\times \div 1.04$
Partition coeff, blood/air (Pba)	12($\times \div 1.5$) $\times \div 1.3$	15.1 $\times \div 1.04$	16.4 $\times \div 1.03$	15.3 $\times \div 1.04$	15.6 $\times \div 1.04$	18.7 $\times \div 1.04$	15.8 $\times \div 1.04$	16.0 $\times \div 1.11$
Partition coeff, well-perfused (Pwp)	4.8($\times \div 1.5$) $\times \div 1.3$	1.83 $\times \div 1.15$	1.98 $\times \div 1.16$	1.95 $\times \div 1.16$	2.00 $\times \div 1.16$	1.83 $\times \div 1.15$	1.83 $\times \div 1.14$	1.92 $\times \div 1.12$
Partition coeff, poorly perfused (Ppp)	1.6($\times \div 1.5$) $\times \div 1.3$	2.94 $\times \div 1.08$	2.59 $\times \div 1.09$	2.51 $\times \div 1.09$	2.76 $\times \div 1.08$	4.06 $\times \div 1.09$	2.96 $\times \div 1.09$	2.90 $\times \div 1.15$
Partition coeff, fat (Pf)	125($\times \div 1.5$) $\times \div 1.3$	82.3 $\times \div 1.08$	69.1 $\times \div 1.08$	73.9 $\times \div 1.08$	49.1 $\times \div 1.08$	171 $\times \div 1.09$	85.4 $\times \div 1.07$	84.1 $\times \div 1.28$
Partition coeff, liver (Pl)	4.8($\times \div 1.5$) $\times \div 1.3$	2.93 $\times \div 1.32$	3.07 $\times \div 1.33$	3.21 $\times \div 1.32$	3.09 $\times \div 1.33$	3.16 $\times \div 1.33$	2.94 $\times \div 1.32$	3.08 $\times \div 1.12$
Max metabolic rate in liver (VMI)	.042($\times \div 10$) $\times \div 2$.0011 $\times \div 1.41$.00139 $\times \div 1.37$.00214 $\times \div 1.30$.00199 $\times \div 1.34$.00415 $\times \div 1.30$.00165 $\times \div 1.38$.00191 $\times \div 1.45$
K_m in liver (KMI)	16($\times \div 10$) $\times \div 1.5$.801 $\times \div 1.63$.754 $\times \div 1.61$.660 $\times \div 1.59$.742 $\times \div 1.57$.650 $\times \div 1.59$.771 $\times \div 1.60$.729 $\times \div 1.20$

NOTE: Prior distributions are lognormal truncated at ± 3 standard deviations, except for the blood flows and organ volumes, which are constrained to sum to constants as described in the text. The prior distributions are expressed as $e^{M(\times \div e^S)} \times \div e^{\Sigma_0}$, where S represents the uncertainty in the population mean and Σ_0 is the prior estimate of the population standard deviation. For example, the prior distribution for KMI (bottom row of the table) shows that the population mean is estimated to be 16, with a one-standard deviation range of [1.6, 160], and the population standard deviation is estimated to be 1.5. Thus the values of KMI for the six subjects are believed to be fairly similar, but their overall level is poorly known in the prior distribution. Posterior distributions for each individual are expressed as geometric mean $\times \div$ geometric standard deviation, as computed from the posterior simulations. Posterior distributions for the population are computed by simulating random additional persons from the population (based on the posterior distribution of the population parameters).

sum to .873 (the fraction of lean body mass not including bones) for each individual. Of these three parameters, θ_{k8} , the volume of the liver, is much smaller than the others and is known relatively precisely. For the purposes of modeling and computation, we transform the model in terms of a new set of parameters ψ_{kl} and then define the physiological parameters, θ_{kl} as follows:

$$\theta_{kl} = \frac{e^{\psi_{kl}}}{e^{\psi_{k2}} + e^{\psi_{k3}} + e^{\psi_{k4}} + e^{\psi_{k5}}}, \quad \text{for } l = 2, 3, 4, 5;$$

$$\theta_{kl} = (.873 - e^{\psi_{k8}}) \frac{e^{\psi_{kl}}}{e^{\psi_{k6}} + e^{\psi_{k7}}}, \quad \text{for } l = 6, 7;$$

and

$$\theta_{kl} = e^{\psi_{kl}}, \quad \text{for } l = 1 \text{ or } l \geq 8. \quad (1)$$

The parameters $\psi_{k2}, \dots, \psi_{k5}$ and ψ_{k6}, ψ_{k7} are not identified (e.g., adding any constant to $\psi_{k2}, \dots, \psi_{k5}$ does not alter the values of the physiological parameters, $\theta_{k2}, \dots, \theta_{k5}$), but they are assigned proper prior distributions, and so we can formally manipulate their posterior distributions. All substantive inferences are obtained on the scale of the θ_{kl} 's. In setting up the transformations, we purposely kept the same number of parameters so that a model of prior independence of the parameters would be scientifically reasonable. For example, if we were to fix the parameters ψ_{k2} to zero, then the range of possible values for θ_{kl} would not change, but setting up independent prior distributions on ψ_{k3}, ψ_{k4} , and ψ_{k5} would lead to an asymmetric and scientifically unreasonable distribution on $\theta_{kl}, l = 2, 3, 4, 5$.

The ψ_{kl} parameters are assumed to follow normal distributions with mean μ_l and variance Σ_l^2 , truncated at a specified number of standard deviations. Modeling on the scale

of ψ respects the constraints on θ while retaining the truncated lognormal distributions for the unconstrained components. Bois et al. (1996) and Gelman (1995) provide more discussion of this choice of constrained model. All computations are performed with the ψ 's, which are then transformed back to θ 's at the end to interpret the results on the natural scales.

In the model, the population distributions for the L parameters are assumed independent, a choice with which we are comfortable because of the parameterization and scaling used (e.g., blood flows as a proportion of the total rather than on absolute scales). In general, setting up independent population distributions means that between-subject information about one parameter will not be used to help estimate other parameters in the model. In a study with only six subjects, very little information will be available from the data on between-subject correlations, and we judge the gain in efficiency from modeling these correlations to be not worth the effort in modeling them, especially because we have set up the model to minimize such correlations.

2.3 Prior Distribution

To fit the population model, we assign prior distributions to the means and variances, μ_l and Σ_l^2 , of the L physiological parameters. We specify a prior distribution for each μ_l (normal with parameters M_l and S_l^2 , and based on substantive knowledge) and Σ_l^2 (inverse- χ^2 , centered at our estimate Σ_{0l}^2 of the true population variance and with a low number of degrees of freedom ν_l —typically set to 2—to indicate large uncertainties).

The hyperparameters M_l , S_l^2 , and Σ_{0l}^2 are based on estimates available in the literature. We set independent prior distributions for the μ_l 's and Σ_l 's because, under our parameterization, our prior information about the parameters is essentially independent to the best of our knowledge. The choice of values for the hyperparameters and the bounds for truncation (expressed as a number of standard deviations to be subtracted or added to the mean) for the PERC example are summarized in the first column of Table 1. In setting uncertainties, we try to be conservative and set the prior variances higher rather than lower when there is ambiguity in the biological literature (e.g., with the partition coefficients); as a result, truncation ranges of $M \pm 3S$ (or, for the case of VMI and KMI, $M \pm 2S$) include the scientifically plausible range of parameter values. We specify the prior distribution for the unconstrained parameters θ_l in terms of $\exp(M_l)$, $\exp(S_l)$, and $\exp(\Sigma_{0l})$, which can be interpreted as geometric means and standard deviations on the natural scale. For the constrained parameters θ_l , $l = 2, \dots, 7$, we specify the exponentials of the prior mean and standard deviation of the population mean on the logarithmic scale, $E(\log \theta_l)$, and the exponential of the prior estimate of the population standard deviation on the logarithmic scale, $SD(\log \theta_l)$.

Details and scientific background for the informative prior distributions for the PERC example are given by Bois et al. (1996). The parameters in a pharmacokinetic model can be divided into two categories: those that describe the

body and those that are characteristic of the compound being studied. The population distribution for any parameter depends on the population of individuals under study. For the PERC study, we were interested in the distributions for young adult white males, but we believe that the method we used to set the prior distributions is instructive in general.

We have some prior knowledge of μ_l and Σ_l^2 , at least in the form of standard values, as defined, for example, by the (Caucasian) "Reference Man" of the International Commission on Radiological Protection (ICRP 1975). To take into account known physiological dependencies between the pharmacokinetic model parameters (e.g., between organ volumes and body mass, or between alveolar ventilation rate and cardiac output), several of these parameters are expressed in proportion to the lean body mass or other parameter values, via scaling functions. The use of these scaling functions goes back at least to Adolph (1949). Some parameters—body mass, lean body mass, minute volume at rest, and age—can be measured directly on each individual and are fixed at their observed values. We label these exactly measured parameters as ϕ in the model.

The PERC-specific parameters in the model are partition coefficients and the Michaelis–Menten coefficients, whose prior distributions are set from the biological literature and allometric scaling from animal measurements, as detailed by Bois et al. (1996). The parameters are set rather roughly, and this is reflected in large variances for their prior distributions.

2.4 Measurement Model

At the individual level, for each subject, a series of measurements can be taken of exhaled air and blood concentrations. We label these as y_{1t} and y_{2t} , with t indexing time. The expected values of the exhaled air and blood concentrations are nonlinear functions, $f_m(\theta_l, \phi_l, E, t)$ of parameters, exposure level (E), and time, with $m = 1$ or 2 indexing blood or air concentration. Given the θ_l , ϕ_l , and E , one can evaluate the pharmacokinetic differential equation over time and compute f_1 and f_2 for all values at which measurements have been taken, thus obtaining the expected values of all the measurements.

The concentrations actually observed in expired air and blood are also affected by measurement errors, which are assumed, as usual, to be independent and lognormally distributed, with mean zero and a variance σ_m^2 (on the log scale). These also implicitly account for errors in the model; we have no particular reason to believe that modeling errors for air and blood measurements will be correlated. The variance vector σ^2 has two components— σ_1^2 for the measurements in blood and σ_2^2 for the measurements in exhaled air—because these measurements have different experimental protocols and thus are likely to have different precisions. For the Bayesian analysis, we assign the standard noninformative prior distribution to the variances, $p(\sigma_1^2, \sigma_2^2) \propto \sigma_1^{-2} \sigma_2^{-2}$ (see, e.g., Box and Tiao 1973). In virtually any application, there will be sufficient data so that the variances can be estimated accurately without requiring any prior information about them.

2.5 Hierarchical Model

For Bayesian inference, we obtain the posterior distribution (up to a multiplicative constant) for all the parameters of interest, given the data and the prior information, by multiplying all the factors in the hierarchical model: the data distribution, $p(y|\psi, \phi, \mathbf{E}, \mathbf{t}, \sigma^2)$; the population model, $p(\psi|\mu, \Sigma^2)p(\mu, \Sigma^2|\mathbf{M}, \mathbf{S}^2, \Sigma_0^2)$; and the prior distribution, $p(\mu, \Sigma^2|\mathbf{M}, \mathbf{S}^2, \Sigma_0^2)p(\sigma^2)$:

$$\begin{aligned} p(\psi, \mu, \Sigma^2, \sigma^2 | \mathbf{y}, \mathbf{E}, \mathbf{t}, \phi, \mathbf{M}, \mathbf{S}^2, \Sigma_0^2, \nu) \\ \propto p(y|\psi, \phi, \mathbf{E}, \mathbf{t}, \sigma^2) p(\psi|\mu, \Sigma^2) p(\mu, \Sigma^2|\mathbf{M}, \mathbf{S}^2, \Sigma_0^2) p(\sigma^2) \\ \propto \left[\prod_{k=1}^K \prod_{m=1}^2 \prod_t N(\log y_{jkm} | \log f_m(\theta_k, \phi_k, \mathbf{E}, \mathbf{t}), \sigma_m^2) \right] \\ \times \left[\prod_{k=1}^K \prod_{l=1}^L N_{\text{trunc}}(\psi_{kl} | \mu_l, \Sigma_l^2) \right] \\ \times \left[\prod_{l=1}^L N(\mu_l | M_l, S_l^2) \text{Inv} - \chi^2(\Sigma_l^2 | \nu_l, \Sigma_{0l}^2) \right] \sigma_1^{-2} \sigma_2^{-2}, \end{aligned} \quad (2)$$

where ψ is the set of vectors of individual-level parameters, μ and Σ^2 are the vectors of population means and variances, σ^2 are the two measurement variances, y are the concentration measurements, \mathbf{E} are the exposure concentrations, \mathbf{t} are the exposure times, ϕ are the individual-level covariates, and $\mathbf{M}, \mathbf{S}, \Sigma$, and ν are the hyperparameters. We use the standard Bayesian notation for probability densities (see, e.g., Gelman, Carlin, Stern, and Rubin 1995), with the additional notation N_{trunc} for the normal distribution truncated at the specified number of standard deviations from the mean. The indexes k, l, m , and t refer to subject, parameter, type of measurement (blood or air), and time of measurement. For experiments such as the PERC study, in which repeated studies are performed on the same subject, the first factor in (2) should include a factor for each study. The parameter θ_k in the first factor in (2) is a function of ψ_k as given by (1). To compute (2) as a function of the parameters, data, and experimental conditions, the function f must be computed numerically over the range of time corresponding to the experimental measurements.

3. BAYESIAN INFERENCE AND MODEL EVALUATION

A Bayesian analysis allows us to combine two forms of information: “prior knowledge” from the scientific literature and experimental data in the context of the physiological compartmental model. Neither source of information is complete. If prior knowledge were sufficient, then the experiments would not have had to be done; but existing data alone are typically insufficient to pin down the parameters to reasonable values. We wish to fit the data using scientifically plausible parameter values, so that the analysis outputs distributions of parameter values that are consistent with both the data and the prior information. Agreement of posterior with data and prior should be checked, as we discuss in Section 3.3.

3.1 Computation

Our goals are twofold: to fit a pharmacokinetic model to experimental data and to use this model to perform inferences about quantities of interest, such as the population distribution of the fraction of the compound metabolized at a given dose. We achieve these goals using random draws of the parameters from the posterior distribution, $p(\psi, \mu, \Sigma^2, \sigma^2 | y, E, t, \phi, M, S^2, \Sigma_0^2, \nu)$. Because the parameter vector has many components, we cannot just calculate the posterior distribution for a grid of reasonable values. Instead, we use a variant of the Gibbs sampler to perform random walks through the posterior distribution. The Gibbs sampler is an iterative procedure that is particularly convenient in the case of hierarchical models. This method belongs to a class of Markov chain Monte Carlo techniques that has recently received much interest. (For review and illustration see Gelfand, Hills, Racine-Poon, and Smith 1990, Gelfand and Smith 1990, Gelfand, Smith, and Lee 1992, Smith 1991, Tanner 1993, Wakefield et al. 1994, and others.)

The Gibbs sampler is based on performing a random walk through the posterior distribution by updating parameters or groups of parameters based on their conditional posterior distribution. For our model, we iteratively update the parameters in the following sequence: $\sigma^2, \Sigma^2, \mu, \psi_1, \dots, \psi_K$. Each of these is actually a vector parameter. The conditional distributions for the components of σ^2, Σ^2 , and μ are inverse-gamma, inverse-gamma, and normal and are well-known results from the Bayesian analysis of hierarchical linear models (see Gelfand et al. 1990). However, the conditional distributions for the parameters ψ have no closed form (because the pharmacokinetic function f is nonlinear), and so we sample from them using steps of the Metropolis algorithm, which requires only the ability to compute the posterior density up to a multiplicative constant, as in (2).

The Metropolis algorithm can jump one component of ψ_{kl} at a time (thus KL jumps in each iteration) or, using vector jumps, one individual at a time (thus K jumps in each iteration, with each jump of a L -dimensional vector ψ_k). Under either jumping rule, the only factors of the posterior density that need to be computed for the Metropolis algorithm are those corresponding to subject k . This is an important concern, because evaluating the function f to obtain expected values of measurements is the costliest part of the computation. The scaling constant of the jumping rule can be set to some reasonable value after some preliminary runs. For example, the scale can be set so that the acceptance rate of the Metropolis algorithm is approximately .44 for the one-component-at-a-time jumping rule or .23 for the vector jumping rule (see Gelman, Roberts, and Gilks 1995).

Five independent simulation runs were performed, with starting points obtained by sampling each ψ_{kl} at random from its prior distribution and then setting the population averages μ_l at their prior means, M_l . The samplers were then begun by drawing σ^2 and Σ^2 . Initially, 15,000 iterations were computed from each run using one-component-at-a-time Metropolis jumping, with a normal candidate distribution for component ψ_{kl} centered at the current value

Table 2. Convergence Monitoring of the Last 5,000 Iterations From Each Five Parallel Runs of Length 15,000.

Parameter	$\sqrt{\hat{R}}$
VPR	1.36
Fwp	1.27
Fpp	1.06
Ff	1.13
FI	1.22
Vwp	1.09
Vpp	1.05
VI	1.06
Pba	1.00
Pwp	1.02
Ppp	1.05
Pf	1.01
PI	1.08
VMI	1.06
KMI	1.13
σ_1	1.00
σ_2	1.00

NOTE: Parameters are described in Table 1. The convergence criterion, $\sqrt{\hat{R}}$, is the estimated potential scale reduction of the posterior interval for each parameter (on the log scale) if simulations were continued indefinitely (Gelman and Rubin 1992). At convergence, $\sqrt{\hat{R}} = 1$. Physiological parameters are shown only for the population geometric mean, (on the log scale), which were the slowest to convergence. Later simulations yielded values of $\sqrt{\hat{R}}$ closer to 1 with very little change in the posterior inferences for the parameters and other estimands of interest.

and with standard deviation equal to $S_l/20$. (The factor of 20 was set after some preliminary runs.) Each run of 15,000 iterations required about 36 hours of CPU time on a Sun Sparc 10 workstation. In practice, the model was gradually implemented and debugged over a period of months, and one reason for our trust in the results is their general consistency with earlier simulations of different variants of the model.

The convergence of the simulations was monitored by comparing the variance between and within sequences for all parameters of interest. Specifically, for each parameter we compute the variance ratio $\hat{R} = (\hat{V}/W)(df/(df - 2))$, where \hat{V} is the estimated posterior variance of the parameter using the simulations from all the sequences, W is the pooled within-sequence variance, and $(df/(df - 2))$ is a correction for sampling variability (see Gelman and Rubin 1992 and the associated discussion). Table 2 presents the convergence monitoring results for the population parameter set and the experimental variance estimates σ_1^2 and σ_2^2 . Convergence was better for the individual subject parameters, ψ_{kl} . At perfect convergence, $\sqrt{\hat{R}}$ values should all be equal to 1; after 15,000 iterations, most were less than 1.1. The highest value is 1.36, for the population ventilation over perfusion ratio (VPR) suggesting that additional simulations might reduce the posterior interval for the population VPR by up to a factor of 1.36.

A later set of $5 \times 50,000$ iterations was computed using one-individual-at-a-time Metropolis jumping, with a normal candidate distribution centered at the current value, with covariance matrix proportional to that of the initial simulation runs, and with scale tuning so that the acceptance rate was approximately .23. The later simulations were about 15 times faster per iteration because they required only $K = 6$ updating steps instead of $KL = 90$ steps per iteration. Be-

cause of storage limitations, every tenth iteration of the parameter vector was saved. After the additional runs, the $\sqrt{\hat{R}}$ values were reduced to all lie below 1.2, with very little change in the posterior inferences compared to the earlier simulations.

3.2 Inference for Quantities of Interest

When the simulations have reached approximate convergence, the posterior distribution for any parameter (or any function of the parameters) can be approximated by the set of draws from the last half of the simulated sequences (see Gelman and Rubin 1992), or a random subset of those draws if computation or storage is a concern. We are typically interested in the following summaries of the posterior distribution.

3.2.1 Extrapolation Scenarios. To obtain the distribution of the fraction of the compound metabolized under a specified scenario (e.g., inhalation with a specified concentration in the air), we can evaluate the differential equation model numerically under the appropriate input conditions. For each individual k , we compute the fraction metabolized for each simulated parameter vector ψ_k ; using the set of simulations yields a distribution of the fraction metabolized for that individual. The variance in the distribution for each individual is due to uncertainty in the posterior distribution of the physiological parameters, ψ_k .

To obtain the distribution of the fraction of PERC metabolized at low exposure and high exposure by each of the six subjects, we simulated continuous exposure to PERC for 3 weeks under two scenarios (.001 ppm and 50 ppm), using as input the posterior simulations of the parameter values. For each simulated scenario, the amount (in milligrams) metabolized the last day was divided by the amount inhaled on that day. The amount inhaled is equal to the alveolar ventilation rate (L/min) \times 1,440 (min) \times the PERC inhalation level (mg/L).

Similar simulations were performed for the population (i.e., predictions for an additional person, exchangeable with the subjects in the study) by simulating random vectors of the physiological parameters from their simulated population distribution. The variance in the population distribution of fraction metabolized includes posterior uncertainty in the parameter estimates and also real variation in the population.

Figure 2 shows the posterior distributions of the PERC fraction metabolized by each subject of the Monster et al. experiment at high (50 ppm) and low (.001 ppm) exposures under our model. The figure also shows the covariance between the high-dose and the low-dose estimates of the fraction metabolized in the six subjects. These simulated vectors are random draws from their joint (multivariate) distributions, not just from the marginal distributions. The population distributions of the fraction metabolized look similar to a mixture of the distributions in Figure 2. Interval estimates can be obtained as percentiles of these distributions. At high exposure, the 95% interval of the fraction metabolized in the population is [.52%, 4.1%]; at low exposure it is [15%, 58%]. Large variations also exist between

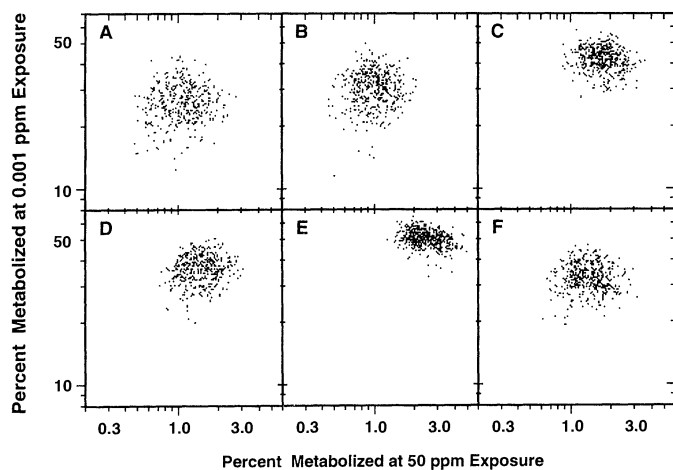


Figure 2. Estimates of the Fraction of PERC Metabolized per Day for a Continuous Inhalation Exposure to 50 ppm Versus the Estimated Fraction Metabolized at .001 ppm, for the Subjects of the Monster et al. (1979) Experiments. For each of the six plots, 500 random simulation draws are displayed. (From Bois et al. 1996. Reprinted with permission. Copyright 1996 Springer-Verlag GmbH & Co.)

individuals; for example, a factor of 2 difference is seen between similar subjects A and E in Figure 2.

Figure 3 shows the relation of fraction of PERC metabolized in 1 day (after 3 weeks continuous inhalation exposure) to exposure level. It also gives the range based on the population simulations. At low exposure levels, the fraction metabolized remains constant, because metabolism is linear. Saturation starts occurring above 1 ppm and is about complete at 10 ppm. At higher levels, the fraction metabolized decreases linearly with exposure, because the quantity metabolized per unit time is at its maximum.

When interpreting these results, one must remember that they are based on a single experiment. This study appears to be one of the best available; however, it includes only six subjects from an homogeneous population, measured at only two exposure levels. Much of the uncertainty associated with the results is due to these experimental limitations. Uncertainty could be reduced by additional analyses, but population variability, which in this study is approximately as large as uncertainty, could increase when a more heterogeneous group of subjects is included.

3.2.2 Estimates of Model Parameters. We can also directly look at the posterior distributions of the individual and population parameters. We transform from ψ_{kl} to θ_{kl} and summarize the distributions by the geometric mean and standard deviation of the posterior simulations of each of the KL components. Table 1 summarizes the distributions of the parameter values from the posterior simulations for the PERC example. Most important, the parameters still retain physiologically plausible values and are consistent with their prior distributions. The standard deviations (i.e., posterior uncertainties) in individual parameters are generally smaller than the prior uncertainties, showing that substantial information has been gained from the experimental data, even at high exposure levels.

Deviations from the prior distributions indicate the specific information brought by the in vivo experiment. Phys-

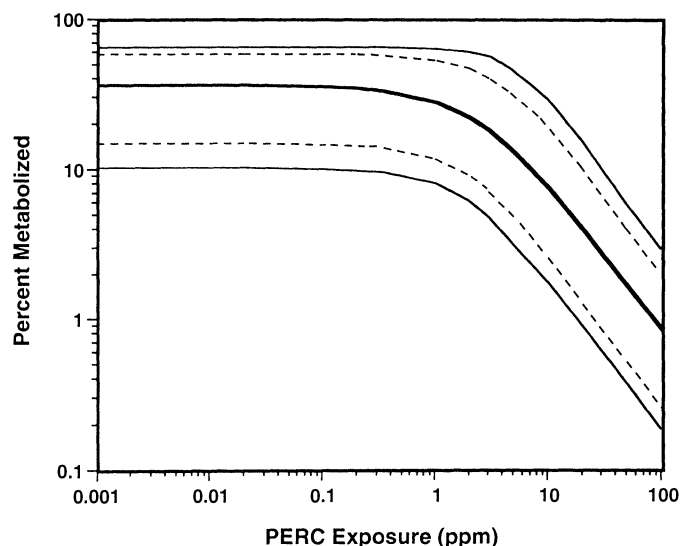


Figure 3. Estimated Fraction of PERC Metabolized as a Function of Inhalation Exposure Level. The thick line corresponds to the mean population model predictions. The thin lines bracket the 99% posterior interval; the dotted lines, the 95% interval for the population, based on the random simulations. (From Bois et al. 1996. Reprinted with permission. Copyright 1996 Springer-Verlag GmbH & Co.)

ical constraints on the mass balance of the compound led to a good identification of the scaling coefficients of the metabolic parameters. (Their posterior standard deviations correspond approximately to a factor of 1.5.) The estimated population mean of V_{ml}, the scaling coefficient of the maximum rate of metabolism, decreased by a factor of 20 from the prior to the posterior distribution. (This is within a reasonable range, considering that the prior geometric standard deviation was 10.) Therefore, the maximum rate of PERC metabolism in humans appears to be much lower than the values extrapolated from rodents on the basis of body mass to the power .7. Estimates of the population variation shows interindividual variation of about a factor of 2, which is not uncommon for metabolic parameters (see, e.g., Opdam 1989).

3.3 Evaluating the Fit and Sensitivity of the Model

In addition to their role in inference *given* the model, the posterior simulations also can be used in several ways to check the accuracy of the model and its sensitivity to prior assumptions.

Most directly, we can examine the errors of measurement and modeling by comparing observed data, y_{ijlt} , to their expectations, $f_i(\theta_l, \phi_l, E, t)$, for all the measurements, based on the posterior simulations of θ . The errors can be summarized quantitatively or used graphically to check for patterns or outliers. A perfect fit is not expected given the analytical measurement errors in the data (and given that the fit is measured by comparison to the posterior distribution, not the best-fit parameters). For the PERC example, the deviations are small (predicted and observed do not deviate by more than 65% of the observed value, with an average deviation of 11% in absolute value), with no apparent pattern (see Bois et al. 1996).

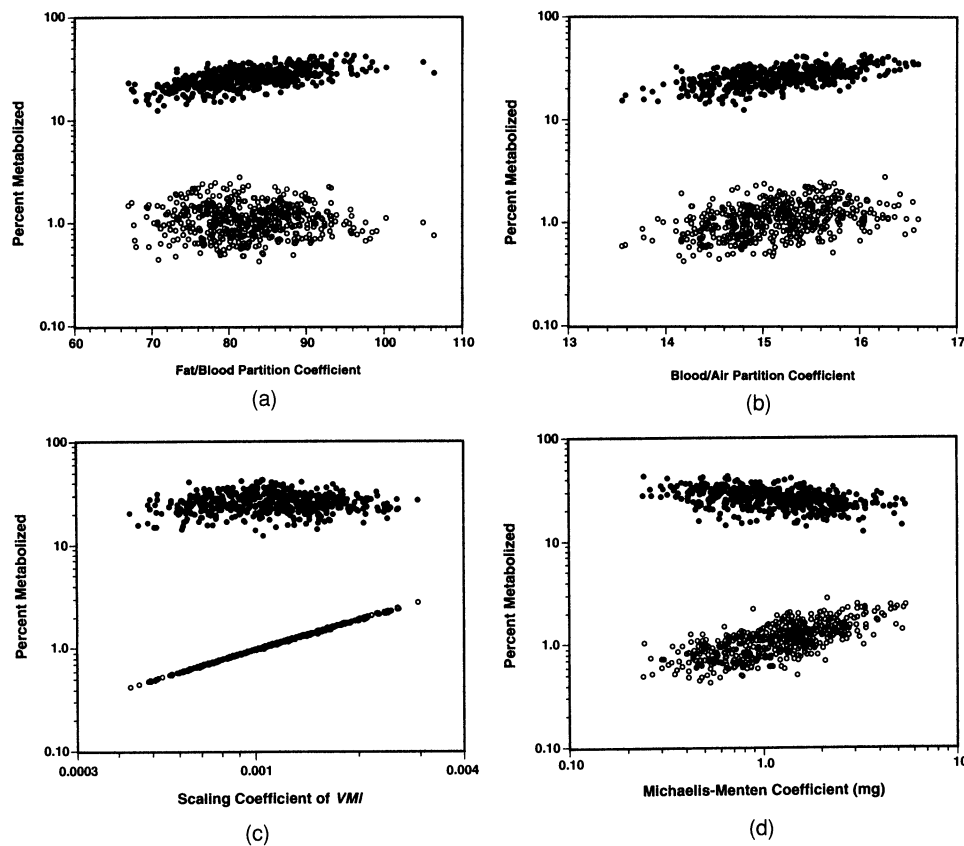


Figure 4. Posterior Simulation Draws of the Fraction Metabolized at High and Low Exposures (\bullet , .001 ppm exposure; \circ , 50 ppm exposure) for Subject A, Plotted Against Simulations of Four Parameters for Subject A: (a) the fat/blood partition coefficient (P_f), (b) the blood/air partition coefficient (P_{ba}), (c) the scaling coefficient of the maximum rate of metabolism (V_{MI}), and (d) the Michaelis–Menten coefficient (K_{MI}).

Another way to check the model is to compare the posterior distributions of parameters to their prior distributions, as we in fact did in the discussion of parameter estimates at the end of Section 3.2. Large shifts from prior to posterior (beyond what might be expected from the prior standard deviation) would suggest a flaw in the model or a misguided prior specification. The sensitivity of the model to the prior distribution can be assessed by comparing variation: If the posterior variation for a parameter is not much less than its prior variation, then the data have supplied little information about that parameter.

Another assumption that can be tested in this model is that of independence between the $L = 15$ individual-level parameters in the population distribution. We can test this assumption by examining the posterior correlations of the parameters, as follows. For each of the $L(L-1)/2$ pairs of parameters $i, j \in 1, \dots, L$ and for each posterior simulation draw, we computed the sample correlation of those parameters across the six subjects: $\hat{\rho}_{ij} = \text{corr}_k(\phi_{ki}, \phi_{kj})$. For each i, j , we then computed the mean and variance of the $\hat{\rho}_{ij}$ values across the 25,000 simulation draws, to yield an estimated population correlation and a standard error. Because the correlations were computed across only six subjects, the standard errors were by necessity large (mostly between .3 and .5). The estimated correlations for the pairs of parameters were all below .5 in absolute value, with the exception of the blood/liver partition coefficient, the ventilation over perfusion ratio (VPR), and the Michaelis–Menten co-

efficient (K_{MI}), all of which had estimated population correlations with each other at about .8 (with standard errors in the correlations of about .12). The high correlations do not invalidate our individual-level pharmacokinetic analysis, but they do imply that an improved model, including a population correlation between these parameters or possibly a reparameterization, would allow the between-subject information to be used more efficiently. To put it another way, the population part of the model is used to express uncertainty in predictions and borrow strength for inferences. By including the correlations in the population model, we would tend to obtain a smaller estimate for population variability and more precise inferences about individual-level parameters.

Another important concern in a Bayesian analysis is the sensitivity of posterior inferences to prior distributions. Consider a quantity of interest Q (typically some function of the model parameters θ ; e.g., the fraction metabolized under some specified exposure) and a particular parameter θ_{kl} . The dependence of Q on θ_{kl} can be assessed by the correlation of Q and θ_{kl} , the regression of Q on θ_{kl} , or, better still, a scatterplot of Q versus θ_{kl} in the posterior simulations. Inference for Q is sensitive to the assumed prior distribution for θ_{kl} if the two quantities are dependent in their posterior distribution and the posterior distribution for θ_{kl} is sensitive to its prior distribution. This sort of sensitivity analysis is also useful when designing future data collection; it is desirable to gather additional information about

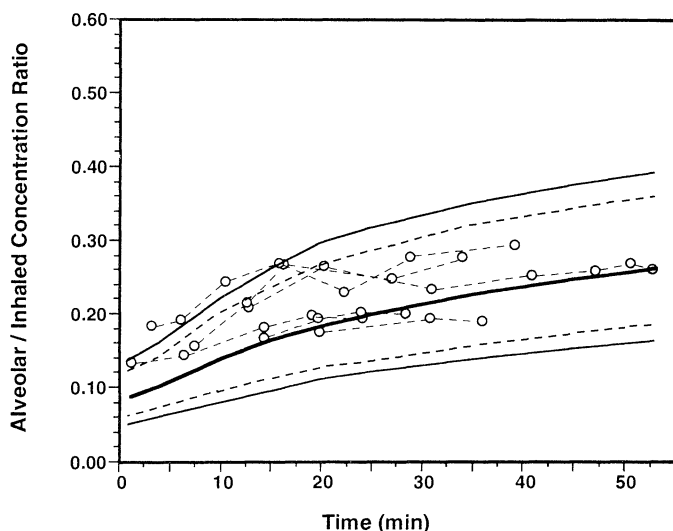


Figure 5. Model Simulation of the Opdam et al. (1986) Inhalation Experiment. Six subjects were exposed to PERC at levels ranging from .5 to 9 ppm. Alveolar/inhalation concentration ratios were reported (open circles). The thick line corresponds to the mean ratios predicted by our population model, adjusted with Monster et al. (1979) data. The solid and dotted lines correspond to 99% and 95% posterior simulation bounds.

the parameters that are highly correlated with summaries of interest. To illustrate, Figure 4 displays the relationship between the fraction metabolized at high dose or low dose and the most influential parameters for subject A (using that individual as an example). Many parameters affect the fraction metabolized at low dose, but VMI essentially determines the fraction metabolized at high dose, because of the saturation phenomenon.

Finally, a model can be checked by comparing its predictions to additional data not used in the original fit. To validate the model, we simulated another inhalation experiment on human volunteers (Opdam and Smolders 1986). Opdam and Smolders exposed six subjects to constant levels of PERC ranging from .5 to 9 ppm and followed alveolar concentration during exposure (up to 50 minutes). We assumed a 5 ppm exposure for every subject, because the published report does not permit matching of exposure levels to subjects. Assuming a level of .5 or 9 ppm changes the results slightly, given the nonlinearity of PERC metabolism indicated by the Monster et al. (1979) data, but not substantially given the noise in the data. The model is not linear, and thus different input levels do change its behavior, but the change in alveolar concentration of PERC appears very small when plotted with the data (i.e., the deviations between models and data dwarf the difference between models). These data were obtained in the same laboratory as those of Monster et al., and the population studied was similar. However, Opdam and Smolders's subjects are at least in part different from Monster et al. volunteers, as three women are included. We used the "normal breathing" data reported in figure 3 of Opdam and Smolders's report. Simulations were performed for the population using the posterior simulation draws of the vectors μ and Σ^2 . Figure 5 presents the data and the model prediction (with 95% and 99% simulation bounds). The model adjustment is good overall, even though exposure levels were 5 to 100 times

lower than those used in the Monster et al. experiments. However, short-term kinetics (i.e., less than 15 minutes after the onset of exposure) are not very well described by the model, which includes only a simple description of pulmonary exchanges. This is unlikely to seriously affect the quality of predictions for long-term, constant exposures.

4. DISCUSSION

Our results illustrate the use of Bayesian statistical techniques to bring together, through physiological modeling, in vivo data on human PERC toxicokinetics and physiological information on specific parameters. The method is very general, and we propose that it could be applied to any similar problem. In this section we discuss the methodological aspects of the work presented here and then address the results obtained and their importance.

4.1 Toxicokinetics and Risk Assessment

Physiologically based toxicokinetic models allow the simulation of various endpoints in specific target organs, while accounting for possible nonlinearities. Thus using these models is often advocated for more accurate risk assessment. One of the distinctive features of physiological modeling is to provide the opportunity to use relevant prior information on parameter values. The general tendency has been to assign specific values to some parameters and to adjust others to achieve a good—sometimes simply visual—agreement with experimental data. Considering the variability within the general population and the uncertainty about many parameters, which are difficult to measure accurately, it can be fundamentally misleading to fix input parameters or present results in the form of a single value (Louis 1991). In controlled human studies, heterogeneity is generally limited by the choice of healthy young males of particular ethnic group, as in the Monster et al. (1979) experiment. This may change with the recent National Institutes of Health requirement to include women and members of various ethnic groups in such studies (National Institutes of Health 1990). In the general population, heterogeneity may result in a wide range of responses. The ability to incorporate this information into risk assessments is still extremely limited.

These concerns should be foremost in risk assessments and instrumental in consideration for reducing population exposures. Much of the "conservatism" used in standard risk assessment procedures was developed to account for population variation. Substitution of a "conservative" methodology by a single modeled value based on a small subset of the human population could result in inadequate public health protection. For these reasons, all modeling results should include an assessment of all levels of uncertainty involved in the model structure and input parameters.

4.2 The Case of PERC

This work was initiated following a California Environmental Protection Agency workshop focusing on the scientific basis for establishing a carcinogenic unit risk value for PERC. A key issue in the unit risk development was the

value of the fraction of PERC metabolized at low exposure levels. This is an issue because of the absence of a total mass balance study of PERC metabolism at low exposure in humans. The estimated fraction metabolized in humans appeared to be dependent in part on the choice of data for analysis and in part on the assumptions made regarding unrecovered PERC in the experiments (Alexeef, Lewis, Zeise, Cox, and Schunk 1991). Thus an effort was made to critically evaluate the best study available and to incorporate as much useful information as possible.

The data of Monster et al. (1979) were thought to be the most scientifically rigorous developed, yet they appeared to suffer in their applicability to risk assessment due to the high exposure concentrations used (five orders of magnitude greater than ambient levels) and the narrow population evaluated (six healthy, young adult Dutch males). In addition to the data of Monster et al., it also was desirable to incorporate preliminary human in vitro information regarding PERC metabolism (Reitz 1992).

Defining prior distributions for the physiological parameters was difficult (Bois et al. 1996). Though it is well known that these parameters exhibit a wide range of inter- or intra-individual variability, the only values readily available (and those always used in physiological modeling) are "reference" values for young Caucasian males. Such reference values artificially reduce the population variance estimates. What is really needed is a database giving access to the population distributions of important physiological parameter values. Such a database would be usable for all types of physiological modeling and for both toxicants and drugs. Due to the current lack of information, we had to use "reference" values to the population means, and we gave reasonable guesses for population standard deviations and truncation limits.

By extrapolating to near-ambient environmental levels (approximately .001 ppm), after deriving the posterior parameter distributions, we estimate the fraction of PERC metabolized at low doses, after inhalation, as in the range [15%, 58%]. These percentages contrast sharply with those estimated from high exposure (50 ppm): a 95% posterior interval of [.5%, 4.1%]. This high-exposure estimate is in agreement with the Monster et al. (1979) figures for the recovery of inhaled PERC; in the experiments, unchanged PERC recovery was 80% to 100%, and approximately 2% of the inhaled dose was recovered in urine as trichloroacetic acid. However, risk assessments using a fraction metabolized calculated directly from the experiment, without considering exposure concentration, would likely underestimate the potential carcinogenic risk at low PERC exposure, by a factor of approximately 20. Hattis, White, Marmorstein, and Koch (1990) reviewed the literature on model-based estimates of the fraction of PERC metabolized at low dose (1 ppm). Previous estimates range from 2% to 86%. None of these were obtained by a complete statistical estimation procedure. Not surprisingly, the lowest estimates were obtained from models that assumed linear metabolism and were parameterized with high exposure data. Our model contains explicitly a nonlinear Michaelis-Menten term for

metabolism but is not constrained to behave nonlinearly. Had metabolism actually been linear for the Monster et al. volunteers, then our estimate of KMI would have been driven to the upper bound of its prior distribution, where the Michaelis-Menten term would behave linearly. This did not happen, and KMI stabilized around a value 2,000 times smaller than its a priori upper bound. Although the exposure levels were high in the experiments (72 and 144 ppm), the course of PERC concentrations in blood and exhaled air was followed with sufficient precision over an extended period and over a large range of tissue concentrations. This experimental design is sufficiently powerful to allow a reasonable identification of VMI and KMI values.

4.3 Statistical Issues

The approach presented here has five key features, all of which work in combination: (1) a physiological model, (2) a population model, (3) prior information on the population physiological parameters, (4) experimental data, and (5) Bayesian inference. If any of these five features are missing, then the model will not work: (1) without a physiological model, there is no good way to obtain prior information on the parameters; (2) without a population model, there generally are not enough data to estimate the model independently on each individual; (3 and 4) the parameters of a multicompartiment physiological model cannot be determined accurately by data or prior information alone; and (5) Bayesian inference yields a distribution of parameters consistent with both prior information and data, if such agreement is possible. Because it automatically includes both inferential uncertainty and population variability, the hierarchical Bayesian approach yields a posterior distribution that can be directly used for an uncertainty analysis of the risk assessment process.

Two other important components of our methodology are computation and model checking. Iterative simulation has been an increasingly useful tool in Bayesian analysis of complex models (see, e.g., Besag and Green 1993, Gilks et al. 1993, and Smith and Roberts 1993), and this example is no exception. Faster simulation algorithms enable one to fit more realistic models, which in turn inspire the development of faster computational algorithms, and so forth. Along with the fitting of a complicated model (and the potential for serious errors in modeling or computation) comes the responsibility to check that the resulting inference is consistent with both data and prior information. We perform these checks using the simulations of the parameters from their posterior distribution.

Wakefield (1996), in a discussion of studies with large numbers of subjects (in the hundreds), considered three kinds of generalizations to the sort of population model we have considered here: using individual-level covariates, modeling population correlations between individual parameters, and going beyond the normal distribution. In an experiment with individuals with measured covariates X_k , our population model can be naturally extended to a regression model of the form $E(\psi_{ki}) = X_k\beta_i$, thus replacing the parameters μ_i by vectors β_i . Correlations between param-

eters can be modeled by replacing the vector of variances, $\Sigma_1^2, \dots, \Sigma_L^2$, by an $L \times L$ matrix, Σ^2 . Finally, the normal distributions on ψ_{kl} can be replaced by wider-tailed distributions such as the t if necessary or modeled nonparametrically (see Davidian and Gallant 1993 and Davidian and Giltinan 1993). These generalizations are all important; correlation modeling is most useful for studies with large numbers of individuals, and using covariates or nonnormal distributions is most useful when fitting data to heterogeneous populations.

Finally, we note that different features of the model are especially useful in different contexts. The population model is important when analyzing data from a large group of subjects and can be embedded within a regression model to allow finer distinctions when studying the effects in the general population under nonlaboratory conditions. The physiological model is important if we wish to use the same model to study other compounds, as many of the parameters can be carried over directly to the new setting. Ultimately, this suggests the prospect of combining information from many laboratories in a program for general population pharmacokinetic modeling.

[Received February 1995. Revised May 1996.]

REFERENCES

- Acton, F. S. (1970), *Numerical Methods That Work*, New York: Harper and Row.
- Adolph, E. F. (1949), "Quantitative Relations in the Physiological Constitution of Mammals," *Science*, 109, 579–585.
- Alexeeff, G. V., Lewis, D. C., Zeise, L., Cox, D. N., and Schunk, H. H. (1991), "Health Effects of Tetrachloroethylene (PCE)," report, California Department of Health Services.
- Andersen, M. E. (1981), "A Physiologically Based Toxicokinetic Description of the Metabolism of Inhaled Gases and Vapors: Analysis at Steady State," *Toxicology and Applied Pharmacology*, 60, 509–526.
- Balant, L. P., and Gex-Fabry, M. (1990), "Physiological Pharmacokinetic Modelling," *Xenobiotica*, 20, 1241–1257.
- Besag, J., and Green, P. J. (1993), "Spatial Statistics and Bayesian Computation" (with discussion), *Journal of the Royal Statistical Society, Ser. B*, 55, 25–102.
- Bogen, K. T., and MacKone, T. E. (1988), "Linking Indoor Air and Pharmacokinetic Models to Assess Tetrachloroethylene Risk," *Risk Analysis*, 8, 509–520.
- Bois, F. Y., Gelman, A., Jiang, J., Maszle, D. R., Zeise, L., and Alexeeff, G. (1996), "Population Toxicokinetics of Tetrachloroethylene," *Archives of Toxicology*, to appear.
- Bois, F. Y., Zeise, L., and Tozer, T. N. (1990), "Precision and Sensitivity Analysis of Pharmacokinetic Models for Cancer Risk Assessment: Tetrachloroethylene in Mice, Rats and Humans," *Toxicology and Applied Pharmacology*, 102, 300–315.
- Box, G. E. P., and Tiao, G. C. (1973), *Bayesian Inference in Statistical Analysis*, New York: Wiley.
- Davidian, M., and Gallant, A. R. (1992), "Smooth Nonparametric Maximum Likelihood Estimation for Population Pharmacokinetics, With Application to Quinidine," *Journal of Pharmacokinetics and Biopharmaceutics*, 20, 529–556.
- Davidian, M., and Giltinan, D. M. (1993), "Some Simple Methods for Estimating Intraindividual Variability in Nonlinear Mixed Effects Models," *Biometrics*, 49, 59–73.
- Fernandez, J. G., Guberman, E., and Caperos, J. (1976), "Experimental Human Exposures to Tetrachloroethylene Vapor and Elimination in Breath After Inhalation," *American Industrial Hygiene Association Journal*, 37, 143–150.
- Gear, C. W. (1971a), "Algorithm 407—DIFSUB for Solution of Ordinary Differential Equations [D2]," *Communications of the Association for Computing Machinery*, 14, 185–190.
- (1971b), "The Automatic Integration of Ordinary Differential Equations," *Communications of the Association for Computing Machinery*, 14, 176–179.
- Gelfand, A. E., Hills, S. E., Racine-Poon, A., and Smith, A. F. M. (1990), "Illustration of Bayesian Inference in Normal Data Models Using Gibbs Sampling," *Journal of the American Statistical Association*, 85, 398–409.
- Gelfand, A. E., and Smith, A. F. M. (1990), "Sampling-Based Approaches to Calculating Marginal Densities," *Journal of the American Statistical Association*, 85, 398–409.
- Gelfand, A. E., Smith, A. F. M., and Lee, T.-M. (1992), "Bayesian Analysis of Constrained Parameter and Truncated Data Problems Using Gibbs Sampling," *Journal of the American Statistical Association*, 87, 523–532.
- Gelman, A. (1995), "Method of Moments Using Monte Carlo Simulations," *Journal of Computational and Graphical Statistics*, 3, 36–54.
- Gelman, A., Carlin, J. B., Stern, H. S., and Rubin, D. B. (1995), *Bayesian Data Analysis*, New York: Chapman and Hall.
- Gelman, A., Roberts, G., and Gilks, W. (1995), "Efficient Metropolis Jumping Rules," in *Bayesian Statistics 5*, eds. J. M. Bernardo et al., New York: Oxford University Press.
- Gelman, A., and Rubin, D. B. (1992), "Inference From Iterative Simulation Using Multiple Sequences" (with discussion), *Statistical Science*, 7, 457–511.
- Gerlowski, L. E., and Jain, R. K. (1983), "Physiologically Based Pharmacokinetic Modeling: Principles and Applications," *Journal of Pharmaceutical Sciences*, 72, 1103–1127.
- Gilks, W. R., Clayton, D. G., Spiegelhalter, D. J., Best, N. G., McNeil, A. J., Sharples, L. D., and Kirby, A. J. (1993), "Modelling Complexity: Applications of Gibbs Sampling in Medicine" (with discussion), *Journal of the Royal Statistical Society, Ser. B*, 55, 39–102.
- Guberman, E., and Fernandez, J. (1974), "Control of Industrial Exposure to Tetrachloroethylene by Measuring Alveolar Concentrations: Theoretical Approach Using a Mathematical Model," *British Journal of Industrial Medicine*, 31, 159–167.
- Hattis, D., White, P., Marmorstein, L., and Koch, P. (1990), "Uncertainties in Pharmacokinetic Modeling for Perchloroethylene. I. Comparison of Model Structure, Parameters, and Predictions for Low-Dose Metabolism Rates for Models Derived by Different Authors," *Risk Analysis*, 10, 449–457.
- ICRP (1975), *Report of the Task Group on Reference Man—A Report Prepared by a Task Group of Committee 2 of the International Commission on Radiological Protection*, Oxford, U.K.: Pergamon.
- Ikeda, M., Ohtsuki, H., Imamura, T., and Komoike, Y. (1972), "Urinary Excretion of Total Trichloro-Compounds, Trichloroethanol, and Trichloroacetic Acid as a Measure of Exposure to Trichloroethylene and Tetrachloroethylene," *British Journal of Industrial Medicine*, 29, 328–333.
- Koizumi, A. (1989), "Potential of Physiologically Based Pharmacokinetics to Amalgamate Kinetic Data of Trichloroethylene and Tetrachloroethylene Obtained in Rats and Man," *British Journal of Industrial Medicine*, 46, 239–249.
- Leung, H.-W. (1991), "Development and Utilization of Physiologically Based Pharmacokinetic Models for Toxicological Applications," *Journal of Toxicology and Environmental Health*, 32, 247–267.
- Louis, T. A. (1991), "Assessing, Accommodating, and Interpreting the Influences of Heterogeneity," *Environmental Health Perspectives*, 90, 215–222.
- Menzel, D. B. (1987), "Physiological Pharmacokinetic Modeling," *Environmental Science and Technology*, 21, 944–950.
- Michaelis, L., and Menten, M. L. (1913), "Die Kinetik der Invertinwirkung," *Biochemische Zeitschrift*, 49, 333–389.
- Monster, A. C., Boersma, G., and Steenweg, H. (1979), "Kinetics of Tetrachloroethylene in Volunteers; Influence of Exposure Concentration and Work Load," *International Archives of Occupational and Environmental Health*, 42, 303–309.
- National Institutes of Health (1990), *NIH Guide for Grants and Contracts*, Washington, DC: U.S. Department of Health and Human Services.
- Ohtsuki, T., Sato, K., Koizumi, A., Kumai, M., and Ikeda, M. (1983), "Limited Capacity of Humans to Metabolize Tetrachloroethylene," *International*

- tional Archives of Occupational and Environmental Health*, 51, 381–390.
- Opdam, J. J. G. (1989), "Intra- and Interindividual Variability in the Kinetics of a Poorly and Highly Metabolising Solvent," *British Journal of Industrial Medicine*, 46, 831–845.
- Opdam, J. J. G., and Smolders, J. F. J. (1986), "Alveolar Sampling and Fast Kinetics of Tetrachloroethylene in Man. 1—Alveolar Sampling," *British Journal of Industrial Medicine*, 43, 814–824.
- (1987), "Alveolar Sampling and Fast Kinetics of Tetrachloroethylene in Man. 2—Fast Kinetics," *British Journal of Industrial Medicine*, 44, 26–34.
- Racine-Poon, A., and Smith, A. F. (1990), "Population Models," in *Statistical Methodology in the Pharmaceutical Sciences*, ed. D. A. Berry, New York: Marcel Dekker, pp. 139–162.
- Reitz, R. H. (1992), "New Data on Tetrachloroethylene Metabolism," paper presented at an OHEEA Advisory Panel Meeting, Cal-EPA, Berkeley.
- Ritschel, W. A., and Banerjee, P. S. (1986), "Physiological Pharmacokinetic Models: Principles, Applications, Limitations and Outlook," *Methods and Findings in Experimental and Clinical Pharmacology*, 8, 603–614.
- Rowland, M., and Tozer, T. N. (1989), *Clinical Pharmacokinetics—Concepts and Applications* (2nd ed.), Philadelphia: Lea and Febiger.
- Sheiner, L. B. (1984), "The Population Approach to Pharmacokinetic Data Analysis: Rationale and Standard Data Analysis Methods," *Drug Metabolism Reviews*, 15, 153–171.
- Sheiner, L. B., Rosenberg, B., and Melmon, K. L. (1972), "Modeling of Individual Pharmacokinetics for Computer-Aided Drug Dosage," *Computers and Biomedical Research*, 5, 441–459.
- Smith, A. F. M. (1991), "Bayesian Computational Methods," *Philosophical Transactions of the Royal Society of London*, Ser. A, 337, 369–386.
- Smith, A. F. M., and Roberts, G. O. (1993), "Bayesian Computation via the Gibbs Sampler and Related Markov Chain Monte Carlo Methods" (with discussion), *Journal of the Royal Statistical Society*, Ser. B, 55, 3–102.
- Tanner, M. A. (1993), *Tools for Statistical Inference: Methods for the Exploration of Posterior Distributions and Likelihood Functions* (2nd ed.), New York: Springer-Verlag.
- Wakefield, J. C. (1996), "The Bayesian Analysis of Population Pharmacokinetic Models," *Journal of the American Statistical Association*, 91, 62–75.
- Wakefield, J. C., Smith, A. F. M., Racine-Poon, A., and Gelfand, A. E. (1994), "Bayesian Analysis of Linear and Non-Linear Population Models Using the Gibbs Sampler," *Applied Statistics*, 43, 201–221.
- Ward, R. C., Travis, C. C., Hetrick, D. M., Andersen, M. E., and Gargas, M. L. (1988), "Pharmacokinetics of Tetrachloroethylene," *Toxicology and Applied Pharmacology*, 93, 108–117.
- Woodruff, T. J., and Bois, F. Y. (1993), "Optimisation Issues in Physiological Toxicokinetic Modeling: A Case Study With Benzene," *Toxicology Letters*, 69, 181–196.
- Yuh, L., Beal, S., Davidian, M., Harrison, F., Hester, A., Kowalski, K., Vonesh, E., and Wolfinger, R. (1994), "Population Pharmacokinetic/Pharmacodynamic Methodology and Applications: A Bibliography," *Biometrics*, 50, 566–575.