

Statistical issues in toxicokinetic modeling: a Bayesian perspective

Pascale Bernillon, Frédéric Y. Bois

▶ **To cite this version:**

Pascale Bernillon, Frédéric Y. Bois. Statistical issues in toxicokinetic modeling: a Bayesian perspective. Environmental Health Perspectives, 2000, 108 (suppl. 5), pp.883-893. ineris-00961852

HAL Id: ineris-00961852
https://ineris.hal.science/ineris-00961852

Submitted on 20 Mar 2014

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Statistical Issues in Toxicokinetic Modeling: A Bayesian Perspective

Pascale Bernillon[1] and Frédéric Y. Bois[2]

[1]B3E INSERM U444, Paris, France; [2]Institut National de l'Environnement Industriel et des Risques (INERIS), Verneuil en Halatte, France

Determining the relationship between an exposure and the resulting target tissue dose is a critical issue encountered in quantitative risk assessment (QRA). Classical or physiologically based toxicokinetic (PBTK) models can be useful in performing that task. Interest in using these models to improve extrapolations between species, routes, and exposure levels in QRA has therefore grown considerably in recent years. In parallel, PBTK models have become increasingly sophisticated. However, development of a strong statistical foundation to support PBTK model calibration and use has received little attention. There is a critical need for methods that address the uncertainties inherent in toxicokinetic data and the variability in the human populations for which risk predictions are made and to take advantage of a priori information on parameters during the calibration process. Natural solutions to these problems can be found in a Bayesian statistical framework with the help of computational techniques such as Markov chain Monte Carlo methods. Within such a framework, we have developed an approach to toxicokinetic modeling that can be applied to heterogeneous human or animal populations. This approach also expands the possibilities for uncertainty analysis. We present a review of these efforts and other developments in these areas. Appropriate statistical treatment of uncertainty and variability within the modeling process will increase confidence in model results and ultimately contribute to an improved scientific basis for the estimation of occupational and environmental health risks. *Key words*: Bayesian analysis, hierarchical models, MCMC methods, toxicokinetic models, uncertainty, variability. — *Environ Health Perspect* 108(suppl 5):883–893 (2000).
http://ehpnet1.niehs.nih.gov/docs/2000/suppl-5/883-893bernillon/abstract.html

Exposure concentrations of chemicals in air, food, water, soil, dust, or other media with which populations are in contact are the dose scales most commonly used when assessing health effects for environmental or occupational pollutants. Yet, external exposure is often only a rough estimate for internal exposure delivered at the critical target in the body. The latter is a more appropriate measure of dose for mechanism-based risk assessments (*1*). For example, many chemicals require metabolic conversion into chemically active species before exhibiting toxicity, and this conversion may be subject to saturation at high doses (*2–5*). In the presence of saturable activation or detoxification pathways, the relationship between administered (external) dose and delivered (internal) dose may be nonlinear. This nonlinearity is of particular concern when performing high-to-low-dose extrapolation of the overall exposure–response relationship. Indeed, a solution is to measure internal biomarkers of exposure, but such direct measurements may not be feasible if the relevant biomarkers are not yet developed. In this case, modeling approaches can be advocated as an alternative. Toxicokinetic (TK) models, for example, are useful tools to relate external exposures to internal measures of dose. TK models describe the behavior of chemicals in the body, e.g. the processes of absorption, distribution, metabolism, and elimination. Two main classes of TK models have been developed: classical toxicokinetic models and physiologically based toxicokinetic (PBTK) models. The development of PBTK models has been particularly active during the last 20 years, and their usefulness for chemical risk assessments is now firmly established (*5–10*). Several major chemical risk assessments currently processed by the U.S. Environmental Protection Agency involve the use of PBTK models.

The focus of PBTK modeling in risk assessment to date has been essentially on improving the scientific basis for extrapolating risk from animals to humans (*11*). Before the introduction of PBTK models, the key default assumption regarding interspecies extrapolation was that animal exposures could be converted to equivalent human exposures by a surface area correction factor. However, several examples can be found in the literature in which comparative modeling of distribution and metabolism in animals and humans has challenged the use of simple surface area conversions for interspecies extrapolations (*12–15*). The ability to tailor the parameter values of PBTK models to a particular animal species offers, in theory, a firm basis for interspecies extrapolation. In practice, for most PBTK models some parameters remain difficult to extrapolate across species. For metabolic parameters, in particular, species-specific values may be difficult to obtain experimentally. The "parallelogram approach" (*16,17*) may provide partial answers to this question, but parameters typically still need to be adjusted through calibration of the model with pharmacokinetic data (usually measured time courses of parent compound and metabolites in biological media such as exhaled air, blood, urine, etc.).

If some uneasiness has been expressed about the use of PBTK models (*18*), it was essentially because of the lack of statistical methods for calibrating them, i.e., for adjusting their parameter values by taking toxicokinetic data into account. Another issue gaining major attention in the scientific and regulatory communities is population variability in toxicokinetic and metabolic processes (*9,19–27*). Considering this variability and the uncertainty about many parameters difficult to measure accurately, using fixed parameter values or presenting results in the form of point estimates can be fundamentally misleading (*28*). Obviously, the task of a correct statistical treatment of PBTK models is daunting, as we are faced with large nonlinear models, small data sets, high uncertainty, and biological variability. Until recently, there was no method for rigourous statistical validation of PBTK models. For example, to our knowledge, the predictions made by such models were never presented with meaningful confidence levels. Consequently, it was impossible to decide whether the fits were acceptable, the models reasonable, and what confidence to place in the results of extrapolation (*29*).

In response to this challenge, our group began to examine the statistical issues associated with the calibration and use of PBTK models and to develop methods to solve them. This article first briefly presents the structure of classical TK and PBTK models. We then discuss the issues of parameterization and calibration of these models and describe a Bayesian approach to their statistical analysis.

This article is part of the monograph on Mathematical Modeling in Environmental Health Studies.

Address correspondence to F.Y. Bois, INERIS, Parc ALATA, BP 2, 60550 Verneuil en Hallate, France. Telephone: 33 03 44 55 65 96. Fax: 33 03 44 55 68 99. E-mail: bois@b3e.jussieu.fr

Funding for this research was from Association pour le Dévelopement de l'Informatique Médicale et l'Epidémiologie Clinique and Institut National de l'Environnement Industriel et des Risques (Convention Industrielle de Formation par la Recherche on Enterprise # 96044).

Received 11 February 2000; accepted 8 August 2000.

Toxicokinetic Models

A variety of TK models have been developed; all are simplified representations of chemical disposition within the human/animal organism. This section describes the structure and characteristics of classical compartmental and physiologically based models, which are the main two classes of TK models found in the literature. We then briefly present minimal physiological models, which recently have been proposed by several investigators.

Classical Compartmental (TK) Models

In classical TK models, the body is represented by several connected compartments (rarely more than three); each compartment is a virtual space (i.e., compartments do not necessarily reflect the anatomy of the species of interest) within which the chemical is assumed to be homogeneously distributed (*30*). Chemical transfers between compartments are described by a set of differential equations. Typical parameters are compartment volumes, exchange rates between them, and clearance (elimination rate). The number of compartments and the parameter values are inferred from fitting the model to TK data. A procedure commonly adopted when using these models consists in first fitting a simple, usually one-compartment model. If the fit is unsatisfactory, another compartment is added to the model, and so on, until an acceptable fit is achieved. Parameter values derived from the analysis of classical compartments models are data dependent and are not intended to represent actual physiological volumes or flows in the organism (*11*). For example, the volume of distribution of a particular compound can be much greater than the volume of the body (this is the case for chemicals with high affinities for tissue proteins) (*31*). These models thus are commonly referred as empirical or data-based models.

Classical compartmental models are reliable tools to predict various surrogates of dose such as areas under the concentration–time curve or the maximum concentration reached in diverse biological media (e.g., exhaled air, venous blood, urine) when the objective is to interpolate from the current data. They are widely used in pharmacokinetic studies to investigate drug disposition in the body. Nevertheless, given their lack of physiological relevance, these models are not indicated to extrapolate kinetic results between species, exposure routes, or exposure conditions. For these reasons, other modeling approaches, including physiologically-based models, have been investigated.

PBTK Models

In PBTK models, the body is subdivided into a series of anatomical or physiological compartments that correspond to specific organs (liver, kidney, lung) or lumped tissue and organ groups (fat, richly perfused, and slowly perfused tissues) (*32–34*). Tissues or organs are lumped when they have, for example, similar blood flow and fat content and when partitioning of the substance of interest between them can be considered homogeneous. Connections between compartments represent the blood or lymphatic circulation and chemical transfers between compartments are described by mass balance differential equations. The time course of transport and transformations of a chemical through the various compartments can be simulated via the resolution of the model's set of equations for any given set of parameter values. PBTK models include three types of parameters: physiological parameters such as breathing rate, blood flows, and tissue volumes; physicochemical parameters, such as partition coefficients that represent the relative solubility of a chemical in specific tissue; and biochemical parameters describing, for example, metabolic processes. The number of compartments to be included in the model depends on the objective of the study and on the possible mode(s) of action and site(s) of toxicity of the chemical studied. In most cases, the kidney or the liver are usually individualized when they are major organs of elimination and/or metabolism. Other organs are also individualized (e.g. bone marrow, brain, bone) when they are known sites of metabolism, storage, or toxicity. In many applications, distribution among compartments is supposed to be limited by perfusion; once in a compartment, the chemical distributes instantaneously and homogeneously throughout the entire volume of the compartment. More complex models describing the diffusion of the compound in subcompartments representing vascular, interstitial, and intracellular spaces (diffusion-limited models) have also been developed (*35–37*).

Figure 1 illustrates a general four-compartment PBTK model for volatile chemicals, which we used to study tetrachloroethylene (TETRA) kinetics in humans (*38,39*). In this model, TETRA distribution among the four compartments (well-perfused tissues, poorly perfused tissues, fat, liver) is supposed to be limited by blood perfusion, as is often the case for apolar solvents. TETRA absorption is supposed to occur by inhalation; elimination takes place by exhalation and metabolism. The only metabolically active compartment included in the model is the liver. TETRA primary metabolism is described by a saturable Michaelis-Menten mechanism. This model encloses 15 parameters.

Toxic effects induced by exposure to chemicals are commonly observed during experimental animal studies under conditions of exposure different than typical human occupational or environmental exposures. PBTK models are promising tools to solve some of the critical extrapolation issues encountered in chemical risk assessments (*6,11,32,33,40–42*):

- for example, dose extrapolation can automatically be achieved with a PBTK model assuming the model correctly captures the linear and nonlinear dynamics involved in the transport and metabolism of the compound studied.
- Interspecies extrapolation is resolved by assuming that the model structure is correct for two or more species (the TETRA model described above, for example, is reasonable for any mammal). Simply changing parameters to values specific to the species of interest operates the extrapolation.
- Interroute (of exposure) extrapolations can be performed. The model presented above (Figure 1) has only one route of entry, the lung; it is possible to simulate intravenous injection by imposing increase in venous blood concentration at any time or by an explicit additional equation; ingestion has been modeled with such models as a direct infusion into the liver compartment or with additional compartments describing the gastrointestinal tract; dermal absorption has also been modeled by addition of a skin compartment (*43,44*).

PBTK models can explain the toxicokinetic behavior of the compound studied under special conditions in humans such as physical activity (*45*) or in specific targeted populations (children, aged, obese) by simply changing the values of some physiological parameters. In addition, these models are

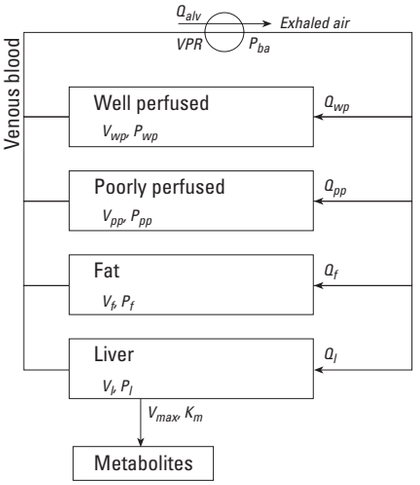


Figure 1. Schematic representation of a PBTK model used for distribution and metabolism of TETRA (*38*). *Q*, blood flows; *V*, compartment volumes; *P*, partition coefficients; *VPR*, ventilation over perfusion ratio; V_{max}, maximum rate of metabolism; K_m, Michaelis-Menten parameter.

attractive when modeling complex situations such as developmental toxicity (*36,46–52*); unusual exposure scenarios such as delayed exposure of a nursing infant due to the mother's inhalation exposure can also be addressed (*53,54*).

The above properties of PBTK models contribute to the development of more accurate estimates of risk. PBTK modeling techniques are now well developed. Nevertheless, such models involve a large number of parameters (typically more than 20), each of which is subject to some degree of uncertainty and variability, which are translated into PBTK model outputs. It is thus necessary to consider the statistical issues raised by such complexity.

Minimal Physiological Models

Minimal (or reduced) physiological models are PBTK models in which some of the compartments (e.g., poorly and well-perfused tissues) have been lumped into a single central compartment. Reducing the number of compartments is a way to simplify the calibration process while partly preserving the model's physiological character. In the case of benzene, it has been shown that a reduced three-compartment PBPK model can be a good alternative to a five-compartment PBPK model (*23*).

Figure 2 presents a minimal physiological TK model of TETRA kinetics in the human. TETRA kinetics in blood and exhaled air are described by a three-compartment TK model linked to a one-compartment submodel for the main TETRA metabolite, trichloroacetic acid (TCA) (*55*). In the model, TETRA enters the body by pulmonary exchange, modeled by first-order transfers between air and the central compartment. TETRA venous blood concentrations are assumed to be predicted by the central compartment concentrations. The second and third compartments are a priori supposed to represent fat and muscles, respectively. Metabolism of TETRA is assumed to take place in the central compartment and is described by a linear process. A fraction of metabolized TETRA leads to TCA; the rest is made of the other TETRA metabolites. TCA is formed from TETRA metabolism but also from endogenous background reactions. Urinary elimination of TCA is assumed to follow first-order kinetics.

Reduced PBTK models may be promising tools to analyze data when many subjects are involved. Recently, Fanning (*56*) used a reduced PBTK model to analyze benzene kinetics data from a cohort of 60 Chinese workers. In that model, the human body was divided into three compartments: a central compartment, fat, and liver. Thomaseth and Salvan (*57*) consider a minimal physiological model to describe long-term kinetics of tetrachlorodibenzo-*p*-dioxin in human tissues using data from a cohort of 359 subjects.

These examples simply illustrate that, as with any modeling exercise, the development of a toxicokinetic model must be guided primarily by the question raised and the data available to deal with it.

Statistical Considerations

PBTK models are considered an effective means to predict target tissue dose, in particular because of the large amount of a priori information available in the literature on their parameters. Nevertheless, calibration of TK models, even when they are physiologically based, is hardly feasible using only a priori information. We mean by calibration the allocation of appropriate values to model parameters to enable that model to behave like the system it represents. In fact, even though reference values can be found in the literature for physiological parameters such as organ volumes or blood flows, experience has shown that the calibration of virtually all PBTK models requires estimation of some parameters (in particular those describing metabolism) using experimental TK data. Such data typically consist of measured parent and/or metabolite amounts or concentrations from various sites in the body (e.g., blood, exhaled air, urine) at various times following the beginning of a controlled exposure to the chemical studied.

We address, in the following subsections, the parameterization of TK models, their calibration, and finally, the predictive use of these models.

Parameterization

It is first necessary to pay careful attention to model parameterization (i.e., to definition of the model parameters) to facilitate their subsequent estimation. Many parameters of a toxicokinetic model (e.g., cardiac output, organ volumes, maximum rate of metabolism) tend to be correlated to some measure of body size. Furthermore, some parameters are correlated to each other (e.g., organ blood flows with cardiac output). Neglecting those correlations might lead to unrealistic physiological properties of the model, particularly when a distribution of behaviors is computed (e.g., through Monte Carlo simulations). It is traditional to implement an a priori deterministic modeling of known physiological dependencies between parameters by using scaling functions and allometric (related-to-size) relationships (*58–63*). For example, volumes can be parameterized as fractions of the lean body weight, flows as fractions of cardiac output, and the maximum rate of metabolism in the liver as a power function of lean body weight. Not all parameters need to be scaled. Unfortunately, there is no unique way to implement scaling, and there does not appear to be a rigorous foundation for any particular choice so far in the literature. We assert that covariate modeling should be guided by statistical considerations, as it is essentially a regression problem, but this is a research topic still to be explored.

Another problem that can be solved through proper parameterization is parameter identifiability. Identifiability characterizes parameters for which reasonable estimates can be obtained given experimental data (*64–66*). Parameter identifiability is a critical issue in compartmental kinetic modeling (*67,68*) and should be examined before any attempt to calibrate the model. Two types of identifiability can be distinguished: structural and statistical. The first is concerned with the formal model's equations and parameters, i.e. the structure of the model. As an example, consider the following model equation:

$$y = \alpha \times \beta \times x, \qquad [1]$$

where α and β are unknown parameters to estimate from observations of x and y. If neither α nor β appears separately in another model-defining equation, α and β will not be separately identifiable whatever the data used to estimate them. Each one can take an infinite

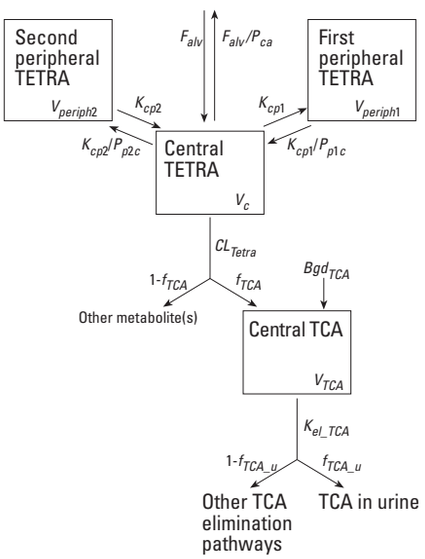


Figure 2. Schematic representation of a linear three-compartment model used for distribution and metabolism of TETRA and distribution of TCA by Bernillon et al. (*55*). The model considers 14 parameters; symbols are: F_{alv}, alveolar flow; V, volumes; K, rates of distribution from central to peripheral compartments; P, partition coefficients; CL_{Tetra}, TETRA metabolic clearance; Bgd_{TCA}, endogenous TCA formation rate; Kel_{TCA}, TCA elimination rate; f_{TCA}, fraction of metabolized TETRA leading to TCA; f_{TCA_u}, fraction of total eliminated TCA excreted into urine.

and unbounded number of values. Only their product $\alpha \times \beta$ can be reasonably estimated. α and β can become identifiable if an informative prior has been specified for one of them in a Bayesian context. Structural identifiability can also be ensured by an adequate parameterization. Structural nonidentifiability was encountered when calibrating the minimal PBTK model for TETRA kinetics depicted in Figure 2, with the two couples of parameters (P_{p1c}, $V_{periph1}$) and (P_{p2}, $V_{periph2}$). Their products, $P_{p1c} \times V_{periph1}$ and $P_{p2c} \times V_{periph2}$, were thus the parameters actually estimated. Putting separate priors on P_{p1c}, $V_{periph1}$, P_{p2c}, and $V_{periph2}$ was another option we could have taken. The second type of identifiability, called data-conditioned, or statistical, characterizes parameters that can be precisely estimated. For a parameter estimate, precision is the inverse of uncertainty, and we treat this question extensively in the next section. In essence, statistical identifiability is ensured by either adequate data to estimate all model parameters or, in a Bayesian context, sufficient prior information to constrain those parameters (*69*). In a TK context, for example, statistical nonidentifiability may be encountered when experimental data do not include observations during or close to the absorption phase. In that case, the parameters characterizing absorption will be poorly estimated given that data set. In a way, statistical identifiability is a milder form of structural identifiability that does not require reparameterization to be treated. Reparameterization can still be useful in case of statistical identifiability, as it can hasten computations during the calibration process.

Calibration, Uncertainty, and Variability

Uncertainty and variability in the TK parameter estimates affect potentially all TK model predictions (e.g., estimates of tissue doses); it is necessary to assess both to make predictions useful for risk analysts and decisionmakers. This helps to determin the confidence to place in those predictions, to assess the range of internal dose likely to occur within the population for which the risk is being evaluated, and finally, to weight the decisions on the basis of TK model predictions (*21,26*).

It is important to distinguish between uncertainty and variability. Variability typically refers to differences in the values of model parameters among individuals (interindividual variability) or across time within a given individual (intraindividual variability). Variability may stem from genetic differences, lifestyles, physiological status, age, etc. (*26*). Uncertainty, on the other hand, essentially is a result of lack of knowledge (*70*) and may have various sources. TK parameters are known only with finite precision. The use of standard values, such as those in the International Commission on Radiological Protection report (*71*), tend to give a false impression of precision for physiological parameter values (and thus for model predictions). At best, such standard or default values are approximate values of the average for a human population. There always will be uncertainty about their true value for a particular group of animals or humans, and even more for a particular individual exposed (from which TK data may be available). In addition, most chemical-specific parameters tend to be imprecise; i.e., they may have been measured *in vitro* rather than *in vivo* or they may be accessible only after fitting a model to TK data. Measurement errors on the experimental data, sparseness of those, and model simplifications or mispecifications are translated into uncertainties in the parameter values determined by statistical fitting to such data. Uncertainty may be reduced by further experiments, the design of which can be formally optimized (*72,73*), or by a better understanding of the actual processes under study.

Variability is inherent in animal and human populations and cannot be reduced. Both variability and uncertainty are usually present in TK data. For example, each TK data set is specific to the subjects studied and extrapolating quantitative information from such small groups to larger and different populations requires careful consideration. In fact, the primary interests of TK analyses rarely lie in the toxicokinetics of the compound studied in any particular individual but resides instead in inferences both about average TK behavior of substances in larger groups (sensitive subpopulations, age groups, whole-country population, etc.) and quantification of the variability in such behavior. It is necessary to take variability explicitely into account within the calibration process. Data pooling (i.e., treating averaged data values as if they characterized a typical member of the population) should be avoided, as it can gravely distort the data (*74,75*).

Given the usual sparseness of TK data and the number of parameters involved in a PBTK model, conventional (e.g., unconstrained least square) parameter estimation is generally not feasible. From a statistical point of view, manually adjusting several parameters is not scientifically acceptable. Neither is the commonly used alternative, which consists of statistically fitting only a subset of the model parameters (assuming that the others are exactly known and set to predefined reference values). It is difficult to arrive at a correct assessment of the uncertainty in model predictions through such an approach. Ignoring the uncertainty in a parameter value implicitly ignores covariances between this particular parameter and all others. This leads to distortions in the entire covariance structure of parameter estimates (because covariance between two parameters vanishes when one is fixed) and finally, to an underestimation of the uncertainty in model outputs (*76*). As an illustration, consider a simple two-dimensional case (Figure 3). Assume that both parameter estimates, given the data θ_1, and θ_2, are marginally (i.e., when considered alone) normally distributed (Figure 3A, C) and that they are highly correlated. The ellipse (Figure 3B) that outlines the 95% joint confidence region for θ_1 and θ_2 illustrates that covariance. A correct joint estimation of θ_1 and θ_2 should give both marginal and joint distributions closely approximating those. However, setting θ_1, for convenience, to a fixed value, *a*, modifies the structure of covariance between θ_1 and θ_2, transforming the initial ellipse into a segment of line, *s*. The 95% confidence interval for θ_2, given $\theta_1 = a$, now ranges from *b* to *c* (panel C). The resulting distribution of θ_2 estimates is much narrower than the correct distribution and its mean may not even be the same, which could later lead to biases and underdispersion in predictions. In higher (imagine 20) dimensions, the situation would be considerably worse.

Three points can be made here: First, fixing the value of some parameters while adjusting the other parameters is not satisfactory unless there is no correlation between the fixed and other parameters. Proving the absence of such correlations is difficult and amounts to first estimating all parameters together (that is, considering all of them uncertain). It is simpler and often more judicious to stop at this first step and acknowledge the complete covariance between all parameters. Second, fixing θ_1 for estimating θ_2 and then relaxing this constraint when

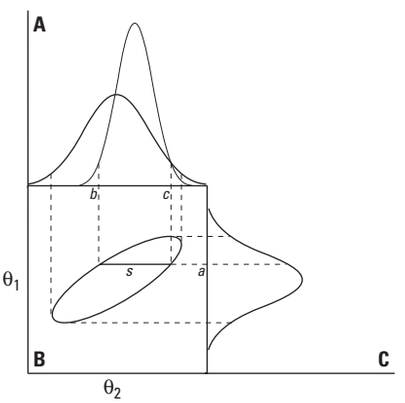


Figure 3. Illustration of covariance effects between two parameter estimates. The correct joint and marginal posteriors (i.e., data conditioned) are drawn as thick lines. Setting θ_1 to the value a before calibration leads to biased and underdispersed estimates of θ_2.

performing predictions is incorrect, as the covariance structure of θ_1 and θ_2 is disrupted in the process and the resulting marginal distribution of θ_2 is a conditional distribution on the assumed known value of θ_1. Third, in most cases, it is not sufficient to check the absence of sensitivity of some of the model predictions to variations in θ_1 values, because what is important is the sensitivity of the entire model calibration process to θ_1. Checking the sensitivity of some predictions also raises a number of questions. For example, which predictions should be checked? And what is the scientific value of a calibration exercise which gives results only for a particular application? Overall, ad hoc calibrations are an impediment to the general applicability of PBTK models to various scenarios of exposures and effects.

Confidence in Model Predictions

Models are usually developed to obtain predictions of various quantities of interest, for example, the quantity of metabolites formed for a given exposure to a specific compound. Ignoring both variability and uncertainty in model parameters renders TK model predictions almost useless for risk assessment. Several attempts have been made recently to study the propagation of variability and uncertainty in model parameters to model predictions using Monte Carlo simulation methods (*13,23 24,27,77,78*). These methods consist of: *a*) specifying a probability distribution for each model parameter; *b*) sampling randomly each model parameter from its specified distribution; *c*) running the model using the sampled parameter values, and computing various model predictions of interest. Repeating steps *b*) and *c*) a large number of times generates many different values for the model predictions. Those values can be used as samples to create histograms approximating the probability distribution of any model prediction (*79*).

Monte Carlo methods can be useful in conducting global sensitivity analysis (*25,80*). It is also possible with such methods to distinguish between variability and uncertainty in model predictions when it is feasible to separate the two for each model parameter (*81*). The validity of Monte Carlo methods is obviously highly dependent on the validity of the assumed parameter distributions. These distributions should adequately characterize the variability and/or uncertainty and covariance in the model parameters to predict the range of TK behavior that could be expected in a population and to determine the confidence to place in model predictions (*76*).

Most Monte Carlo simulations of TK models performed to date assumed for convenience that all input parameters were independent—partial exceptions can be found in Farrar et al. (*20*), Bois et al. (*21*), and Smith et al. (*76*). Most of the time this assumption is not valid. Neglecting covariances typically leads to very large confidence intervals in model predictions, overestimating their actual spread (*19*).

A better and more comprehensive approach is to sample all parameter values from their joint probability distribution (*39*). Empirically specifying such a joint distribution is a difficult task, particularly when some parameters have been estimated. In the next section we describe how a Bayesian statistical analysis can yield naturally such a joint distribution (called joint posterior distribution). Such an analysis leads to more relevant and useful Monte Carlo simulations, taking into account dependencies among all TK parameters when computing model predictions.

Multilevel Bayesian Analysis of TK Models

The above considerations led us to develop and use multilevel models of uncertainty/variability in a Bayesian framework for the statistical analysis of TK models. Those tools are described in the following sections.

Multilevel Modeling

As stated above, inferring the TK behavior of a compound in a population necessitates extracting quantitative information from individual TK data. A hierarchical structure distinguishing individual and population levels of variability is a convenient and efficient way to perform that task.

Population analyses were first introduced in the context of pharmacokinetic studies for drug development and evaluation (*74,82,83*). Their advantages have long been discussed in this context and a number of applications can be found in the pharmacokinetic literature. Yuh et al. (*84*) published a comprehensive bibliography on methodological aspects and applications of population models and a detailed review of these methods can be found in Davidian and Giltinian (*85*). However, until recently, very little attention has been paid to these approaches in a TK context. Before the work of Bois et al. (*38,86*) and Gelman et al. (*39*), the only attempt to link population principles to TK modeling was achieved by Droz et al. (*87,88*), but no fitting to experimental data was performed in this exercice.

The objective of population models is to obtain from individual data a quantitative description of the variability of the kinetic behavior of a given compound within a population. Such a model is illustrated in Figure 4. This hierarchical model has been used for TETRA and benzene by Bois and colleagues (*38,39,86*). The basic idea is that the same kinetic model f can describe the concentration–time profiles of the compound and its metabolites for each individual, and that the model parameters may vary from individual to individual. Interindividual variability is then described by assuming that each individual's parameter set ψ_i arises independently from the other individual parameter sets, from a common multivariate probability distribution.

The model presented in Figure 4 has two major components: the subject level and the population level. At the subject level, chemical concentrations are measured on each of the I individuals studied (e.g., n_i measurements for individual i). Let: $\mathbf{y}_1 = \{y_{i1}, y_{i2}, \ldots, y_{in_i}\}$ denote the set of concentration measurements made on individual i, and $\mathbf{t}_i = \{t_{i1}, t_{i2}, \ldots, t_{in_i}\}$ the set of their associated sampling times; the kinetic model, denoted f, can predict concentration–time profiles for given exposure characteristics (E), individual TK parameters (ψ_i), and physiological covariables (φ_i) (e.g., body weight, age, sex, etc.). At this level, the measurement error model is defined. Various error models have been suggested to model the differences between observed and model-predicted concentrations. One possible lognormal error model frequently used for concentration measurements is given by:

$$\log y_{ij} = \ln f(t_{ij}, E_i, \psi_i, \varphi_i) + \varepsilon_{ij}, \qquad [2]$$

where the error terms ε_{ij} are assumed to be independent and normal random variables,

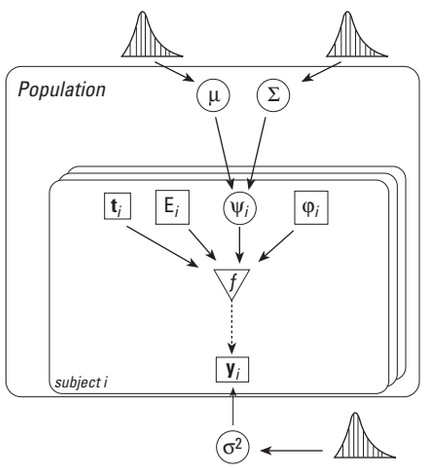


Figure 4. Graph of a hierarchical population model describing dependencies between groups of variables. Three types of nodes are featured: square nodes represent variables for which the values are known by observation, such as $\mathbf{y}_i$ or φ_i, or were fixed by the experimenters, such as E_i and $\mathbf{t}_i$; circle nodes represent unknown variables such as ψ_i, σ^2, μ, Σ; the triangle represents the deterministic TK model f. A plain arrow between two nodes indicates a direct statistical dependency between the variables of those nodes, a dashed arrow represents a deterministic link. Symbols are given in the text.

with mean 0 and variance σ^2. These error terms may represent assay precision, model mispecification, and/or random intraindividual variability in model parameters and covariables. σ^2 can be set to a fixed known value (e.g., when the analytical error is known and modeling error assumed negligible) or assumed to be a known function of actual concentration in the biological sample; it can also be estimated along with the other model parameters.

At the population level, interindividual variability is described by considering that the vectors of individual TK parameters $\psi_1, \ldots, \psi_I$ are independent realizations from a multivariate distribution, *F,* with mean vector μ and scale matrix Σ:

$$\psi_i \sim F(\mu,\Sigma),\ i = 1, \ldots, I. \qquad [3]$$

The population parameters, μ and Σ, are a priori affected by uncertainty.

Within the population framework, information is gained on the average population behavior, but at the same time information on each subject is reinforced by borrowing strength from the other subjects. The latter effect is particularly useful when few observations are available per subject, which is typical in pharmacokinetic and TK studies. This framework is flexible and, for example, can be extended to separate within-subject, between-subject, and measurement uncertainty.

A number of methods have been proposed to estimate kinetic parameters within a population framework. They can be separated into two broad classes, parametric and nonparametric. In the parametric case, the shape of the population distribution is assumed to be known, and the population parameter values are quantities to estimate (e.g., multivariate normal distribution with unknown vector of means and variance/covariance matrix). Typically, assumed distributions for the population distribution *F* include multivariate normal, lognormal, Student-*t*, or mixtures of such distributions. In the nonparametric approach, no assumption about the shape of the population distribution is made, and the entire distribution (both the shape and the parameters) is estimated from the population data (*89–91*). The parametric approach is the most commonly used because the data are often too sparse for nonparametric estimation.

The above approaches rely typically on maximum likelihood techniques (*82,92,93*) or on Bayesian principles (*94,95*). Bayesian approaches prove to be very efficient when dealing with the complexities brought about by the large number of parameters of the hierarchical structure (a 10-parameter TK model applied to 50 subjects leads to more than 500 parameters to estimate) and by the nonlinearities typically present in the subject-specific kinetic models (*94,95*). They are particularly appealing in the case of PBTK models, as they allow prior physiological information to be incorporated explicitely into the analysis (*39*).

The Bayesian Approach

A Bayesian statistical analysis allows combination of two forms of information: prior knowledge about parameter values drawn from the scientific literature, and data from TK experiments (*96,97*). In this section we first present the general principles of Bayesian statistical analyses and then we explain how likelihood and prior distributions can be elicited in the case of a population TK model; then, using the Metropolis Hastings (MH) sampler as an illustration, we explain how Markov chain Monte Carlo (MCMC) algorithms proceed, and finally, we illustrate how inference and predictions can be made from MCMC outputs.

General principles. In a Bayesian setting, all model unknowns, θ, are considered random variables. The probability distribution of an unknown model is interpreted in terms of degrees of belief about possible values of that quantity. Before conducting an experimental study, a prior probability distribution, $p(\theta)$ is constructed to reflect current knowledge of θ. This prior distribution is then updated using the data values, y, to yield a posterior probability distribution of model unknowns. Bayes' theorem indicates that the posterior probability distribution of θ given y, $p(\theta|y)$, is proportional to the product of the likelihood by the prior $p(\theta)$. In a Bayesian framework, the likelihood of a particular data set, given an assumed model and its parameter values, is the conditional probability distribution of the observed data, $p(y|\theta)$. Bayes' rule can be viewed as a formula that shows how existing beliefs, formally expressed as probability distributions, are modified by new information (*98*):

$$p(\theta|y) \propto p(y|\theta) \times p(\theta). \qquad [4]$$

All inferences about θ or a particular set of its components follow from the posterior distribution $p(\theta|y)$: means, standard deviations (SDs), quantiles, correlations, and marginal distributions are obtainable by integration of $p(\theta|y)$. In the case of the above population TK model, model unknowns, θ, comprise the individual TK parameter vectors ψ_i, $i = 1, \ldots, I$ (denoted ψ in the following, for convenience), the population parameter vector or matrix, μ and Σ, and a variance term for the measurement errors, σ^2. Hence, Equation 4 takes the form:

$$p(\psi,\mu,\Sigma,\sigma^2|y) \propto p(y|\psi,\mu,\Sigma,\sigma^2) \times p(\psi,\mu,\Sigma,\sigma^2). \qquad [5]$$

It is possible to rewrite Equation 5 as a product of simpler conditional distributions, considering the following arguments, which are legitimate in any similar hierarchical model. First, arguments of conditional independence (*99–101*) allow us to simplify the first term of the right side of Equation 5:

$$p(y|\psi,\mu,\Sigma,\sigma^2) = p(y|\psi,\sigma^2). \qquad [6]$$

Second, assuming that the experimental error variance is a priori independent of (ψ, μ, Σ) (which is reasonable in most applications), and by the definition of conditional probabilities, the second term of the right side of Equation 5 can be written as:

$$\begin{aligned} p(\psi,\mu,\Sigma,\sigma^2) &= p(\psi,\mu,\Sigma) \times p(\sigma^2) \\ &= p(\psi|\mu,\Sigma) \times p(\mu,\Sigma) \times p(\sigma^2), \end{aligned} \qquad [7]$$

leading to the following expression for $p(\theta|y)$:

$$\begin{aligned} p(\psi,\mu,\Sigma,\sigma^2|y) \propto\ & p(y|\psi,\sigma^2) \times \\ & p(\psi|\mu,\Sigma) \times p(\mu,\Sigma) \times p(\sigma^2); \end{aligned} \qquad [8]$$

finally, μ and Σ are usually assumed to be a priori independent, allowing the factorization of $p(\mu, \Sigma)$:

$$\begin{aligned} p(\psi,\mu,\Sigma,\sigma^2|y) \propto\ & p(y|\psi,\sigma^2) \times \\ & p(\psi|\mu,\Sigma) \times p(\mu) \times p(\Sigma) \times p(\sigma^2). \end{aligned} \qquad [9]$$

Even though Equation 9 involves simpler components than Equation 5, it is impossible to obtain an analytical expression for $p(\theta|y)$, in particular because of the nonlinear form of the TK model. Until recently, this intractability has been a strong impediment to practical Bayesian analyses. Fortunately, the recent advent of MCMC methods has alleviated a number of these difficulties. These methods are powerful tools to provide samples of parameter values from $p(\theta|y)$ even without knowledge of the analytical expression of $p(\theta|y)$ (*102,103*). They have largely contributed to the recent proliferation of Bayesian analyses in applied statistics. These methods require the functional forms for $p(y|\psi,\sigma^2)$ (the likelihood), $p(\psi|\mu,\Sigma)$ (the population model per se), and $p(\mu)$, $p(\Sigma)$, and $p(\sigma^2)$ (the priors).

Specification of likelihood and prior distributions. The choice of a likelihood function embodies an error model and has been discussed previously. For example, consider the 15-parameter ($K = 15$) PBTK model depicted in Figure 1 and the associated population model in Figure 4 (*38*). This model was calibrated using TETRA concentration data measured in the exhaled air and venous blood of I = 6 human volunteers exposed to 70 and 144 ppm of TETRA for 4 hr (*104*). To specify a likelihood on the observed individual concentrations (n_i observations for subject i), it was assumed that analytical

errors were independent and lognormally distributed. The first-stage likelihood $p(\mathbf{y}|\psi,\sigma^2)$ was then:

$$\prod_{i=1}^{I} p\left(\mathbf{y}_i|\psi_i,\sigma^2\right) = \prod_{i=1}^{I}\prod_{j=1}^{n_i} p\left(y_{ij}|\psi_i,\sigma^2\right) = \prod_{i=1}^{I}\prod_{j=1}^{n_i} N\left[\ln f\left(t_{ij},E_i,\psi_i,\varphi_i\right),\sigma^2\right]. \quad [10]$$

The variance term σ^2 was a vector with two components: σ_1^2 for the measurements in blood, and σ_2^2 for the measurements in exhaled air; these measurements had different experimental protocols and were thus likely to have different precisions.

Interindividual variability was described by assuming that each component ψ_{ik} of the individual TK parameter sets ψ_i, was distributed lognormally, with population parameters μ_k and Σ_{kk}^2 (in log-space). So the population model $p(\psi|\mu,\Sigma)$ was given by:

$$p(\psi|\mu,\Sigma) = \prod_{i=1}^{I}\prod_{k=1}^{K} p\left(\psi_{ik}|\mu_k,\Sigma_{kk}^2\right), \quad [11]$$

and

$$\ln\psi_{ik} \sim N(\ln\mu_k,\Sigma_{kk}^2). \quad [12]$$

The use of a probability distribution to represent prior knowledge offers a unified way to account for precise as well as vague information available before experiments. There is a large body of literature on the elicitation of prior distributions (*96,105,106*). When little prior knowledge is available on a particular parameter value, a noninformative distribution (i.e., flat-shaped) can be used. Conversely, in the case of stronger prior knowledge, the distribution shape, location, and dispersion parameters should be chosen to represent that information as appropriately as possible. As stated previously, there is usually very little prior knowledge of the parameter values for a particular individual because information obtainable from the literature relates more often to average values or variances for a human population. Within a population model, such information can be directly used under the form of prior distributions for the population parameters μ and Σ (Figure 4). For example, when a PBTK model is considered, reference values available in the literature (*71,107*) for physiological parameters such as blood flows and organ volumes yield a sound basis for setting up informative prior distributions. For chemical-specific parameters, prior information is often less precise but can still be obtained from previously published experiments (for example, *in vitro* determinations of partition coefficients or estimates of metabolic parameters obtained from *in vivo* experiments). In all cases, it is preferable when setting uncertainties on population parameters to be conservative and set the prior population variances higher rather than lower when there is ambiguity. To stay in physiological or biochemical plausible ranges, truncated distributions can be used.

In the example cited above, all population parameters were assumed a priori independent (after scaling functions had been applied). The prior distribution assigned to each term, μ_k, of the population mean μ was a truncated normal distribution (with parameters M_k and S_k^2 in log-space). So the prior on the mean population vector $p(\mu)$ was given by:

$$p(\mu) = \prod_{k=1}^{K} p(\mu_k) \quad [13]$$

and

$$\mu_k \sim N(M_k,S_k^2) \quad [14]$$

The prior distribution assigned to each population variance Σ_{kk}^2 was an inverse-gamma distribution with shape parameter equal to one (to indicate large uncertainties) and scale parameter ω_k^2 (ω_k^2 was the prior estimate of the true population variance). $p(\Sigma)$ was therefore given by:

$$p(\Sigma) = \prod_{k=1}^{K} p\left(\Sigma_{kk}^2\right) \quad [15]$$

where

$$\Sigma_{kk}^2 \sim InvGamma(1,\omega_k^2) \quad [16]$$

The quantities M_k, S_k^2, and ω_k^2 are called hyperparameters. They directly embody prior knowledge. For each k, M_k relates to the location of the population mean μ_k, S_k^2 characterizes the (prior) uncertainty associated with this location; ω_k^2 is used to represent prior beliefs on the variability of the TK parameter ψ_k within the population. Variability and uncertainty were thus formally distinguished. As an example, consider the metabolic parameter V_m, which represents the maximum rate of TETRA metabolism. To reflect the large prior uncertainty associated with that parameter, $\exp(S_k^2)$ was set to 10, while $\exp(\omega_k^2)$ was set to 2. This indicates that this parameter was believed to vary by a factor 2 in the population studied, but its true population mean was uncertain by a factor of 10. It would be difficult to express such uncertainties without an explicit hierarchical model.

Finally, the prior distributions assigned to the experimental error variances σ_1^2 and σ_2^2 were standard noninformative priors for variances (*108*):

$$p(\sigma^2) = p(\sigma_1^2,\sigma_2^2) \propto \sigma_1^{-2} \times \sigma_2^{-2}. \quad [17]$$

Computational aspects—MCMC sampling. MCMC methods are iterative sampling schemes that provide direct approximations to a complex joint posterior distribution (by random draws from it sets of parameter values). These methods consist in constructing a Markov chain that converges, as iterations progress, toward the posterior distribution, $p(\theta|\mathbf{y})$. Various MCMC methods have been developed to date; they differ in ways the Markov chain is constructed. The MH algorithm proved in our experience to be very efficient in dealing with Bayesian population TK models. Briefly, it proceeds as follows: at the beginning, all model unknowns are assigned values; for example, by sampling from their respective prior distribution. At each following iteration step, each component, θ_k, of the parameter vector θ is eventually updated according to the following acception/rejection rule: a proposed value, θ_k', is sampled from a "proposal" distribution (e.g., normal, centered on the current value θ_k). The joint posterior density is then computed at θ_k and θ_k' (up to a proportionality constant, using Equation 9). Label these two density values π and π'. If the ratio π'/π exceeds 1, the new value θ_k' is accepted and replaces θ_k; otherwise, θ_k' is accepted only with probability π'/π. In case of rejection of θ_k', the value θ_k is kept. After (eventually) upating all components of θ sequentially, their values are recorded, which completes one iteration of the Markov chain. Many iterations are typically needed. It has been shown that under some regularity conditions (i.e., the Markov chain has to be irreducible, aperiodic, and positive recurrent), the chain converges toward the distribution of interest, i.e., the joint posterior distribution. This means that after a sufficient number of iterations, the samples generated by such a process can be considered approximate samples from $p(\theta|\mathbf{y})$, and thus can be used to make inferences. Implementation details (e.g., how to choose the proposal distribution) and techniques to monitor convergence are well described in the literature (*102,103,109–111*) and will not be discussed here. Briefly, the main desired characteristics of the proposal distribution are related to ease of simulation and approximation of the posterior distribution.

MCMC methods provide joint distributions of parameter estimates appropriate for uncertainty analyses because they account for the full dependence structure between parameter estimates. They are directly usable as inputs for uncertainty analysis of exposure–dose relationships.

MCMC methods can actually be viewed as extensions of the traditional or standard Monte Carlo methods for uncertainty analysis

(*9*). Figure 5 is an illustration of MCMC sampling, compared to simple Monte Carlo sampling. In the former, the values drawn for each parameter start from its prior distribution and converge, as iterations progress, to a data-adjusted posterior distribution. In the case of simple Monte Carlo sampling, the values are always drawn from the prior distribution. The use of uniform priors in a Bayesian analysis leads to posteriors strictly proportional to the data likelihood and is equivalent to the frequentist maximum likelihood estimation. Conversely, if the data do not convey information about some parameters, the corresponding posteriors are equal to the priors and this approach becomes equivalent to standard Monte Carlo sampling from the priors.

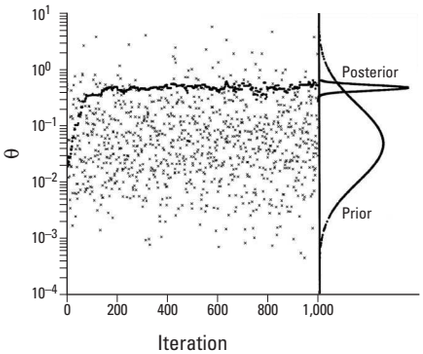


Figure 5. Illustration of MCMC sampling compared to simple Monte Carlo sampling. In MCMC sampling (dots), the values drawn for any parameter θ start from the prior distribution and converge, as iterations progress, to a data-adjusted posterior distribution. The posterior density corresponds to the product of the prior density by the data likelihood. In the case of simple Monte Carlo sampling (crosses), the values are always drawn from the prior distribution. Reproduced from Bois (*119*) with permission from *Environmental Health Perspectives*.

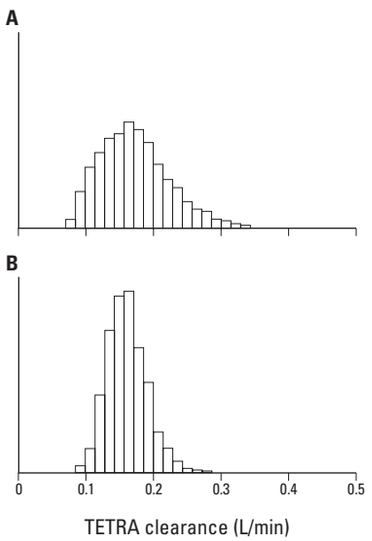


Figure 6. Posterior distributions of TETRA clearance for subject A (*A*) and mean population clearance (*B*) obtained by calibrating the TK model depicted in Figure 2 to the data of Bernillon et al. (*55*). The histograms bin 3,000 parameter values sampled from their posterior distribution.

Analyzing MCMC outputs. MCMC sampling methods provide joint multivariate samples of parameter values from $p(\theta|y)$. Posterior means, SDs, and quantiles of any single (or set of) component(s) of θ can be estimated from these samples by computing their equivalent in the MCMC output. As an illustration, we present some results obtained by calibrating, within a Bayesain hierarchical framework, the TK model depicted in Figure 2 with the data of Bernillon et al. (*55*). These data consisted in concentration–time profiles of TETRA in air and blood and TCA in blood and urine, followed for 1 week after exposure for six volunteers exposed once or twice to 1 ppm TETRA for 6 hr.

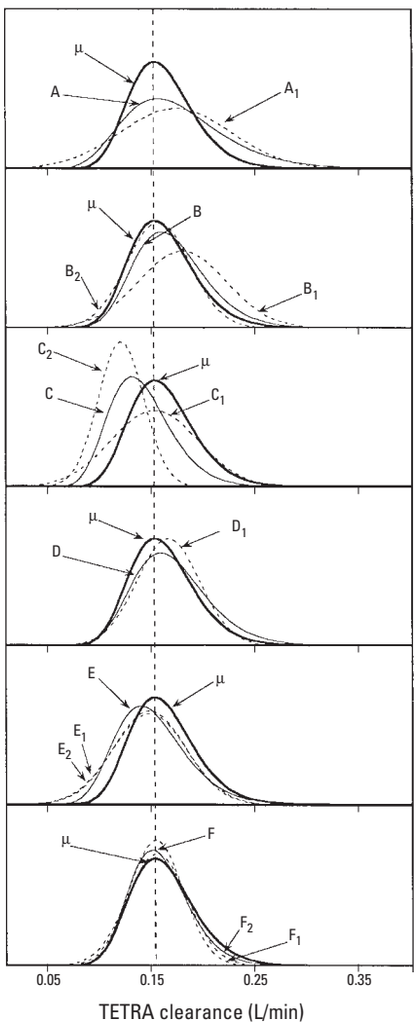


Figure 7. Posterior distributions of TETRA clearance parameters (occasion-specific, individual mean, and population mean) estimated by calibration of the TK model depicted on Figure 2 with the data of Bernillon et al. (*55*). The population mean distribution is denoted by μ (thick line), individual means by unsubscripted capitals A–F (thin lines), occasion-specific distribution values by subscripted capitals A_1–F_2 (dashed lines). The mode of the distribution of the population mean is indicated by the vertical dashed line.

Histograms can be used to represent the marginal distribution of individual or population parameters. Figure 6A is a histogram of the estimates of TETRA clearance (mean, 0.17 L/min[1]; SD, 0.05) for subject A. Figure 6B is a histogram of the distribution of the population mean of TETRA clearance (mean, 0.16 L/min; SD, 0.03) for the six subjects studied. The distribution of the population mean was slightly skewed.

In Bernillon et al. (*55*), four of the six subjects studied underwent repeated exposures to TETRA; intraindividual variability was thus assessed in addition to interindividual variability. This was achieved by adding one level to the hierarchy: the first level was the occasion level (where the observations were available); at the individual level, occasion-specific parameters were supposed to vary among exposures around a subject-specific mean; at the population level, these subject-specific means were supposed to vary among individuals. Figure 7 illustrates the posterior distributions of TETRA clearance at each of these three levels: occasion, individual mean, and population mean. For display purposes, the distributions are approximated by normal (occasion-level parameters) or lognormal (individual or population mean parameters) distributions, whose means and SDs have been computed on the posterior samples. A subject's average clearance estimate is not only influenced by his occasion-specific estimates but also by the population mean. In other words, sharing information across individuals and occasions occurs through the hierarchical structure.

In the six-subject sample studied, interindividual differences in clearance reached a factor of 1.2 (between subjects A and C); intraindividual (i.e., interoccasion) variability reached a factor of 1.3 (subject C). Inference for a larger population is given by the size of the posterior interindividual and intraindividual variances, whose corresponding coefficients of variation were estimated at 33 and 18%, respectively. Using the population parameter estimates, it is possible to simulate additional hypothetical individuals.

Correlations between parameters can also be studied through correlation matrices (estimated from the MCMC output) or simple scatterplots. Figure 8 illustrates the correlation between the occasion-specific estimates of TETRA clearance and the TETRA central/air partition coefficient for subject A. These two parameters are negatively correlated ($\hat{r} = -0.33$).

Distributions of model predictions of toxicologic interest can also be easily obtained.

Figure 9 gives the posterior distributions of the fraction of TETRA metabolized by the six individuals studied, for each occasion of exposure. These results were obtained by running the TK model once for each of 3,000 vectors from the MCMC ouptuts for each individual and occasion-specific parameters simulating a 1 ppm exposure to TETRA for 6 hr. Large intraindividual variability is observed following those of TETRA clearance estimates (Figure 7). Estimates of the fraction of TETRA metabolized after 1 ppm 6-hr exposures range from 14% (subject C, occasion 2) to 25% (subject D, occasion 1). This tends to confirm our results obtained by extrapolating the results obtained by calibration of a PBTK model to high-exposure data (70 and 140 ppm) (*38*).

After calibration, residual (posterior) uncertainty is reflected by the range of parameter estimates. For each parameter, an estimate of variability is given by the location of the corresponding (intraindividual, interindividual) variance parameter. Note that the variance parameters are themselves affected by uncertainty.

Sensitivity analyses. Bayesian analyses call for the specification of prior distributions. A legitimate question is the sensitivity of the results (either posterior parameter estimates or model predictions) with respect to the choice of those priors. It is therefore recommended that diagnostics of sensitivity be performed. That is quite easy to do after MCMC sampling, for example, by plotting and analyzing the correlations between results and sampled parameter values (*39*). The prior assumptions about the parameters that strongly influence the results can be further evaluated. It is possible, for example, to study the effect of changing prior specifications.

The question of sensitivity to the priors can also be answered generally. Note first that the choice of priors should always be transparent, i.e., fully stated, to allow useful discussion of the results. Clearly, the choice of priors for noninfluential parameters should not be a problem. For influential parameters, the priors used can either be vague or informative (i.e., biologically motivated and knowledge based) and the posteriors distributions can be either prior- or data-driven. Four major cases can therefore be envisioned:

- Both prior and posterior are vague (the data brought no information). The results will be prior dependent and more information should be acquired either in form of prior knowledge or additional data.
- The prior is vague, but the posterior is informative (the data are driving the estimates). In that case, the conclusions should be robust with respect to the prior.
- The prior is informative and reinforced by the data (data and prior agree). Here also, if the data are strongly informative, the conclusions should be robust. When the relative weight of prior and data is less clearcut, further checks of sensitivity are warranted.
- The prior is informative but the data conflict with it. The conclusions will be prior dependent, but in that case the model, the prior, and the data should be questioned to resolve the discrepancy.

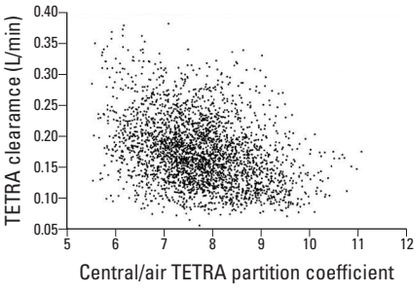


Figure 8. Posterior correlations between TETRA clearance and central/air TETRA partition coefficient estimates (subject A) obtained by calibrating the TK model depicted in Figure 2 to the data of Bernillon et al. (*55*). The dots represent 3,000 pairs of parameter values from the MCMC outputs. The correlation coefficient is estimated at –0.33.

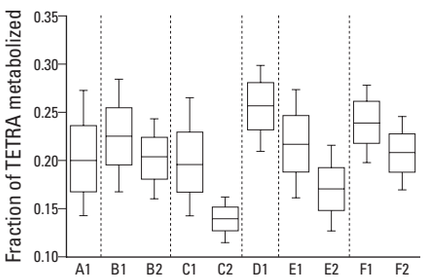


Figure 9. Prediction of the fraction of TETRA metabolized after an exposure to 1 ppm TETRA for 6 hr by the six subjects studied at each occasion. Box plots of 3,000 estimates are displayed. Each box encloses 50% of the predictions; the lines extending from the top and bottom of each box mark the 10th and 90th percentiles. The line inside the box represents the median value.

Perspectives

The objective of this article was to present a Bayesian statistical approach to the analysis of population TK models, which offers a number of advantages over previously proposed approaches. The proposed methodology is far from definitely established and several issues still must be resolved. For example, several modeling assumptions are involved, particularly within the population model, that should be further validated. Hopefully, experience gained from population analyses using simpler models can be transposed to PBPK models. Another task is to reduce the associated computational burden so as to quickly explore alternative models. Simple MCMC samplers can sometimes converge slowly or even fail to converge. Improved algorithms recently have been proposed (*103,112–115*) and should be tested. In addition, the definition of prior distributions is not always an easy task because of data accessibility. Although it is well known that TK parameters exhibit inter- or intraindividual variability in humans (and other species), the only values readily available and those commonly used on physiological modeling are reference values for young Caucasian males. The use of reference values artificially reduces estimates of population variance. Information about population variability may eventually be found by researching the original publications, but even these often lack adequate statistical treatment. Data about the shape of the distributions are even harder to find. A database giving population distributions of important physiological parameter values together with their correlations is needed. Such a database would be useful for all types of physiological modeling and for both toxicants and drugs. Despite these limitations, experience has been sufficient to demonstrate that the Bayesian approach provides reasonable estimates of uncertainty and variability of TK model parameters and predictions.

Is it worth making the computational effort implied by such an approach? The answer to that question is a matter of choice and scientific judgment. A PBTK model with standard parameter values might be sufficient for simple exploratory analyses or for theoretical work on model structure. Alternatively, for the full analysis of a data set or for risk assessment purposes, a formal model calibration along the lines presented here appears necessary. For example, in a scientific data analysis context, Smith (*116*) illustrated, in the case of occupational exposure to gasoline, how PBTK and pharmacodynamic models could be combined with epidemiological data to evaluate concurrent reasonable hypotheses concerning mechanisms of toxicity. The statistical framework described in this article is convenient in such a setting because distributions of internal doses are more suitable than average point estimates in comparing alternative hypotheses. We have shown how to extend the methodology for optimal designing of occupational monitoring programs or TK studies (*73*).

This approach is also particularly suited for risk assessment and decision making. For example, population extrapolations can be accomplished if the factors responsible for heterogeneity (e.g., body weight, breathing rate, metabolic constants) are among the list of model parameters. In this case, the population distributions obtained for these parameters, from fitting datasets with few volunteers,

can be changed to better reflect the range of values found in the general population or in sensitive subpopulations (e.g., children). Alternatively, if these parameters are measured for each volunteer and introduced as fixed covariates during the calibration procedure (*117*), they can be replaced by distributions when making predictions. In this way, the predictions of interest (e.g., internal metabolite dose, fraction metabolized) can be extrapolated beyond the group of individuals used for calibrating the model. Note that this statistical approach also can be used to calibrate the other (e.g., fate and transport, cancer effects) models used in risk assessment (*118*).

Finally, one of the challenges of toxicological modeling is the full exploitation of the numerous datasets collected during epidemiological or occupational hygiene studies, generally in settings in which exposure concentrations are unknown. Most of the time, exposure is represented in a considerably simplified and approximate manner, which can be misleading. It is possible using the above statistical framework to consider exposure as one of the estimands. A major problem, however, resides in accounting fully for the uncertainties stemming from unknown time-varying exposures. The impact of particular functional form for the time evolution of exposure has not yet been thoroughly studied and validated. We are confident that as progress is made toward answering such questions, TK modeling will become a more powerful and widespread tool for toxicity and risk assessments.

References and Notes

1. Greim H, Csanady G, Filser JG, Kreuzer P, Schwarz L, Wolff T, Werner S. Biomarkers as tools in human health risk assessment. Clin Chem 41:1804–1808 (1995).
2. Conney AH. Induction of microsomal enzymes by foreign chemicals and carcinogenesis by polycyclic aromatic hydrocarbons: G. H. A. Clowes Memorial Lecture. Cancer Res 42:4875–917 (1982).
3. Gehring PJ, Blau GE. Mechanisms of carcinogenesis: dose response. J Environ Pathol Toxicol 1:163–179 (1977).
4. Hoel DG, Kaplan NL, Anderson MW. Implication of nonlinear kinetics on risk estimation in carcinogenesis. Science 219:1032–1037 (1983).
5. Johannsen FR. Risk assessment of carcinogenic and noncarcinogenic chemicals. CRC Crit Rev Toxicol 20:341–367 (1990).
6. Anderson MW, Hoel DG, Kaplan NL. A general scheme for the incorporation of pharmacokinetics in low-dose risk estimation for chemical carcinogenesis: example--vinyl chloride. Toxicol Appl Pharmacol 55:154–161 (1980).
7. Sharma R, Coulombe R. Pharmacokinetics and Risk Assessment. In: Toxicology and Risk Assessment; Principles, Methods and Applications (Fan A, Chang L, eds). New York:Dekker, 1996;81–99.
8. Page NP, Singh DV, Farland W, Goodman JI, Conolly RB, Andersen ME, Clewell HJ, Frederick CB, Yamasaki H, Lucier G. Implementation of EPA Revised Cancer Assessment Guidelines: Incorporation of Mechanistic and Pharmacokinetic Data. Fundam Appl Toxicol 37:16–36 (1997).
9. Bailer AJ, Dankovic DA. An introduction to the use of physiologically based pharmacokinetic models in risk assessment. Stat Methods Med Res 6:341–358 (1997).
10. United States Department of Labor. 29 CFR Parts 1910, 1915 and 1926. Fed Reg 1997.
11. Gargas ML, Medinsky MA, Andersen ME. Pharmacokinetic modeling approaches for describing the uptake, systemic distribution, and disposition of inhaled chemicals. Crit Rev Toxicol 25:237–254 (1995).
12. Andersen ME, Clewell HJ III, Gargas ML, Smith FA, Reitz RH. Physiologically based pharmacokinetics and the risk assessment process for methylene chloride. Toxicol Appl Pharmacol 87:185–205 (1987).
13. Cronin WJ, Oswald EJ, Shelley ML, Fisher JW, Flemming CD. A trichloroethylene risk assessment using a Monte Carlo analysis of parameter uncertainty in conjunction with physiologically-based pharmacokinetic modeling. Risk Anal 15:555–565 (1995).
14. Kohn MC, Melnick RL. Effects of the structure of a toxicokinetic model of butadiene inhalation exposure on computed production of carcinogenic intermediates. Toxicology 113:31–39 (1996).
15. Watanabe KH, Bois FY. Interspecies extrapolation of physiological pharmacokinetic parameter distributions. Risk Anal 16:741–754 (1996).
16. Adler ID, Cochrane J, Osterman-Golkar S, Skopek TR, Sorsa M, Vogel E. 1,3-Butadiene working group report. Mutat Res 330:101–114 (1995).
17. Morris JB, Robinson DE, Vollmuth TA, Brown RP, Domeyer BE. A parallelogram approach for safety evaluation of ingested acetaldehyde. Regul Toxicol Pharmacol 24:251–263 (1996).
18. Kohn MC. Achieving credibility in risk assessment models. Toxicol Lett 79:107–114 (1995).
19. Portier CJ, Kaplan NL. Variability of safe dose estimates when using complicated models of the carcinogenic process. Fundam Appl Toxicol 13:533–544 (1989).
20. Farrar D, Allen B, Crump K, Shipp A. Evaluation of uncertainty in input parameters to pharmacokinetic models and the resulting uncertainty in output. Toxicol Lett 49:371–385 (1989).
21. Bois FY, Zeise L, Tozer TN. Precision and sensitivity of pharmacokinetic models for cancer risk assessment: tetrachloroethylene in mice, rats, and humans. Toxicol Appl Pharmacol 102:300–315 (1990).
22. Allen BC, Covington TR, Clewell HJ. Investigation of the impact of pharmacokinetic variability and uncertainty on risks predicted with a pharmacokinetic model for chloroform. Toxicology 111:289–303 (1996).
23. Woodruff TJ, Bois FY, Auslander D, Spear RC. Structure and parameterization of pharmacokinetic models: their impact on model predictions. Risk Anal 12:189–201 (1992).
24. Gearhart JM, Mahle DA, Greene RJ, Seckel CS, Flemming CD, Fisher JW, Clewell HJD. Variability of physiologically based pharmacokinetic (PBPK) model parameters and their effects on PBPK model predictions in a risk assessment for perchloroethylene (PCE). Toxicol Lett 68:131–144 (1993).
25. Spear RC, Bois FY. Parameter variability and the interpretation of physiologically based pharmacokinetic modeling results. Environ Health Perspect 102(suppl 11):61–66 (1994).
26. Krewski D, Wang Y, Bartlett S, Krishnan K. Uncertainty, variability, and sensitivity analysis in physiological pharmacokinetic models. J Biopharm Stat 5:245–271 (1995).
27. Thomas RS, Bigelow PL, Keefe TJ, Yang RS. Variability in biological exposure indices using physiologically based pharmacokinetic modeling and Monte Carlo simulation. Am Ind Hyg Assoc J 57:23–32 (1996).
28. Louis TA. Assessing, accommodating, and interpreting the influences of heterogeneity. Environ Health Perspect 90:215–222 (1991).
29. Smith AE, Evans JS. Uncertainty in fitted estimates of apparent in vivo metabolic constants for chloroform. Fundam Appl Toxicol 25:29–44 (1995).
30. Gibaldi M, Perrier D. Pharmacokinetics. New York:Dekker, 1982.
31. Balant LP, Gex-Fabry M. Physiological pharmacokinetic modelling. Xenobiotica 20:1241–1257 (1990).
32. Fernandez JG, Droz PO, Humbert BE, Caperos JR. Trichloroethylene exposure - Simulation of uptake, excretion, and metabolism using a mathematical model. Br J Ind Med 34:43–55 (1977).
33. Andersen ME, Ramsey JC. A physiologically based description of the inhalation pharmacokinetics of styrene in rats and humans. Dev Toxicol Environ Sci 11:415–418 (1983).
34. Droz PO, Guillemin MP. Human styrene exposure - V. Development of a model for biological monitoring. Int Arch Occup Environ Health 53:19–36 (1983).
35. Ritschel WA, Banerjee PS. Physiological pharmacokinetic models: principles, applications, limitations and outlook. Methods Find Exp Clin Pharmacol 8:603–614 (1986).
36. Gray DG. A physiologically based pharmacokinetic model for methyl mercury in the pregnant rat and fetus. Toxicol Appl Pharmacol 132:91–102 (1995).
37. Kohn MC. The importance of anatomical realism for validation of physiological models of disposition of inhaled toxicants. Toxicol Appl Pharmacol 147:448–458 (1997).
38. Bois FY, Gelman A, Jiang J, Maszle D, Zeise L, Alexeef G. Population toxicokinetics of tetrachloroethylene. Arch Toxicol 70:347–355 (1996).
39. Gelman A, Bois FY, Jiang J. Physiological pharmacokinetic analysis using population modeling and informative prior distributions. J Am Stat Assoc 91:1400–1412 (1996).
40. Angelo MJ, Bischoff KB, Pritchard AB, Presser MA. A physiological model of the pharmacokinetics of methylene chloride in B6C3F1 mice following I.V. administrations. J Pharmacokinet Biopharm 12:413–436 (1984).
41. Lutz RJ, Dedrick RL, Tuey D, Sipes G, Anderson MW, Matthews HB. Comparison of the pharmacokinetics of several polychlorinated biphenyls in mouse, rat, dog, and monkey by means of a physiological pharmacokinetic model. Drug Metab Dispos 12:527–535 (1984).
42. Johanson G. Physiologically based pharmacokinetic modeling of inhaled 2-butoxyethanol in man. Toxicol Lett 34:23–31 (1986).
43. Rao HV, Brown DR. A physiologically based pharmacokinetic assessment of tetrachloroethylene in groundwater for a bathing and showering determination. Risk Anal 13:37–49 (1993).
44. Roy A, Weisel CP, Lioy PJ, Georgopoulos PG. A distributed parameter physiologically-based pharmacokinetic model for dermal and inhalation exposure to volatile organic compounds. Risk Anal 16:147–160 (1996).
45. Åstrand I. Effect of physical exercise on uptake, distribution and elimination of vapors in man. In: Modeling of Inhalation Exposure to Vapors: Uptake, Distribution, and Elimination (Fiserova-Bergerova F, ed). Boca Raton, Florida:CRC Press, 1983;107–130.
46. O'Flaherty EJ, Scott W, Schreiner C, Beliles RP. A physiologically based kinetic model of rat and mouse gestation: disposition of a weak acid. Toxicol Appl Pharmacol 112:245–256 (1992).
47. Luecke RH, Wosilait WD, Pearce BA, Young JF. A physiologically based pharmacokinetic computer model for human pregnancy. Teratology 49:90–103 (1994).
48. O'Flaherty EJ. Physiologically based pharmacokinetic models in developmental toxicology. Risk Anal 14:605–611 (1994).
49. Terry KK, Elswick BA, Welsch F, Conolly RB. Development of a physiologically based pharmacokinetic model describing 2-methoxyacetic acid disposition in the pregnant mouse. Toxicol Appl Pharmacol 132:103–114 (1995).
50. O'Flaherty EJ, Polák J, Andriot MD. Incorporation of temporal factors into physiologically based kinetic models for risk assessment. Inhal Toxicol 7:917–925 (1995).
51. Welsch F, Blumenthal GM, Conolly RB. Physiologically based pharmacokinetic models applicable to organogenesis: extrapolation between species and potential use in prenatal toxicity risk assessments. Toxicol Lett 82/83:539–547 (1995).
52. Luecke RH, Wosilait WD, Pearce BA, Young JF. A computer model and program for xenobiotic disposition during pregnancy. Comput Methods Programs Biomed 53:201–224 (1997).
53. Shelley ML, Andersen ME, Fisher JW. An inhalation distribution model for the lactating mother and nursing child. Toxicol Lett 43:23–29 (1988).
54. Byczkowski JZ, Fisher JW. Lactational transfer of tetrachloroethylene in rats. Risk Anal 14:339–349 (1994).
55. Bernillon P, Pullens R, Kezic S, Auburtin G, Dujardin R, Monster A, Bois FY. Unpublished data.
56. Fanning EW. Improving the biological basis for carcinogen risk assessment: benzene-induced leukemia as a case study [PhD Thesis]. Berkeley, CA:University of California, 1998.
57. Thomaseth K, Salvan A. Estimation of occupational exposure to 2,3,7,8-tetrachlorodibenzo-*p*-dioxin using a minimal physiologic toxicokinetic model. Environ Health Perspect 106(suppl 2):743–753 (1998).
58. Adolph EF. Quantitative relations in the physiological constitution of mammals. Science 109:579–585 (1949).
59. Boxenbaum H. Interspecies scaling, allometry, physiological time, and the ground plan of pharmacokinetics. J Pharmacokinet Biopharm 10:201–227 (1982).
60. Boxenbaum H. Interspecies pharmacokinetic scaling and the evolutionary-comparative paradigm. Drug Metab Rev 15:1071–1121 (1984).
61. Davidson IWF, Parker JC, Beliles RP. Biological basis for extrapolation across mammalian species. Regul Toxicol Pharmacol 6:211–237 (1986).
62. Mordenti J. Man versus beast: pharmacokinetic scaling in mammals. J Pharm Sci 75:1028–1040 (1986).
63. Ings RM. Interspecies scaling and comparisons in drug development and toxicokinetics. Xenobiotica 20:1201–1231 (1990).
64. Cobelli C, DiStefano JJ. Parameter and structural identifiability concepts and ambiguities: a critical review and analysis. Am J Physiol 242(5):R7–R24 (1980).
65. Jacquez JA. Identifiability: the first step in parameter estimation. Fed Proc 46 (8):2477–2480 (1987).

66. Jacquez JA, Perry T. Parameter estimation: local identifiability of parameters. Am J Physiol 258(4):E727–E736 (1990).
67. Godfrey KR, Chapman MJ, Vajda S. Identifiability and indistinguishability of nonlinear pharmacokinetic models. J Pharmacokinet Biopharm 22 (3):229–251 (1994).
68. Slob W, Janssen PHM, Van den Hof JM. Structural identifiability of PBPK models: practical consequences for modeling strategies and study designs. Crit Rev Toxicol 27:261–272 (1997).
69. Omlin M, Reichert P. A comparison of techniques for the estimation of model prediction uncertainty. Ecol Model 115:45–59 (1999).
70. Rowe WD. Understanding uncertainty. Risk Anal 14:743–750 (1994).
71. ICRP. 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:Pergamon Press, 1975.
72. Beatty D, Piegorsch W. Optimal statistical design for toxicokinetic studies. Stat Methods Med Res 6:359–376 (1997).
73. Bois FY, Smith TJ, Gelman A, Chang HY, Smith AE. Optimal design for a study of butadiene toxicokinetics in humans. Toxicol Sci 49:213–224 (1999).
74. Beal SL, Sheiner LB. Methodology of population pharmacokinetics. In: Drug Fate and Metabolism, Vol 5 (Garret ER, Hirtz J, eds). New York:Dekker, 1985;135–183.
75. Racine-Poon A, Smith AFM. Population models. In: Statistical Methodology in the Pharmaceutical Sciences (Berry DA, ed). New York:Dekker, 1990;139–162.
76. Smith AE, Ryan PB, Evans JS. The effect of neglecting correlations when propagating uncertainty and estimating the population distribution of risk. Risk Anal 12:467–474 (1992).
77. Spear RC, Bois FY, Woodruff T, Auslander D, Parker J, Selvin S. Modeling benzene pharmacokinetics across three sets of animal data: parametric sensitivity and risk implications. Risk Anal 11:641–654 (1991).
78. Hetrick DM, Jarabek AM, Travis CC. Sensitivity analysis for physiologically based pharmacokinetic models. J Pharmacokinet Biopharm 19:1–20 (1991).
79. Rubinstein R. Simulation and the Monte Carlo Method. New-York:Wiley, 1981.
80. Hammersley JM, Handscomb DC. Monte Carlo Methods. London:Chapman & Hall, 1964.
81. Hoffman FO, Hammonds JS. Propagation of uncertainty in risk assessments: the need to distinguish between uncertainty due to lack of knowledge and uncertainty due to variability. Risk Anal 14:707–712 (1994).
82. Beal SL, Sheiner LB. Estimating population kinetics. CRC Crit Rev Biomed Eng 8:195–222 (1982).
83. Sheiner LB. The population approach to pharmacokinetic data analysis: rationale and standard data analysis methods. Drug Metab Rev 15:153–171 (1984).
84. Yuh L, Beal S, Davidian M, Harrison F, Hester A, Kowalski K, Vonesh E, Wolfinger R. Population pharmacokinetic/pharmacodynamic methodology and applications: a bibliography. Biometrics 50:566–575 (1994).
85. Davidian M, Giltinian D. Nonlinear Models for Repeated Measurements Data. New York:Chapman & Hall, 1995.
86. Bois FY, Jackson ET, Pekari K, Smith MT. Population toxicokinetic of benzene. Environ Health Perspect 104:1405–1411 (1996).
87. Droz PO, Wu MM, Cumberland WG, Berode M. Variability in biological monitoring of solvent exposure. I: Development and validation of a population model. Br J Ind Med 46:447–460 (1989).
88. Droz PO, Wu MM, Cumberland WG. Variability in biological monitoring of organic solvent exposure. I: Application of a population physiological model. Br J Ind Med 46:547–558 (1989).
89. Mallet A, Mentre F, Steimer JL, Lokiec F. Nonparametric maximum likelihood estimation for population pharmacokinetics, with application to cyclosporine. J Pharmacokinet Biopharm 16:311–327 (1988).
90. Schumitzky A. Nonparametric EM algorithms for estimating prior distributions. Appl Math Comput 45:141–157 (1991).
91. Davidian M, Gallant AR. Smooth nonparametric maximum likelihood estimation for population pharmacokinetics, with application to quinidine. J Pharmacokinet Biopharm 20:529–556 (1992).
92. Mallet A. A maximum likelihood estimation method for random coefficient regression models. Biometrika 73:645–656 (1986).
93. Lindstrom MJ, Bates DM. Nonlinear mixed effects models for repeated measures data. Biometrics 46:673–687 (1990).
94. Racine-Poon A. A Bayesian approach to nonlinear random effects models. Biometrics 41:1015–1023 (1985).
95. Wakefield JC, Smith AFM, Racine-Poon A, Gelfand AE. Bayesian analysis of linear and non-linear population models by using the Gibbs sampler. Appl Stat – J Roy Stat Soc C 43:201–221 (1994).
96. Bernardo JM, Smith AFM. Bayesian Theory. New York:Wiley, 1994.
97. Gelman A, Carlin B, Stern H, Rubin D. Bayesian Data Analysis. London:Chapman & Hall, 1995.
98. Spiegelhalter DJ, Myles JP, Jones DR, Abrams KR. An introduction to bayesian methods in health technology assessment. Br Med J 319:508–512 (1999).
99. Lauritzen S, Spiegelhalter D. Local computations with probabilities on graphical structures and their application to expert systems. J Roy Stat Soc B 50:157–194 (1988).
100. Whittaker J. Graphical Models in Applied Mutivariate Analysis. Chichester:Wiley, 1990.
101. Pearl J. Causal diagrams for empirical research (with discussion). Biometrika 82:669–710 (1995).
102. Gelman A, Rubin DB. Markov chain Monte Carlo methods in biostatistics. Stat Methods Med Res 5:339–355 (1996).
103. Gilks WR, Richardson S, Spiegelhalter DJ. Markov Chain Monte Carlo in Practice. London:Chapman & Hall, 1996.
104. Monster AC, Boersma G, Steenweg H. Kinetics of tetrachlorothylene in volunteers; Influence of exposure concentration and work load. Int Arch Occup Environ Health 42:303–309 (1979).
105. Datta G, Ghosh M. Some remarks on noninformative priors. J Am Stat Assoc 90:1357–1363 (1995).
106. Kass RE, Wasserman L. The selection of prior distributions by formal rules. J Am Stat Assoc 91:1343–1370 (1996).
107. ILSI. Physiological Parameter Values for PBPK models. Washington, DC:International Life Sciences Institute, Risk Science Institute, 1994.
108. Box GEP, Tiao GC. Bayesian Inference in Statistical Analysis. New York:Wiley, 1973.
109. Smith AFM. Bayesian computational methods. Philos Trans R Soc Lond A 337:369–386 (1991).
110. Gelman A, Rubin D. Inference from Iterative Simulation Using Multiple Sequences. Stat Sci 7:457–511 (1992).
111. Cowles MK, Carlin BP. Markov chain Monte Carlo convergence diagnostics: a comparative review. J Am Stat Assoc 91:883–904 (1996).
112. Geyer CJ, Thompson EA. Annealing Markov chain Monte Carlo with applications to ancestral inference. J Am Stat Assoc 90:909–920 (1995).
113. Gilks WR, Robert C. Strategies for improving MCMC. In: Markov Chain Monte Carlo in Practice (Gilks WR, Richardson S, Speigelhalter DJ, ed). London:Chapman & Hall, 1996;89–114.
114. Tierney L, Mira A. Some adaptive Monte Carlo methods for Bayesian inference. Stat Med 18:2507–2515 (1888).
115. Sargent DJ, Hodges JS, Carlin BD. Structured Monte Carlo Markov chains. J Comput Graph Stat 9: 217–234 (2000).
116. Smith TJ. Pharmacokinetic models in the development of exposure indicators in epidemiology. Ann Occup Hyg 35:543–560 (1991).
117. Wakefield J, Bennett J. The Bayesian modeling of covariates for population pharmacokinetic models. J Am Stat Assoc 91:917–927 (1996).
118. Bois FY. Analysis of PBPK models for risk characterization. Ann NY Acad Sci 895:317–337 (1999).
119. Bois FY. Statistical analysis of Clewell et al. PBPK model of trichloroethylene kinetics. Environ Health Perspect 108(suppl 2):307–316 (2000).