Statistical Methods in Medical Research 1997; **6**: 341–358

An introduction to the use of physiologically based pharmacokinetic models in risk assessment

AJ Bailer Department of Mathematics and Statistics, Miami University, Oxford, Ohio, USA and National Institute for Occupational Safety and Health, USA and **DA Dankovic** National Institute for Occupational Safety and Health, USA

Many extrapolation issues surface in quantitative risk assessments. The extrapolation from high-dose animal studies to low-dose human exposures is of particular concern. Physiologically based pharmacokinetic (PBPK) models are often proposed as tools to mitigate the problems of extrapolation. These models provide a representation of the disposition, metabolism, and excretion of xenobiotics that are believed to possess the potential of inducing adverse human health responses. Given a model of xenobiotic disposition that is applicable for multiple species and appropriate for nonlinearity of the xenobiotic biotransformation process, better extrapolation may be possible. Unfortunately, the true structure of these models (e.g. number of compartments, type of metabolism, etc.) is seldom known, and attributes of these models (tissue volumes, partition coefficients, etc.) are often experimentally determined and often only central measures of these quantities are reported. We describe the use of PBPK models in risk assessment, the structural and parameter uncertainty in these models, and provide a simple illustration of how these characteristics can be incorporated in a statistical analysis of PBPK models. Additional complexity in the analysis of variability in the models is also outlined. This discussion is illustrated using data from methylene chloride.

1 Introduction

Quantitative risk estimation for the protection of human health is an exercise in which the best available data sources, animal experiments or human observational studies, are used to evaluate the potential of some hazard to induce an adverse human health response. The hazard may be a chemical, mixture of chemicals or some other agent while the response of interest may be cancer, decrements in lung function, teratogenicity or a host of other possibilities. At its most primitive level, quantitative risk assessment (QRA) is a regression problem – of interest is either the response associated with a specified level of exposure, or the concentration associated with a specified level of response. Thus, availability of adequate dose/concentration–response data is a prerequisite to doing a QRA.

A fundamental question in QRA revolves around the distinction between 'exposure' and 'dose'. The word 'exposure' is usually reserved for the external (outside of the organism) concentration of a toxicant while 'dose' usually refers to an internal concentration of either the toxicant itself, or a biologically active metabolite derived from the toxicant. In the absence of information concerning the relationship between exposure and dose, a QRA will proceed by studying exposure–response relationships. One means by which exposure–dose relationships are better understood is via the use of models relating these quantities.

Address for correspondence: AJ Bailer, Department of Mathematics and Statistics, Miami University, Oxford, OH 45056, USA. E-mail: ajbailer@muohio.edu

 0962-2802(97)SM143RA

If the best available data source is an animal toxicology experiment, then many questions surface regarding the use and relevance of such data for humans. Since these experiments are often conducted at high exposure levels, often much higher than exposure levels of human concern, the question of inferring low-dose behaviour from these high-dose studies arises. The animal toxicology studies typically use laboratory rodents from selectively bred strains. The differences in the biological activity of a hazardous substance in rodents versus its activity in humans is a natural concern while the notion of differential variability in humans versus inbred rodents is also vexing. Because of the use of animal data as the basis of some risk assessments, tools and techniques for addressing these dose and species extrapolation questions are needed.

Physiologically based pharmacokinetic (PBPK) models provide one such tool. In the discussion that follows, we introduce the general structure of PBPK models; discuss their use in quantitative risk assessment focusing on how PBPK models may address the extrapolation issues raised previously; describe an 'ideal' circumstance in which these models might be used; comment on the sources of uncertainty and variability that cause us to be far from the 'ideal'; describe an initial means by which these sources should be incorporated and displayed; illustrate the implication of uncertainty and variation on estimated risks; and, finally, consider how more complicated statistical tools might be needed to address these concerns. This discussion is highlighted throughout by a model of the disposition of methylene chloride.

2 What are PBPK models?

PBPK models are representations of a physiological system – usually rodents and/or humans. These representations include compartments that represent tissues or collections of tissues along with connections between these compartments. In a therapeutic context, the PBPK model is used to describe the time course of a beneficial drug while in a risk assessment context, the PBPK model is used to describe the time course of some potentially hazardous substance. The time course of the disposition of a substance in the body can be subdivided into the amount of the substance in the various compartments of interest at each point in time. A key element of this model is the dynamic, time-varying aspect of the amount of substance present in each compartment.

In order to specify such models, attributes of each compartment along with properties of the transition between compartments must be considered. In PBPK models, compartments (i.e. tissues or collection of tissues) are characterized by volumes, partition coefficients, and metabolism rates while the transition between compartments is characterized by attributes, such as blood flows and respiration rates. Given the description of compartment properties and relationships between compartments, PBPK models are specified by a collection of differential equations in which the change in the amount or concentration of the substance of interest in each compartment is expressed as the difference in the rate of introduction and assimilation of the substance and the rate of elimination and biotransformation of the substance. The 'rate of introduction' represents both blood concentrations and blood flow to the tissue while 'assimilation' represents the incorporation of the substance into the tissue as

might be expressed in terms of partition coefficients. The 'rate of elimination' represents a combination of the tissue concentration and blood flow from the tissue while 'biotransformation' represents metabolic transformation of the parent substance into metabolic products.

A typical display of a PBPK model for methylene chloride (dichloromethane, MC) is provide in Figure 1, based on a model developed by Reitz *et al.*,[1] as modified by Dankovic and Bailer.[2] In this figure, compartments are represented as boxes with inter-relationships between compartments represented as arrows. Five compartments are considered. Two of the compartments correspond to specific tissues ('lung' and 'liver') while the remaining three compartments correspond to tissue groups ('fat', 'slowly perfused', and 'rapidly perfused'). In addition, the lung compartment is subdivided into an additional component, a 'gas exchange' area. Methylene chloride enters this model via inhalation, hence the arrows entering and departing the lung compartment by ventilation. Once methylene chloride enters the lung, it redistributes

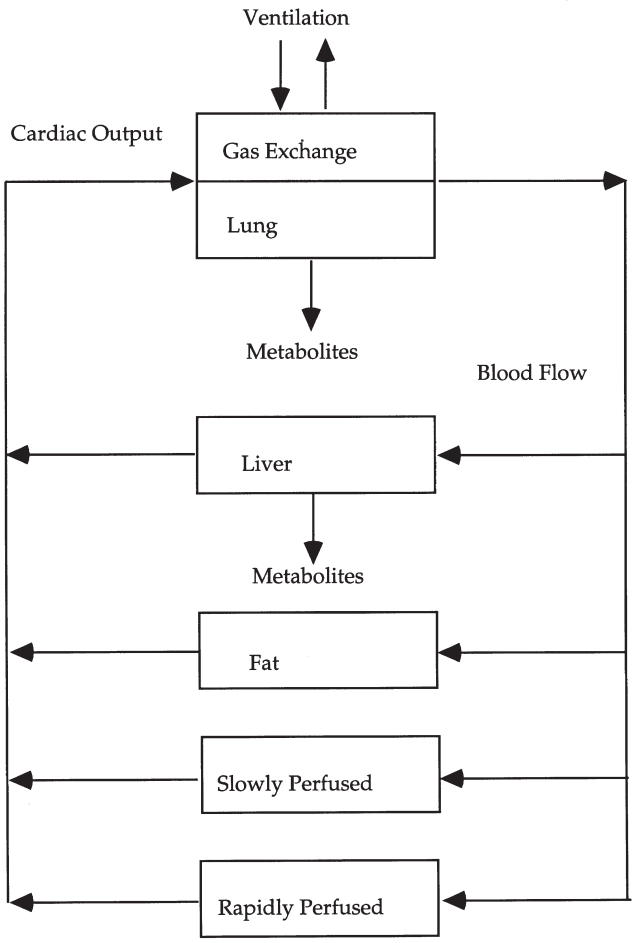

Figure 1 PBPK model for methylene chloride (based on Figure 1 in Dankovic and Bailer[2])

to the other four compartments via the blood stream. Methylene chloride either partitions into these compartments or returns to the lung through the blood stream. In the model, there are multiple ways in which MC is removed from the body – metabolism in the lung or in the liver, or exhalation in the breath. In Table 1, we identify the components of the MC PBPK model along with constituent ingredients of the differential equations that describe the time course of MC in the compartments of this model. For example, the differential equation for the change in the amount of MC in the fat compartment, A_F, can be written as

$$\frac{dA_F}{dt} = Q_F C_A - \left(\frac{[Q_F A_F]}{[V_F P_F]}\right)$$

where this equation is from the difference of the 'inflow' and 'outflow' of MC in the fat (see Table 1). The term $Q_F C_A$ is the blood flow to the fat (Q_F) times the arterial concentration of MC (C_A) while the $[Q_F A_F]/[V_F P_F]$ is comprised of: Q_F (blood flow out must equal blood flow into the tissue unless the tissue volume is expanding!); A_F, the amount of MC in the fat compartment; P_F, the partition coefficient that represents the steady-state ratio of the concentration of MC in fat to the concentration of MC in the blood; and V_F, the volume of the fat compartment. Eight other differential equations are needed to fully specify this model: four are needed for the other

Table 1 Table of methylene chloride PBPK model components, parameters, and relationships

Component	Amount	Volume	Concentration	Partition coefficient	'Inflow'	'Outflow'
Lung	A_{LU}	V_{LU}	$C_{LU} = A_{LU}/V_{LU}$	P_{LU}	$Q_C C_{AL}$	$Q_C C_{LU}/V_{LU}$ + $M1'(LU)$ + $M2'(LU)$
Fat	A_F	V_F	$C_F = A_F/V_F$	P_F	$Q_F C_A$	$Q_F A_F/V_F P_F$
Liver	A_L	V_L	$C_L = A_L/V_L$	P_L	$Q_L C_A$	$Q_L A_L/V_L P_L$ + $M1'(LI)$ + $M2'(LI)$
Rapidly perfused	A_R	V_R	$C_R = A_R/V_R$	P_R	$Q_R C_A$	$Q_R C_R/V_R P_R$
Slowly perfused	A_S	V_S	$C_S = A_S/V_S$	P_S	$Q_S C_A$	$Q_S C_S/V_S P_S$
Liver metabolite 1	M1(LI)	V_L	na	na	$k_f(LI)*C_L$	0
Liver metabolite 2	M2(LI)	V_L	na	na	$k_m(LI)*C_L/(V1_{max} + C_L)$	0
Lung metabolite 1	M1(LU)	V_{LU}	na	na	$k_f(LU)*C_{LU}$	0
Lung metabolite 2	M2(LU)	V_{LU}	na	na	$k_m(LI)*C_{LU}/(V_{max} + C_{LU})$	0

Notes:
na = not applicable.
1) Metabolite 1 = MFO pathway.
2) Metabolite 2 = GST pathway.
3) The differential equations describing the rate of change of the amount of substance, either parent compound or metabolites, are constructed from the difference between the entries in the 'inflow' and 'outflow' columns of this table.
4) M1′(L1) is the 'inflow' of metabolite 1 in the liver (similar notation is used for metabolite 2 and for both metabolites in the lung (*LU*)).
5) The parameter values were taken from Reitz *et al*.,[1] as modified by Dankovic and Bailer.[2]

compartments while four additional equations are needed for the two metabolic products of MC (in the two compartments, the liver and the lung, where metabolism occurs). Notice that the 'outflow' of the liver and lung compartments also includes the loss of MC to biotransformation to its metabolites.

3 How are PBPK models used in QRA?

PBPK models have been used for about 10 years in risk assessment, particularly in cancer risk assessment, as a means of reducing uncertainty (or perhaps, better representing uncertainty) in risk assessment. Andersen *et al*.[3] provide an excellent review of background with a rich set of references including various compounds that have been assessed using these models. Frederick[4] provides a table in which he describes the sources of uncertainty in risk assessment, including test species, dose measure and expression, selection of tumour response in test species, dose–response model, interspecies extrapolation, human dose, target human, exposure concentration and environmental exposure (see Table 1 in Frederick[4]). The selected test species should be the best model of the effects of the substance of interest in humans. The use of a dose–response model begs the questions of the best dose representation and most appropriate response. The 'dose' used in the dose–response analysis should be the appropriate metric for the extrapolation across species and to lower doses. Often, it is a metabolic product of some parent compound interacting with some internal tissue group that causes a toxic response. If this is true, then 'dose' for a dose–response analysis might be defined as the level of a metabolite observed in this internal tissue. As an aside, this type of dose is sometimes called a *delivered dose*, *target-tissue dose* or *biologically effective dose*. Thus, the proper tumour response for comparing a mammalian model to humans is a basic biomedical question and is intimately tied to the selection of the most appropriate dose measure. Finally, what are the characteristics of the human population that is assessed for risk? The sources of uncertainty in risk assessment which can be addressed by PBPK models can be summarized into a few broad themes: species/strain extrapolation, dose extrapolation, and route extrapolation.

Laboratory animals, used as the basis of risk assessment, are often inbred strains of certain species of rodents. Obvious differences exist between humans and rodents – organism size, tissue volumes, ventilation rates, etc. – while more subtle differences also exist – differences in metabolic rates, the existence of organs that are not in humans (e.g. Zymbal gland). What is the basis for using animal data for assessing human risk? In other words, how do we address the so-called problem of *species extrapolation*? One strategy that is commonly advocated[4,5] is to determine an 'equivalent' dose between rodents and humans where 'equivalent' implies that the dose–response relationship is identical across species if dose is measured on this scale. PBPK models are used under the assumption that species differences in metabolism, physiology (blood flow, tissue size, respiration, etc.) are, at least partially, the determinants of cross-species dose equivalency. To illustrate, one model would be developed for mammals in which the only difference between species would be the values of the parameters of this model.

In the context of applying a MC PBPK model developed for rodents to humans, the calculation of an appropriate dose for humans involves only changing the parameters (Q, P, etc. as given in Table 1) of the PBPK model from the rodent values to the human values. Although many of the parameter values are well known (i.e. physiological parameters) or generally similar across species (partition coefficients), parameter values for metabolism of foreign compounds do not vary across species in a readily predictable manner.[6] Although there is a general tendency toward lower metabolic rates in larger species (with maximal rates of metabolism varying roughly in proportion to body weight to the three-quarters power[7,8]), parameter values describing metabolic rates should be determined experimentally whenever possible.

Toxicology studies in rodents are typically conducted at fairly high exposure levels – often much higher exposure levels than are expected in human occupational or environmental exposures. Thus, how can studies conducted with only high-dose experimental groups be used to infer responses at low dose levels? This is the problem of *low-dose extrapolation*. Of particular concern is the notion that a nonlinear relationship exists between the exposure to a hazard and delivered doses at target tissues. This nonlinearity may be a result of the saturation of clearance mechanisms prior to the formation of putative constituents of the hazard. This can be reflected in the PBPK model by incorporating nonlinear clearance and/or metabolic rates.

In the MC PBPK model, nonlinear metabolism is used for the mixed function oxidase (MFO) pathway ('metabolite 1' in Table 1). Figure 2 illustrates the PBPK model predictions for the relationship between MC exposure concentration and the formation of MC metabolites in the $B6C3F_1$ mouse. For low exposures, the rate of MFO metabolism is approximately proportional to exposure, while at high exposures the rate of MFO metabolism asymptotically approaches its maximal rate. The second metabolic pathway for MC, the glutathione-*S*-transferase (GST) pathway, is essentially nonsaturable in the exposure range of interest.[1] As shown in Figure 2, the exposure–dose relationship for the GST metabolite of MC is also curved, due to competition between the two metabolic pathways. In this case, nonlinearity in MC metabolism by the GST pathway reflects the saturation of the MFO pathway, such that a larger percentage of the inhaled dose is metabolized via the GST pathway ('metabolite 2' in Table 1) at high exposures. This nonlinearity in GST metabolism has obvious implications for high-dose to low-dose extrapolation of risk estimates in the $B6C3F_1$ mouse.

Metabolism at low exposures is frequently controlled by transport limitations, rather than enzymatic capabilities, resulting in a linear relationship between exposure and metabolism. This linear relationship breaks down abruptly when the exposure is sufficiently high that metabolic capability becomes the rate-limiting factor in metabolic clearance.[6] Nonlinearity in metabolism due to saturation of metabolic clearance is frequently observed at the high exposures used in animal toxicology studies. The expected effect of this nonlinearity on high-exposure to low exposure extrapolation of toxicity depends on whether toxicity is mediated by metabolism. In the case of agents whose toxicity is mediated by a capacity-limited pathway of metabolic activation, one might anticipate a greater probability of harm at low exposures (in the test species) than would be predicted by a simple linear extrapolation from high exposure to low. This issue is very important because the parent compound or any metabolic product could be used as a 'dose' in a QRA.

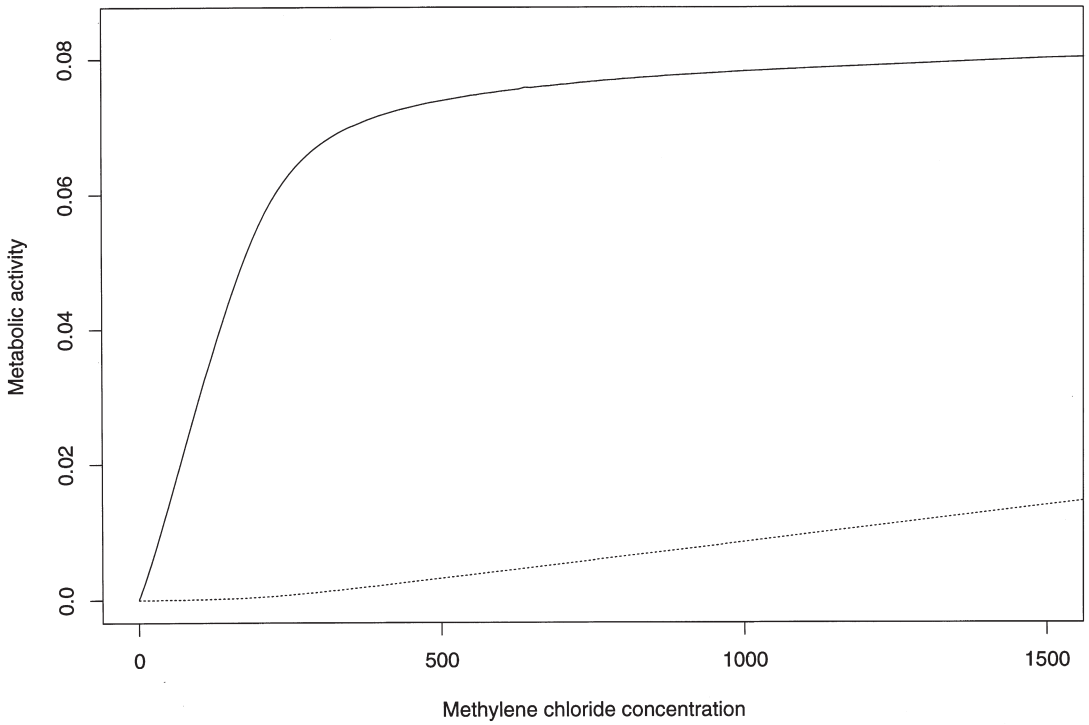

Figure 2 Cumulative MFO (solid line) and GST (dashed line) metabolite production (μmoles/day) in the liver of a B6C3F$_1$ mouse exposed to MC (ppm) for 6 h and then followed for and additional 18 h. Based on the Reitz *et al*. PBPK model for MC[1]

The last issue of concern in using animal studies for describing human risks is that animal studies are often conducted using a route of exposure differing from the route of exposure which is of interest for humans. For example, many toxicology experiments may use a gavage route of exposure with the hazard introduced in a corn oil medium while the typical human exposure is via inhalation or dermal contact. Given a PBPK representation of a mammalian model which describes the disposition of a hazard once it enters the system, the actual route of entry to the model becomes a simple addition to the model. For example, if exposure might occur via inhalation, gavage, or dermal contact, then separate equations relating these routes of exposure to the models can be included. In a PBPK model, different routes of exposure are incorporated by different routes of entry of the hazard into the model. The MC PBPK model, as shown in Figure 1, allows for only inhalation exposure. To generalize this model to another route of exposure, say through a gavage exposure, additional information concerning absorption of MC into the blood from the GI tract is needed. Once the chemical is in the blood, the biotransformation and disposition of the chemical proceeds as before.

In addition to the extrapolation issues discussed above, PBPK models can be used to assess the impact of intermittent exposure. Intermittent exposure is commonly encountered in occupational settings. For example, workers often will have approxi-

mately 8 h of exposure per day for a five-day work week. Between each period of exposure is 16 h of zero exposure. Further, the weekends provide a two-day break in possible exposure. The PBPK model for MC incorporates this intermittent exposure pattern by setting the MC concentration in the air to zero for all nonwork-shift times. In addition, these models can be used to evaluate doses to retired workers exposed to chemicals with long half-lives or that sequester in certain compartments.

4 'Ideals' for both QRA and PBPK models

What data provide the best support for a QRA, and if PBPK models were used as part of this process, what information would make these models most useful? The following list presents such information:

1) The putative agent is identified, and is the same in both the test species and the target species (usually humans). For example, if a metabolite is the carcinogenic species of a particular hazard, then levels of this metabolite would be used as the 'dose' in the dose–response models.
2) The dose of the putative agent at all points in time is known (including times when exposure is not occurring).
3) The true relationship between the dose of the putative agent and the response of interest in the test organism is known.
4) Pharmacodynamic differences between species are well understood so that extrapolation between different species was well understood.

If the QRA for humans is based on animal experiments, then (ideally) the relationship between dose and response is the same in both species assuming that an appropriate dose has been identified. Thus, the availability of a PBPK model may prove quite important for resolving species differences, by facilitating comparison of the exposure–dose relationship in the test species to that in the target species (usually human). It would be ideal if all aspects of a PBPK model were determined independently of the exposure–response data to which it will be applied. For example, if a PBPK model is developed to derive a delivered dose estimate for use in modelling a tumourigenicity experiment, then it would be desirable if the parameters of the PBPK model (including tissue volumes, partition coefficients, metabolic rates, etc.) were determined separately from and independently of the response data. As a precursor to the application of a PBPK model in QRA, the quality of the model would be evaluated and validated using external data sources. In particular, if a PBPK model is to be used for extrapolating from rodents to humans, then greater confidence may be placed in this model if it is shown to describe the time course of a hazard in multiple species.

5 What are the unknowns in using PBPK models in QRA?

In previous sections, we mentioned one view of uncertainty as tied to the extrapolation process in risk assessment. We expand on this discussion by describing uncertainty in PBPK models from a systematic and from a stochastic perspective. The systematic perspective relates to the structural components of the model. Questions relating the

structural aspects of the model include:[9,10] how many compartments should be in the model? In which compartments does metabolism occur? Is metabolism linear or non-linear? What is the proper dose metric for use in a risk assessment? The stochastic component of the model relates to the fact that parameters of PBPK models are estimated from data. In addition, these parameters exhibit *variability* in the sense that they naturally exhibit a distribution of possible values in a population of organisms and these quantities are measured with error.

5.1 Number of compartments

The choice of the number of compartments in a model is not completely arbitrary. Although it is tempting to include as many individual tissues in the model as possible, in practice some grouping of tissues is necessary for mathematical tractability, and to reduce the data requirements for model construction to a feasible level. In general, even the most minimal PBPK model (for aqueous or lipid soluble substances) will generally contain at least three compartments: rapidly perfused tissues, slowly perfused tissues and fat. These tissue groupings are chosen by virtue of similarity in flow-rate to volume ratio and chemical composition – properties which jointly govern the rate of equilibration between chemicals transported in the blood and those sequestered in tissues.[11] A model with only these three compartments was used successfully for modelling occupational exposures to benzene.[12]

We demonstrate the implications of the choice of tissue groupings in Figure 3, which illustrates the concentration–time profile for MC in venous blood, using three separate PBPK models with either three or five compartments. The first example is a five-compartment model with MC metabolism occurring in both the lung and the liver, as shown in Figure 1 and with results displayed in Figure 3 (see the solid line with plotting symbol '5'). The second example is a three-compartment model which collapses the lung and liver compartments into the rapidly perfused tissue compartment (see the dashed line with plotting symbol '3' in Figure 3). In this example, the metabolic capability of the liver was assigned to the rapidly perfused tissue group, and the metabolic capability of the lung was simply ignored! Notice that this three-compartment model yields a concentration–time profile similar to that of the more complex five-compartment model, and would in all likelihood provide an adequate representation of the venous concentration of MC if that were the dose metric of interest – for example, for biomonitoring purposes. In contrast, the third example is a three-compartment model with a separate lung and liver, with the same metabolic capabilities as the lung and liver in the five-compartment model. However, the rapidly perfused, slowly perfused, and fat tissue groups have been inappropriately lumped together. As displayed as a dotted line with symbol 'L' in Figure 3, this third model is unable to reproduce the concentration–time profile of the five-compartment model. It should be noted that the second model, though adequate for representation of the concentration of parent MC, cannot provide an estimate of MC metabolized by the lung, which may be of interest in modelling the carcinogenicity of MC. This serves to illustrate one of the most common reasons for considering inclusion of additional tissues in the model, beyond the minimal set described above: to provide internal dose

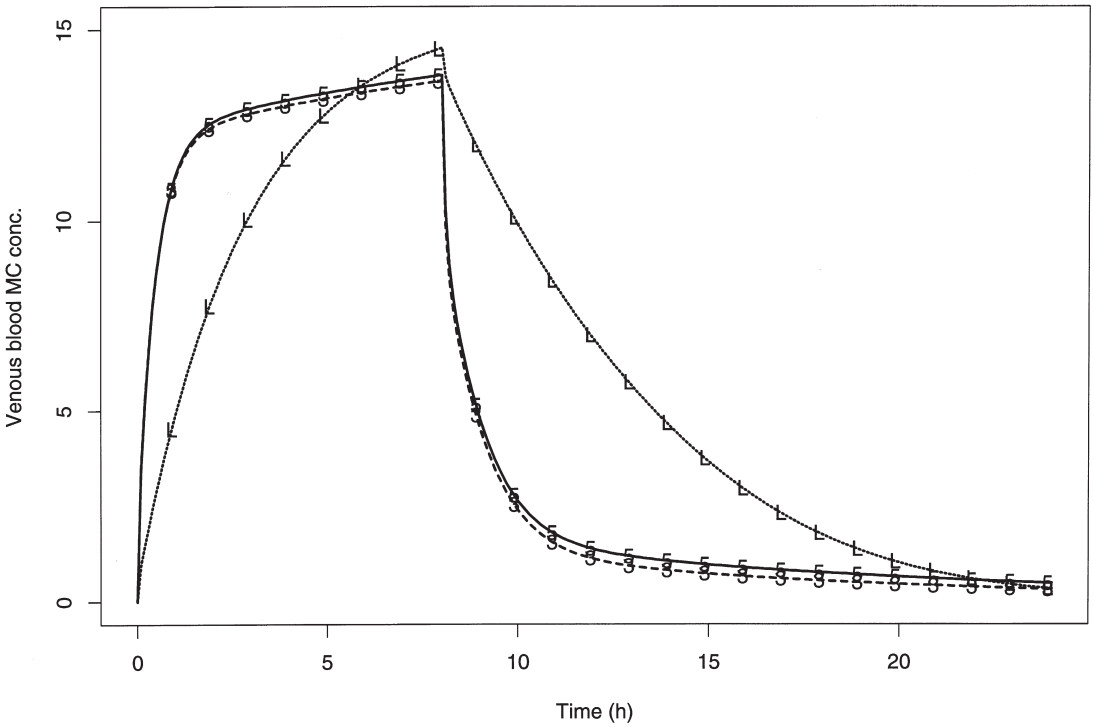

Figure 3 Venous blood methylene chloride concentration as a function of time for three different PBPK model formulations (solid line, symbol '5' = five-compartment model; dashed line, symbol '3' = three-compartment model; and dotted line, symbol 'L' = three-compartment 'lumped' model)

estimates for specific target tissues. As with any modelling exercise, parsimony in model complexity aids in the interpretation of a model; however, this illustration suggests that an overly simple model may completely lose its utility.

5.2 Systematic variability

In discussions of PBPK models, consideration of variability often relates to the predictions from the models. In particular, do we wish to predict the behaviour of a PBPK model for some typical/'median' individual or for members of sensitive subpopulations? We illustrate the differences in total metabolite production based on the five-compartment model of MC. In this illustration, we have constructed several different 'prototypes' of exposed humans:

1) median body fat, regular MFO and GST metabolism;
2) median body fat, fast MFO metabolism, regular GST metabolism;
3) median body fat, slow MFO metabolism, regular GST metabolism;
4) smaller fat compartment, slow MFO and fast GST metabolism;
5) larger fat compartment, fast MFO metabolism, slow GST metabolism.

A body fat content of 23.1% was used as 'median' (as in the original Reitz *et al.* model for methylene chloride[1]); 10% body fat was used in the 'lean' simulation, and

35% body fat in the 'obese' simulation. The MFO and GST parameter values were largely based on those estimated by Dankovic and Bailer.[2] For the 'fast' MFO simulations, V_{max} was multiplied by 2.0, and K_m was multiplied by 0.44, relative to the 'regular' MFO parameter values. For the 'slow' MFO simulation, V_{max} was multiplied by 0.23, and K_m was multiplied by 0.95, relative to the 'regular' MFO parameter values. The K_f parameter governing GST metabolism was multiplied by 1.44 for the 'fast GST' simulation, and by 0.5 for the 'slow GST' simulation, relative to the 'regular GST' parameter value. Recall that all parameters for this model are available in the Reitz *et al.*[1] model with modified parameter values given in Dankovic and Bailer.[2] Figure 4 displays the time course of metabolite levels in these 'individuals'.

Figure 4 shows that GST metabolism in the liver is greatest for those individuals with slow MFO metabolism, rapid GST metabolism, and low body fat (i.e. have a fat compartment with small volume). The smallest level of GST metabolism in the liver is observed in relatively obese (high fat volume) individuals with fast MFO and slow GST metabolism. The implication of these observations is that the metabolic and

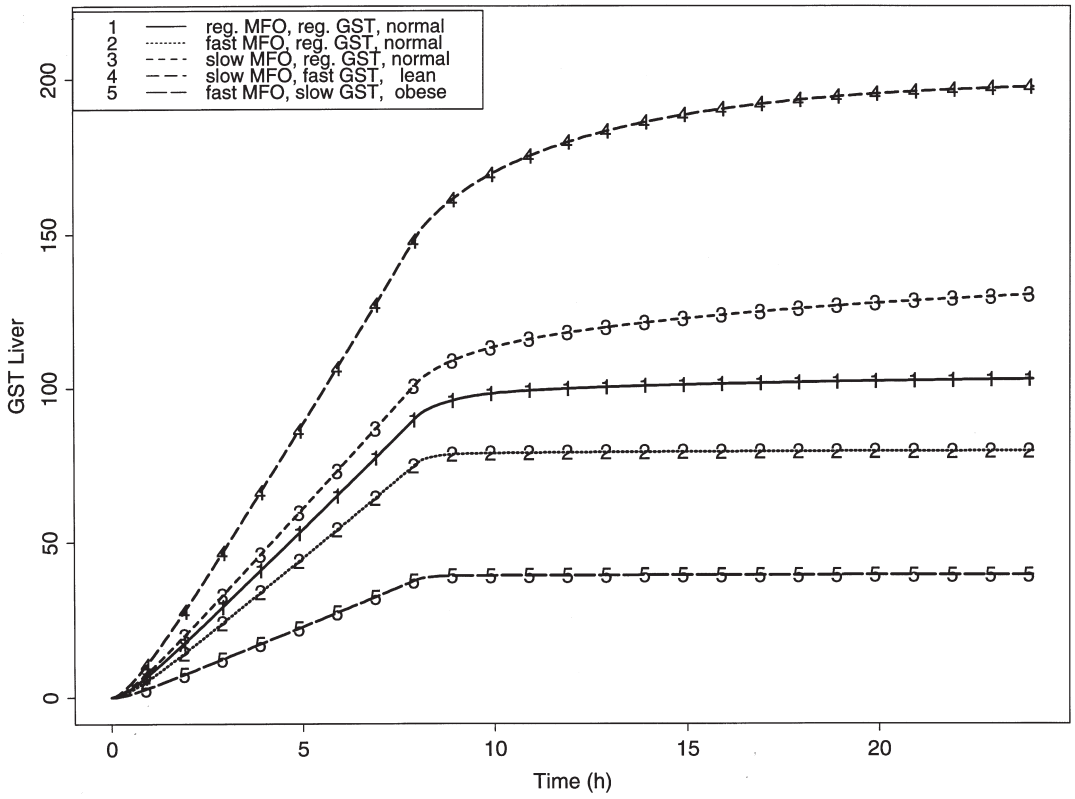

Figure 4 GST metabolite production (μmoles/day) in the liver for individuals with different physiological characteristics based on the five-compartment model. Assumes 8-h exposure at 500 ppm under working conditions. Body types are classified as normal (23.1% body fat), lean (10% body fat); and obese (35% body fat). MFO metabolism is classified as normal, fast (V_{max} = 2.0∗normal, K_m = 0.44∗normal), or slow (V_{max} = 0.23∗normal), K_m = 0.95∗normal). GST metabolism is classified as normal, fast (1.44∗normal), or slow (0.5∗normal).

physiologic characteristics of an individual may have strong influence on the production of hazardous metabolites. In the five 'prototypes' considered above, the ultimate production of GST liver metabolites differed by almost a factor of five at its most extreme.

In addition to physiological differences between humans that might affect the metabolism of MC, the conditions under which MC exposure occurs may also alter this response systematically. The difference in metabolic output was explored for 'working' (higher respiration, blood flow condition) and 'resting' states. This is a relevant exercise for exploring occupational exposure conditions to a particular hazard. The amount of GST metabolite produced in the liver based upon applying the five-compartment model (see Figure 1) is displayed for these two conditions in Figure 5.

In Figure 5, working conditions (solid line, plotting symbol 'W') lead to higher levels of GST metabolism in the liver relative to resting conditions (dotted line, plotting symbol 'R'). A more detailed analysis of metabolite production under differing work conditions can be found in Dankovic and Bailer.[2]

A natural statistical question is how might we estimate population risk from such an assessment of individual exposure scenarios. One possibility is to classify exposed individuals according to a reasonable subgroups of similar traits. Given individuals of similar traits (i.e. can be described by the same PBPK model), estimate their exposure levels (or possibly risks). The population effect estimate would then be constructed

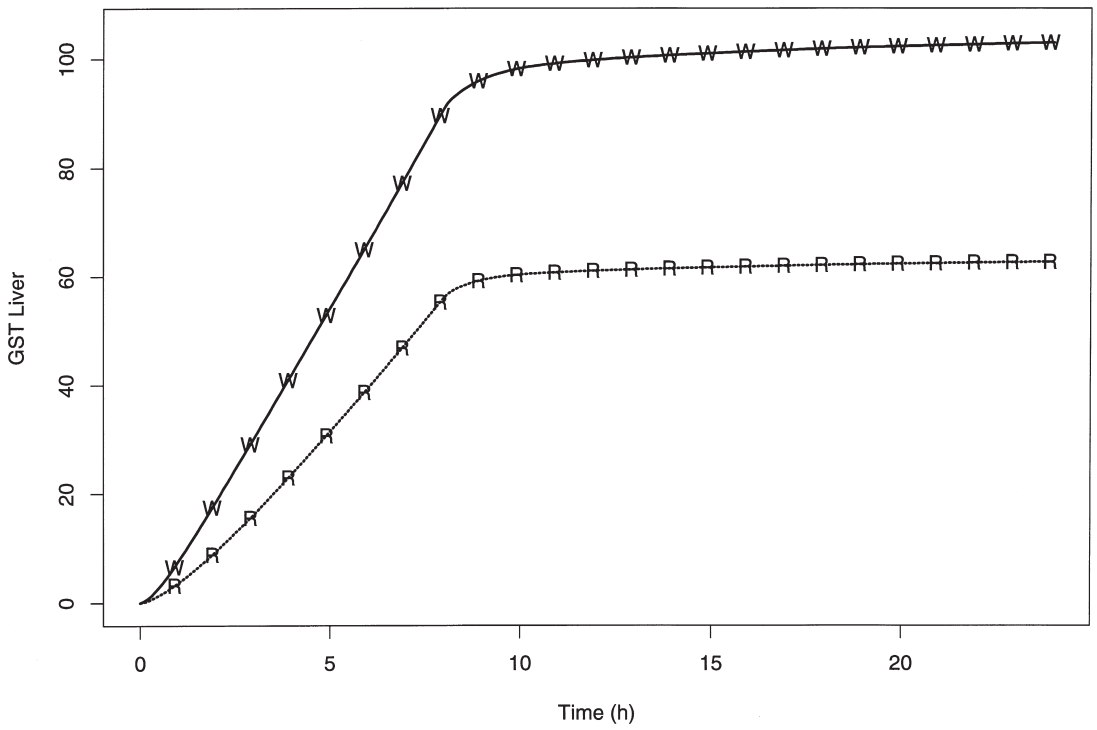

Figure 5 GST metabolism production in the liver (μmoles/day) for working or resting physiological conditions based on the five-compartment model. Assumes 8 h exposure at 500 ppm.

from combining the estimates from the various scenarios weighted by their relative likelihood of occurrence in the population. Frederick[4] employed different exposure scenarios to develop a composite estimate of population risk.

5.3 Monte Carlo simulation

We note that even given individuals in the same subcategory of similar tissue size, blood flows, etc., variability in response would be expected. Variability of attributes within a PBPK model can be incorporated in a variety of ways including the use of Monte Carlo methods (see, for example, Portier and Kaplan[5]). In this exercise, the attributes of the PBPK models are assumed to possess a distribution in the population at large (see, for example, Lipton *et al*.[13]). The nature of this distribution should reflect the physiological constraints of the model. For example, blood flows must be non-negative and the total blood flow to all tissues in the model must be equal to the total cardiac output. This suggests that variability in both the total cardiac output and in its partitioning into the flows to all tissue groups must be reflected in any uncertainty analysis of the flow to tissues. In our example, this could be accommodated by partitioning the total cardiac flow, Q_C, into tissue flows by generating a vector of multinomial probabilities where each cell probability corresponds to the proportion of Q_C going to a particular tissue or tissue group.

For illustration, we have imposed a normal distribution on the metabolic rates, blood flows, and tissue volumes in the five-compartment model for the mouse, using mean parameter values taken from Reitz *et al*.,[1] and standard deviations based on literature values: Reitz *et al*. for the metabolic parameters;[1] Gargas *et al*. for the partition coefficients;[14] and Farrar *et al*.[15] for tissue volumes and blood flows. One thousand simulations were run in order to estimate the hepatic dose of the GST metabolite (metabolite 2 in Table 1) formed from a 6-h exposure to 2000 ppm MC. A relative frequency histogram of the 1000-dose estimates is shown in Figure 6. The results suggest a median dose of 9.18 μmoles/day/ml of liver, with the fifth and 95th percentiles of the distribution being 7.51 and 10.91, respectively.

Note that Monte Carlo methods raise questions concerning the statistical distributions used to describe variability, and the appropriate measures of dispersion (i.e. standard deviation versus standard error, in the normal distribution). In general the true distributional forms of the parameters in PBPK models are unknown; typically, normal, log-normal or uniform distributions are assumed. The question of the use of standard error versus standard deviations relates to the intended impact on the model. For example, standard errors of mean model attributes relate to the sampling distribution of this summary statistic while standard deviations relate to the variation of an attribute in the population. Farrar *et al*.[15] suggested that the uncertainty measures of interest for the test species are those for the average parameter values, because the PBPK dose estimates will in general be applied to an aggregate response. For example, the standard error or some other summary measure of the sampling distribution of a statistic such as the proportion of animals in a group which develop tumours is considered. In contrast, uncertainty estimates for humans are generally applied to questions of interindividual variability, so that the relevant measures of dispersion are those for individuals rather than groups.

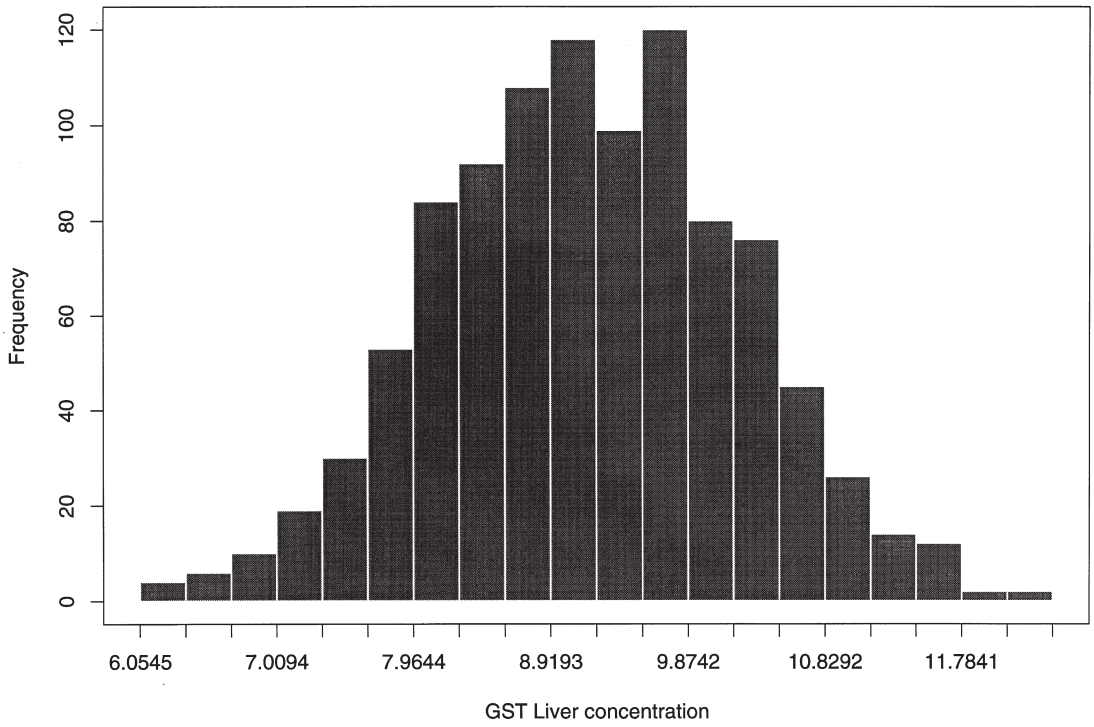

Figure 6 Histogram of PBPK model predicted GST liver metabolism (μmoles/day/ml of liver) generated in a Monte Carlo simulation of 2000 ppm methylene chloride exposure in mice

6 Implications of PBPK parameter variability on QRA?

Finally, and perhaps most importantly, what are the implications of these models on the estimates of risk for humans? We need to start with a model relating delivered dose to response. In our example, and in much of our discussion, we focused on a quantal response, the presence/absence of a tumour or cancer. A host of binary regression models are available including logistic and probit regression; however, we focus on a model that has been employed for many years for carcinogenic risk estimation – the so-called quantal multistage model. While the development of this model was motivated by a simple mechanistic concept of carcinogenesis, it can be viewed as a generalized linear model with a binomial response distribution and a complementary log link and with non-negative constraints on the parameters. The form of this model with two stages can be written as

$$\pi(d) = 1 - \exp\{-\theta_0 - \theta_1 d - \theta_2 d^2\}$$

where $\pi(d)$ is the probability of cancer in organisms exposed to dose d and the regression parameters, θ, are all constrained to be non-negative. We begin with data from a long-term carcinogenicity study of MC conducted by the National Toxicology Program.[16] The MC exposures in this study were 0, 2000 and 4000 ppm, and the observed quantal responses (number of liver adenomas or carcinomas/number of mice

on test) in female mice associated with these exposure concentrations were 3/50, 16/48 and 40/48, respectively. As described above, simulations of the PBPK model were conducted to generate 1000 different estimates of the daily GST liver metabolite dose associated with 2000 and 4000 ppm exposure in mice. The multistage model was fit to each of the 1000 data sets defined by the simulated doses and the associated quantal responses. Human occupational exposure at 25 ppm was converted to a daily GST liver dose. Given this human GST dose and the 1000 regression model fits, we estimated the excess risk as $P(d) - P(0)$ where d corresponds to the human GST dose and $P(d)$ is the model predicted probability of tumour at dose d. The distribution of these excess risks estimates are presented in Figure 7.

In Figure 7 we see that an 8-h occupational exposure to 25 ppm MC for a working lifetime (assumed to be five days per week for 45 years over a 70-year lifespan) suggests a median excess risk of 6.94×10^{-8} with the fifth and 95th percentiles of the excess risk distribution being 4.72×10^{-8} and 9.03×10^{-5}, respectively. Simple percentile calculations may not be an adequate summary of the bimodal distribution of added risks observed in Figure 7. The bimodal appearance of this figure reflects two classes of multistage model fits which were observed: those with a zero coefficient for the linear dose term (approximately 60% of the model fits with corresponding risk estimates less than 10^{-7}) and those with a positive linear dose term (approximately 40% of the models with corresponding risk estimates between 10^{-7} and 10^{-3}).

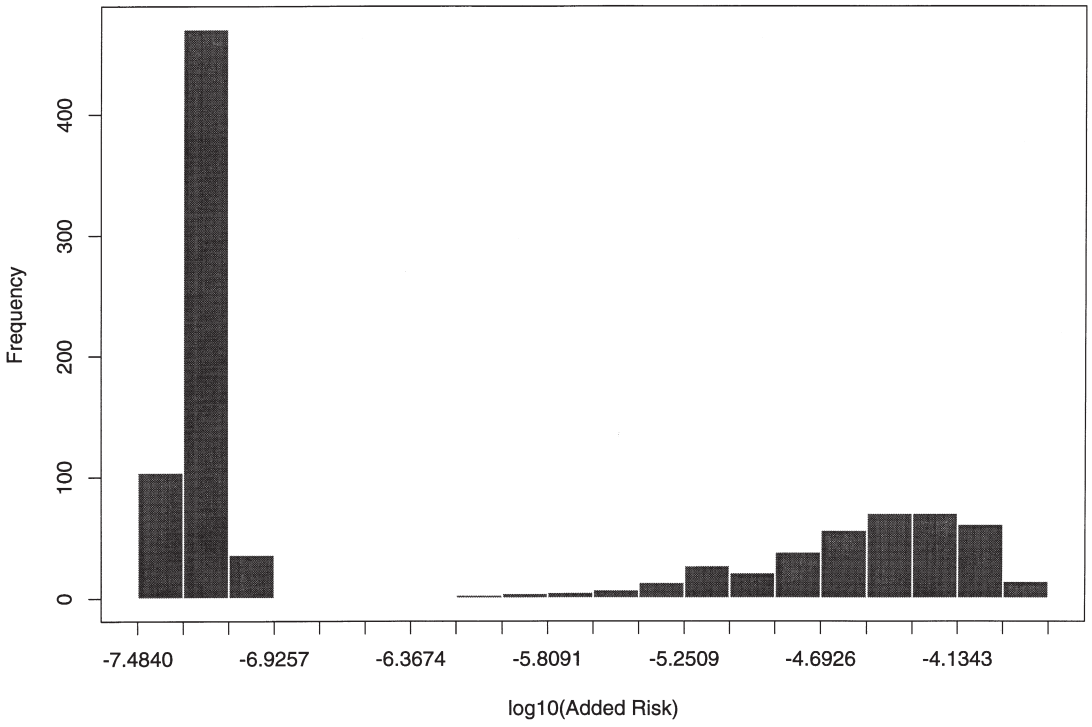

Figure 7 Histogram of excess risk estimates for humans exposed occupationally to 25 ppm MC based on a multistage model for mice using GST metabolite in the liver as a dose metric

(Essentially, 40% of the simulated delivered doses yielded a spacing of doses in which a significant θ_1, linear, term was obtained in the $\pi(d)$ regression model.) Thus, although the median dose estimate from the PBPK model suggests that the risk of liver tumours is less than 10^{-7}, the Monte Carlo analysis suggests a substantial probability that the risk may in fact be one to three orders of magnitude greater than the median estimate.

The analysis described above could be extended in a variety of ways. We restricted our attention only to the variability of the mouse delivered doses. While this provides simplicity in presentation, other variation could and should be considered in these models. First, we accepted the tumour responses as given. A more general simulation would also generate random bioassay results in which the tumour responses could also have been regenerated to reflect the binomial sampling variability in these quantities, see Portier and Kaplan.[5] Second, in the simulation reported above we converted the human 25 ppm exposure into a delivered dose of GST metabolite to the liver. As an alternative, Monte Carlo methods could be employed in conjunction with the PBPK human model to generate a GST metabolite to the liver dose based on lifetime exposure to 25 ppm MC. Thus, a more general simulation would involve 1000 separate simulated rodent metabolite doses for 2000 and 4000 ppm, randomly paired with 1000 separate simulated bioassay results which in turn yielded 1000 model fits. Then, 1000 separate human metabolite doses would be randomly paired with the 1000 model fits yielding 1000 separate excess risk estimates for humans.

7 Extensions

The host of parameters describing the MC PBPK model were all simulated as jointly independent. While this has been a common assumption in many PBPK simulation studies, it may not be a reasonable assumption. For example, it appears reasonable that the five tissues volumes would be positively correlated, i.e. individuals with a large liver may also tend to have a large rapidly perfused compartment. Assuming we can determine the population variances and covariances between all of the parameters of the PBPK model and assuming that the parameters (possibly log-transformed) are normally distributed, then this information can readily be incorporated into the simulation. Suppose μ is a vector of means and Σ is the matrix of variances and covariances for all of the PBPK parameters. Then, sets of parameter estimates for each PBPK model run could be generated as $Y = \mu + Lz$, where $\Sigma = LL'$, where Y is the vector of generated PBPK parameters, z is a vector of standard normal deviates, and L is a lower triangular matrix obtained from a Cholesky decomposition of Σ.[18] One concern in employing this analysis is that the covariances of these quantities may be unknown. While covariances between tissue volumes may be available in both rodents and humans, covariances for partition coefficients and metabolic rate constants are likely to be unavailable. In fact, the variances for many of the parameters in these models are often established by informal rules, for example a variance associated with a coefficient of variation of 10% might be employed with minimal data to support this selection.

If distributions of parameters can be well described in the PBPK model and additional experimental data are available for updating these distributions, then a

Bayesian analysis strategy[19–22] may provide an interesting and natural step for incorporating parameter variability into these models. In the context of PBPK models, a Bayesian analysis might employ a hierarchical population model in which a parameter in the PBPK model is sampled from some distribution that might be described by hyperparameters.[23] In our example, prior distributions could be assigned to the means and variances of the PBPK model parameters. One benefit of this strategy, as suggested by Gelman *et al.*,[20] is that the quality of the specification of prior distributions can be assessed by comparing the posterior and prior distributions. In the analysis we describe above, we are not updating prior distributions on the PBPK model parameters in light of any new data. In this case, the posterior distributions of the parameters would be identical to their prior distributions.[21]

8 Summary

In this paper we have focused on presenting how PBPK models are used to address some of the challenges inherent in evaluating the risks associated with exposure to potential hazards in the workplace or in the environment. We present the basic structure of PBPK models, and then discuss the uncertainty associated with the definition of these models along with consideration of how the variability in both model parameters and population characteristics should be expressed in light of these models. A next step after a basic introduction, such as we present, is to pursue some of the advanced methods that are now being employed in the analysis of PBPK models. Monte Carlo methods, either with or without Bayesian updating, offer one approach for extending the basic PBPK modelling technique to address questions of parameter and population variability.

Acknowledgements

The authors would like to thank Jim Deddens, Eileen Kuempel and Mark Toraason for their suggestions and comments on a previous draft of this manuscript.

References

1 Reitz RH, Mendrala AL, Park CN, Andersen ME, Guengerich FP. Incorporation of *in vitro* enzyme data into the physiologically-based pharmacokinetic model for methylene chloride: implications for risk assessment. *Toxicology Letters* 1988; **43**: 97–116.

2 Dankovic D, Bailer AJ. The impact of exercise and intersubject variability on dose estimates for dichloromethane derived from a physiologically-based pharmacokinetic model. *Fundamental and Applied Toxicology* 1994; **22**: 20–25.

3 Andersen ME, Krewski D, Withey JR. Physiological pharmacokinetics and cancer risk assessment. *Cancer Letters* 1993; **69**: 1–14.

4 Frederick CB. Limiting the uncertainty in risk assessment by the development of physiologically based pharmacokinetic and pharmacodynamic models *Toxicology Letters* 1993; **68**: 159–75.

5 Portier CJ, Kaplan N. Variability of safe dose estimates when using complicated models of the carcinogenic process, a case study: methylene chloride. *Fundamental and Applied Toxicology* 1989; **13**: 533–44.

6 Andersen ME. Flow-limited clearance. In: Fiserova-Bergerova V ed. *Modeling of inhalation exposure to vapors: uptake, distribution, and elimination*, vol II. Boca Raton, FL: CRC Press, 1983: 67–95.

7 Boxenbaum H. Inter-species scaling, allometry, physiologic time, and the ground plan of pharmacokinetics. *Journal of Pharmacokinetics and Biopharmaceutics* 1982; **10**: 201–27.
8 O'Flaherty EJ. Interspecies conversion of kinetically equivalent doses. *Risk Analysis* 1989; **9**: 587–98.
9 Edler L, Portier CJ. Uncertainty in physiological pharmacokinetic modeling and its impact on statistical risk estimation of 2,3,7,8 TCDD. *Chemosphere* 1992; **25**: 239–42.
10 Woodruff TJ, Bois FY, Auslander D, Spear RC. Structure and parameterization of pharmacokinetic models: their impact on model predictions. *Risk Analysis* 1992; **12**: 189–201.
11 Fiserova-Bergerova V ed. Physiological models for pulmonary administration and elimination of inert vapors and gases. In: *Modeling of inhalation exposure to vapors: uptake, distribution, and elimination*, vol I. Boca Raton, FL: CRC Press, 1983: 73–100.
12 Fiserova-Bergerova V, Vlach J, Singhal K. Simulation and prediction of uptake, distribution, and exhalation of organic sovents. *British Journal of Industrial Medicine* 1974; **31**: 45–52.
13 Lipton J, Shaw WD, Holmes J, Patterson A. Short communication: selecting input distributions for use in Monte Carlo simulations. *Regulatory Toxicology and Pharmacology* 1995; **21**: 192–98.
14 Gargas ML, Clewell HJ, Andersen ME. Metabolism of inhaled dihalomethanes *in vivo:* differentiation of kinetic constants for two independent pathways. *Toxicology and Applied Pharmacology* 1986; **82**: 211–23.
15 Farrar D, Allen B, Crump K, Shipp A. Evaluation of uncertainty in input parameters to pharmacokinetic models and the resulting uncertainty in output. *Toxicology Letters* 1989; **49**: 371–85.
16 National Toxicology Program. National Toxicology Program technical report on toxicology and carcinogenesis studies of dichloromethane in F-344/N rats and B6C3F$_1$ mice (inhalation studies). NTP-TR-306, 1985.
17 Thomas RS, Bigelow PL, Keefe TJ, Yang RSH. Variability in biological exposure indices using physiologically based pharmacokinetic modeling and Monte Carlo simulation. *American Industrial Hygiene Association Journal* 1996; **57**: 23–32.
18 Dagpunar J. *Principles of random variate generation*. Oxford: Clarendon Press, 1988.
19 Bois FY, Gelman A, Jiang J, Maszle DR, Zeise L, Alexeef G. Population toxicokinetics of tetrachloroethylene. *Archives of Toxicology* 1996; **70**: 347–55.
20 Gelman A, Bois FY, Jiang J. Physiological pharmacokinetic analysis using population modeling and informative prior distributions. *Journal of the American Statistical Association* 1996; **91**: 1400–12.
21 Bois FY, Smith AE. Evaluation of variability in predictions of cancer risk from occupational exposure to methylene chloride as a consequence of variability in model input parameters. Final report prepared for the Health Standards Program, Occupational Safety and Health Administration. Washington, DC: US Department of Labor, 1995.
22 Wakefield J. The Bayesian analysis of population pharmacokinetic models. *Journal of the American Statistical Association* 1996; **91**: 62–75.
23 Gelman A, Carlin JB, Stern HS, Rubin DB. *Bayesian data analysis*. London: Chapman & Hall, 1995.