Statistical Analysis of Fisher et al. PBPK Model of Trichloroethylene Kinetics

Frédéric Yves Bois

Institut National de l'Environnement Industriel et des Risques (INERIS), Verneuil-en-Halatte, France

Two physiologically based pharmacokinetic models for trichloroethylene (TCE) in mice and humans were calibrated with new toxicokinetic data sets. Calibration is an important step in model development, essential to a legitimate use of models for research or regulatory purposes. A Bayesian statistical framework was used to combine prior information about the model parameters with the data likelihood to yield posterior parameter distributions. For mice, these distributions represent uncertainty. For humans, the use of a population statistical model yielded estimates of both variability and uncertainty in human toxicokinetics of TCE. After adjustment of the models by Markov chain Monte Carlo sampling, the mouse model agreed with a large part of the data. Yet, some data on secondary metabolites were not fit well. The posterior parameter distributions obtained for mice were quite narrow (coefficient of variation [CV] of about 10 or 20%), but these CVs might be underestimated because of the incomplete fit of the model. The data fit, for humans, was better than for mice. Yet, some improvement of the model is needed to correctly describe trichloroethanol concentrations over long time periods. Posterior uncertainties about the population means corresponded to 10–20% CV. In terms of human population variability, volumes and flows varied across subject by approximately 20% CV. The variability was somewhat higher for partition coefficients (between 30 and 40%) and much higher for the metabolic parameters (standard deviations representing about a factor of 2). Finally, the analysis points to differences between human males and females in the toxicokinetics of TCE. The significance of these differences in terms of risk remains to be investigated. *Key words*: Bayesian, human, Markov chain Monte Carlo, mouse, PBPK model, TCE, toxicokinetics, tricholoroethylene, uncertainty analysis, variability.
— *Environ Health Perspect* 108(suppl 2):275–282 (2000).
http://ehpnet1.niehs.nih.gov/docs/2000/suppl-2/275-282bois/abstract.html

New physiologically based pharmacokinetic (PBPK) models of trichloroethylene (TCE) in mice and humans have recently been published (*1,2*), together with new toxicokinetic data sets. These models, prior to their use in research or regulation, should be carefully calibrated by confrontation with the data. It is most efficient in terms of information use and from a statistical point of view to combine two forms of information: prior knowledge and data. Neither source of information is complete. If prior knowledge were sufficient, experiments would not need to be performed, but data alone are insufficient to pin down all parameters of a complex PBPK model to reasonable values. Prior knowledge can be summarized as prior parameter distributions. These can be obtained from the scientific literature, from specific *in vitro* experiments, and, as is the case here, from fitting previous data. The data [in this case the new data provided to us by Fisher (*2*)] are introduced through the use of a likelihood function. A Bayesian statistical framework gives rigorous rules to combine the prior distributions with the data likelihood, to yield posterior parameter distributions. In the case of complex models, like those described here, the posterior parameter distributions can be obtained by numerical Markov chain Monte Carlo (MCMC) techniques (*3,4*). This statistical methodology is gaining interest and is establishing itself for the calibration and validation of PBPK models (*3,5–8*). Beyond improving the fit, this method also provides distributions of prediction estimates directly usable as inputs for uncertainty analysis of cancer dose–response relationships.

An interesting aspect of the human data analyzed here is that they have been collected individually on a number of male and female volunteers. A hierarchical population model (*4,5*) deconvolves the various levels of variability present in the data. Such a model, of which the PBPK model is just a component, is easily calibrated with Bayesian numerical methods. This offers a unique opportunity to examine separately the important issues of variability and uncertainty in human toxicokinetics of trichloroethylene.

Methods

Data

Mice. Groups of male $B6C3F_1$ mice (body weight [bw] 25–30 g) were exposed to TCE or its metabolites by intravenous (i.v.) injection (*9*) or by oral gavage (*1*).

All i.v. doses were equal to 100 mg of the compound administered per kilogram of body weight. After chloral hydrate (CH) administration, venous concentrations of CH, free trichloroethanol (TCOH), trichloroacetic acid (TCA), and dichloroacetic acid (DCA), as well as the amount of urine excreted metabolites, were recorded at various times after dosing. Similarly, after TCA i.v. dosing, venous concentrations of TCA and DCA, as well as the amount of DCA excreted in urine, were recorded. Following TCOH dosing, venous concentrations of CH, TCOH, TCA, and the urine-excreted amounts of TCOH and glucuronidated trichloroethanol (TCOG), were measured. Finally, after DCA dosing, venous concentrations and urine-excreted amounts of DCA were measured.

When trichloroethylene was administered by gavage in corn oil, doses of 300, 600, 1,200, and 2,000 mg/kg bw were used. TCE concentrations were measured in venous blood, liver, fat, lung, and kidney. CH and TCOH concentrations were recorded in venous blood, liver, lung, and kidney. TCA, DCA, and TCOG concentrations were obtained in venous blood, liver, and kidney. The quantities of TCA and TCOG excreted in urine were also recorded.

Additional data, published by Templin et al. (*10*), were partly used. Groups of four male $B6C3F_1$ mice (bw 27 g) were gavaged with 3.8 mmol/kg (500 mg/kg) TCE. The venous blood concentrations of TCE, TCA, and DCA were measured at various times.

Humans. A group of 21 volunteers (10 females and 11 males) were exposed by inhalation to 50, 60, or 100 ppm TCE for 4 hr (*2*). Body weight and adiposity were recorded for each subject. The venous blood concentrations of TCE, TCOH, and TCA were measured during and after exposure. For some subjects, the exhaled air concentration of TCE was also recorded at various times after exposure. Finally, the cumulated quantities of TCA and TCOH glucuronide excreted in urine were recorded. The data for each individual were available to us.

Toxicokinetic and Statistical Models

Mice. The description of the physiological model used for mice can be found in Abbas and Fisher (*1*). The model equations were transcribed to a format suitable for MCSim (*11*). Three modifications were made to the

This article is part of the monograph on Trichloroethylene Toxicity.
Address correspondence to F.Y. Bois, INERIS, Parc ALATA, BP 2, 60550 Verneuil-en-Halatte, France. Telephone: 33 3 44 55 65 96. Fax: 33 3 44 55 68 99. E-mail: frederic.bois@ineris.fr
This work has been supported by the U.S. Environmental Protection Agency.
Received 20 October 1999; accepted 10 February 2000.

model, after consultation with the original authors. *a)* Metabolism of TCE in the lung was turned off. *b)* For several parameters, the scaling power of body weight was changed from 0.74 to 0.75 for standardization with the model of Clewell et al. (*12*). *c)* The volumes of the "body" and poorly perfused compartments, and the blood flow to the body compartment, were computed by difference at each iteration so that the sum of the organ volumes equaled 82% of the body weight and the sum of organ flows equaled cardiac output. Given this reparameterization, the model had a total of 57 independent parameters.

The statistical model used for mice lumps uncertainty and variability, since the animal data were aggregated (only the mean and standard deviation [SD] for several animals were provided). For a schematic of the model, see for example, Bois (*13*). Various concentrations or quantities (**y**) were measured. The expected values of these measurements are a function (f) of exposure level (**E**), time (**t**), a set of physiological parameters of unknown values ($\boldsymbol{\theta}$), and a set of measured, covariate parameters ($\boldsymbol{\varphi}$), such as body weight. **E**, **t**, $\boldsymbol{\theta}$, and $\boldsymbol{\varphi}$ are experiment specific. All animals in an experiment were supposed to have behaved similarly, from a toxicokinetic point of view. The function f is the pharmacokinetic model described above. The concentrations or quantities actually observed are also affected by measurement error and interindividual variability. Only aggregated data were available, and the corresponding errors were assumed to be independent and log-normally distributed, with mean zero and variance $\boldsymbol{\sigma}^2$ (on the log scale). The variance vector $\boldsymbol{\sigma}^2$ had 27 components for mice, most of them known by repeating the experiments (Table 1). The unknown experimental variances were sampled with a standard noninformative prior distribution $P(\sigma_i^2) \sim \sigma_i^{-2}$ (*14*). It was assumed that each θ parameter was a priori distributed log-normally (except for PCCH, the fraction TCE converted to CH, distributed normally), with averages **M** and variances **S** (in log scale).

Table 1. Values (*2*) or distributions used for the assay SDs, for each of the measured end points of the mouse data set.

Parameter	SD
CV	1.25
CL	1.25
CLU	LogUniform distribution
CF	1.39
CK	1.33
CCHL	1.11
CCHLU	1.27
CCHV	1.25
CCHK	1.13
CTCL	1.16
CTCV	1.19
CTCK	1.21
CDCL	1.24
CDCV	1.24
CDCK	1.40
COHL	1.17
COHLU	1.16
COHV	1.21
COHK	1.19
COGL	1.41
COGV	1.27
COGK	1.34
ACHKR	LogUniform distribution
ATCKR	LogUniform distribution
ADCKR	LogUniform distribution
AOHKR	LogUniform distribution
AOGKR	LogUniform distribution

A major advantage of physiological modeling is to provide a priori information on several of the mean parameter values for mice. Values for the hyperparameters **M** were set on the basis of the parameter values used by Fisher (*2*), or when applicable, on the basis of the posterior parameter distributions obtained from the analysis of the Clewell et al. model (*12,13*). To set **S**, a distinction was made between the physiological parameters or partition coefficients (which are quite well known) and the other metabolic or pharmacokinetic parameters (which are model specific and little known a priori). For the first group of parameters uncertainties of the order of 20–50% were assigned (*3,5,6*). For the second group of parameters, a "vague" distribution was assumed and **S** was set to correspond to a factor between 2 ("quite uncertain") or 5 ("very uncertain"). All priors on $\boldsymbol{\mu}$ were truncated to $\pm 2 \times S$ or $\pm 3 \times S$ to avoid reaching unrealistic values. Table 2 gives for mice the values used for exp(**M**) (i.e., the geometric means) and exp(**S**) (i.e., the geometric standard deviation [GSD]), both of which lie on the natural scale.

Humans. The PBPK model for humans is a simplification of the mouse model. Basically, in addition to TCE, only the distribution of TCOH and TCA in blood, liver, kidney, and rest of the body is described by physiological compartments. CH is not modeled explicitly: a fraction of TCE metabolized forms TCOH while the rest forms TCA. Formation of DCA is not modeled. A fraction of TCOH is glucuronidated to give TCOG, which is supposed to be immediately excreted in urine. Four modifications to the original model were introduced. For several parameters, the scaling power of body weight was changed from 0.74 to 0.75 for standardization with the model of Clewell et al. (*12*). The volumes of the body and poorly perfused compartments and the blood flow to the body compartment were computed by differences at each iteration so that the sum of the organ volumes equaled 82% of the body weight and the sum of organ flows equaled cardiac output; the quantity of TCOG excreted is in TCOH equivalents, rather than in TCOG equivalents. The Michaelis-Menten coefficient K_m (labeled KMTCOH) for the formation of TCOG was reexpressed as the ratio V_{max}/K_m for easier Monte Carlo sampling (preliminary runs showed that V_{max} and K_m for that reaction are highly correlated). The model had 33 independent parameters.

The statistical model describing uncertainties and variabilities in human data was constructed using a hierarchical population approach (*3,4*). The model has two major components: the individual level and the population level. At the individual level, various concentrations or quantities (**y**) were measured. The expected values of these measurements are a function (f) of exposure level (**E**), time (**t**), a set of physiological parameters of unknown values ($\boldsymbol{\theta}$), and a set of measured, covariate parameters ($\boldsymbol{\varphi}$), such as body weight. **E**, **t**, $\boldsymbol{\theta}$, and $\boldsymbol{\varphi}$ are experiment specific. The function f is the human pharmacokinetic model. The concentrations or quantities actually observed are also affected by measurement errors. These errors were assumed to be independent and log-normally distributed, with mean zero and variance $\boldsymbol{\sigma}^2$ (on the log scale). The variance vector $\boldsymbol{\sigma}^2$ had six components for humans.

At the population level, each component of the $\boldsymbol{\theta}$ parameter set was assumed to be distributed log-normally, with averages $\boldsymbol{\mu}$ and variances $\boldsymbol{\Sigma}^2$ (in log scale). A priori knowledge of $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}^2$ is available under the form of standard values for some parameters. Since uncertainty in these average values and variances has to be acknowledged, a priori log-normal distributions were assumed for the population means $\boldsymbol{\mu}$ (with hyperparameters **M** and **S**) and a standard inverse gamma distribution (*15*) for the population variances $\boldsymbol{\Sigma}^2$.

Values for the hyperparameters **M** were set on the basis of the parameter values used by Fisher (*2*), or when applicable, on the basis of the posterior parameter distributions obtained from the analysis of the Clewell et al. model (*12,13*). To set **S**, a distinction was made between the physiological parameters or partition coefficients (which are quite well known) and the other metabolic or pharmacokinetic parameters (which are model specific and little known a priori). For first group of parameters uncertainties of the order of 20–50% were assigned (*3,5,6*). For the second group of parameters, a vague distribution was assumed and S was set to correspond to a factor of 5. All priors on $\boldsymbol{\mu}$ were truncated to $\pm 2 \times S$.

The β coefficient of the inverse-gamma prior on $\boldsymbol{\Sigma}^2$ was set to the posterior values found in the previous analysis (*13*) when applicable and to 0.22 otherwise (this corresponds to a CV of 50%). The α coefficients of the inverse-gamma priors were set to 1 or 6 [the latter when a posterior distribution

had already been obtained from analysis of the Clewell et al. data (*13*), which had several groups of subjects]. Table 3 gives for humans the values used for exp(**M**)—i.e., the geometric means—exp(**S**) —the GSD—and the α and β coefficients. For $\boldsymbol{\sigma}^2$, a standard noninformative prior distribution $P(\sigma_1^2,\ldots,\sigma_n^2) \sim \sigma_1^{-2} \times \ldots \times \sigma_n^{-2}$ was used.

Statistical Computation of Posterior Parameter Distributions

From Bayes' theorem, the joint posterior distribution of the parameters to estimate, $P(\boldsymbol{\theta}, \boldsymbol{\sigma}^2, \boldsymbol{\mu}, \boldsymbol{\Sigma}^2|\mathbf{y}, \boldsymbol{\varphi}, \mathbf{E}, \mathbf{t}, \mathbf{M}, \mathbf{S}, \boldsymbol{\Sigma}_0)$, in the case of humans, is proportional to the likelihood of the data multiplied by the parameters' priors:

$$P(\boldsymbol{\theta}, \boldsymbol{\sigma}^2, \boldsymbol{\mu}, \boldsymbol{\Sigma}^2|\mathbf{y}, \boldsymbol{\varphi}, \mathbf{E}, \mathbf{t}, \mathbf{M}, \mathbf{S}, \boldsymbol{\Sigma}_0^2) \sim P(\mathbf{y}|\boldsymbol{\theta},\boldsymbol{\sigma}^2,\boldsymbol{\varphi},\mathbf{E},\mathbf{t}) \bullet P(\boldsymbol{\theta}|\boldsymbol{\mu},\boldsymbol{\Sigma}^2) \bullet P(\boldsymbol{\sigma}^2) \bullet P(\boldsymbol{\mu}|\mathbf{M},\mathbf{S}) \bullet P(\boldsymbol{\Sigma}^2|\boldsymbol{\Sigma}_0^2) \quad [1]$$

A similar, but simpler, expression can be obtained in the case of the mouse model. For mice and humans, the likelihood term is given by the normal measurement model:

$$\log(\mathbf{y}) \sim N(\log f(\boldsymbol{\theta},\boldsymbol{\varphi},\mathbf{E},\mathbf{t},)\boldsymbol{\sigma}^2) \quad [2]$$

As mentioned above, the prior distribution for $\boldsymbol{\sigma}^2$ is: $P(\sigma_1^2,\ldots,\sigma_n^2) \sim \sigma_1^{-2} \times \ldots \times \sigma_n^{-2}$, with eventually some fixed components.

For humans, the prior distribution of each component of $\boldsymbol{\theta}$ is an independent normal distribution in log space:

$$\log(\theta) \sim N(\mu,\Sigma^2) \quad [3]$$

with truncation constraints. Each component of $\boldsymbol{\mu}$, or $\boldsymbol{\Sigma}^2$ is assigned an independent hyper-prior distribution, $\mu \sim N(M,S^2)$ and $\Sigma^2 \sim$ inverse-gamma(α,β), as described in the previous section. For mice, the $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}^2$ levels are not defined and prior distribution of $\boldsymbol{\theta}$ depends directly on **M** and **S**.

Current practice in Bayesian statistics is to summarize a complicated high-dimensional posterior distribution by random draws of the vector of parameters. This is currently the most effective way to perform high-dimensional numerical integration. Because there are many parameters to estimate, Metropolis-Hasting sampling was used to perform a random walk through the posterior distribution. This iterative sampling procedure is particularly convenient in the case of hierarchical models. It belongs to a class of MCMC techniques that has recently received much interest (*3,16–21*). Three independent MCMC runs were performed for each species. Convergence was monitored using the method of Gelman and Rubin (*22*).

Table 2. Prior and posterior distributions for the scaling coefficients of the mouse PBPK model (*1*). All distributions are log-normal with truncations at ±2 SDs, except where indicated.

	Prior distribution		Posterior distribution			
Parameter	Geometric mean	Geometric SD	Geometric mean	Geometric SD	2.5% percentile	97.5% percentile
BW[a]	0.03	1.11	0.0354	1.02	0.0334	0.0365
QCC[b]	20.1	1.08	14.4	1.05	13.1	16.1
QPC[b]	21.8	1.08	19.3	1.07	16.8	22.0
QLC[b]	0.2	1.08	0.259	1.05	0.235	0.284
QKC[c]	0.09	1.22	0.0944	1.21	0.0646	0.129
QFC[b]	0.054	1.14	0.0455	1.11	0.0373	0.0561
VLUC[c]	0.007	1.22	0.00654	1.18	0.00480	0.00900
VLC[b]	0.05	1.07	0.0478	1.06	0.0424	0.0529
VKC[c]	0.018	1.22	0.0221	1.12	0.0180	0.0265
VFC[b]	0.063	1.12	0.0773	1.10	0.0646	0.0944
VRC[b]	0.05	1.07	0.0503	1.06	0.0442	0.0567
K1[b]	1	2.72	4.18	1.14	3.22	5.47
K2[b]	17.3	1.24	3.46	1.10	2.86	4.14
K3[b]	1.14	1.36	0.0916	1.01	0.0898	0.0944
PB[b]	16.4	1.13	22.9	1.08	19.7	26.6
PLU[c]	2.61	1.6	1.81	1.10	1.50	2.19
PL[b]	1.74	1.2	3.10	1.04	2.83	3.35
PK[c,d]	2.07	1.6	4.81	1.06	4.31	5.42
PF[b]	30.6	1.15	15.6	1.05	14.3	17.5
PS[b]	0.756	1.21	1.41	1.13	1.08	1.75
PR[b]	1.75	1.21	1.97	1.22	1.31	2.89
PCHLU[c]	1.65	1.6	1.96	1.07	1.71	2.23
PCHL[c]	1.42	1.6	1.61	1.05	1.45	1.78
PCHK[c,d]	0.98	1.6	0.361	1.06	0.323	0.399
PCHB[c,d]	1.35	1.6	1.51	1.08	1.29	1.74
POHLU[c]	0.78	1.6	1.71	1.04	1.58	1.85
POHL[c,d]	1.3	1.6	3.97	1.04	3.71	4.31
POHK[c,d]	1.02	1.6	3.90	1.04	3.63	4.14
POHB[c]	1.11	1.6	0.595	1.08	0.516	0.683
POGLU[c]	1.06	2	1.01	1.73	0.361	3.00
POGL[c]	0.56	2	0.897	1.08	0.771	1.04
POGK[c]	1.44	2	2.65	1.13	2.12	3.42
POGB[c,d]	1.11	2	0.449	1.25	0.275	0.668
PTCLU[c]	0.54	1.6	0.518	1.46	0.249	1.07
PTCL[c]	1.18	1.6	0.780	1.03	0.733	0.830
PTCK[c]	0.74	1.6	0.758	1.04	0.706	0.812
PTCB[c]	0.88	1.6	0.857	1.05	0.773	0.945
PDCLU[c]	1.23	1.6	1.14	1.43	0.572	2.33
PDCL[c,d]	1.08	1.6	0.636	1.05	0.576	0.700
PDCK[c]	0.74	1.6	0.843	1.08	0.730	0.973
PDCB[c]	0.37	1.6	0.548	1.06	0.487	0.619
VMAXC[b]	38.1	1.15	45.2	1.03	42.5	47.5
KM[b]	0.47	1.57	10.1	1.05	9.21	11.1
PCCH[a,e]	0.99	0.01	0.995	0.00380	0.986	1.00
KRCHC[c]	0.06	5	0.0493	1.17	0.0365	0.0679
PCTCOC[c]	309	5	403	1.09	347	478
PCTCAC[c]	115	5	119	1.08	103	141
VMTCOC[c]	16.5	5	74.4	1.11	60.9	90.9
KMTCOH[c]	15.7	5	49.9	1.14	39.3	64.1
KOCHC[c]	1.32	5	5.58	1.13	4.39	7.03
KROHC[c]	1.14	5	0.133	1.18	0.0983	0.183
KFTCC[c]	0.35	5	0.546	1.08	0.472	0.639
KRTCC[c]	1.55	5	0.596	1.12	0.489	0.749
KFDCC[c]	20.5	5	41.7	1.09	35.5	48.4
KRDCC[a]	1	2.72	2.62	1.14	2.13	3.49
KROGC[c]	32.8	5	19.9	1.16	15.2	26.8
KGBLC[c]	4.61	5	16.4	1.08	14.2	19.1

[a]Prior set by ourselves. [b]Prior based on the posterior distribution obtained previously (*3*). [c]Prior based on Fisher's values (*2*). [d]Truncation at ±3 SDs. [e]For this parameter a normal distribution was used, with truncation at 0.95 and 1.

Table 3. Prior and posterior distributions for the population averages (μ), and population SDs (Σ), of the scaling coefficients for the human PBPK model.

Parameter	Prior on μ Geometric mean (GSD)	Prior on Σ Geometric mean (GSD)	Posterior on μ Geometric mean (GSD)	Posterior on Σ Geometric mean (GSD)
QCC[a]	15.2 (1.16)	1.41 (1.12)	16.3 (1.08)	1.30 (1.15)
QPC[a]	16.1 (1.16)	1.39 (1.10)	16.9 (1.08)	1.36 (1.17)
QLC[a]	0.232 (1.17)	1.42 (1.11)	0.244 (1.07)	1.20 (1.12)
QKC[b]	0.198 (1.22)	–	0.194 (1.08)	1.26 (1.14)
QFC[a]	0.0523 (1.16)	1.47 (1.12)	0.0639 (1.10)	1.19 (1.12)
VLUC[b]	0.014 (1.22)	–	0.0141 (1.11)	1.25 (1.14)
VLC[a]	0.026 (1.13)	1.37 (1.10)	0.0257 (1.10)	1.14 (1.08)
VKC[b]	0.00401 (1.22)	–	0.00401 (1.13)	1.26 (1.14)
VFC[a]	0.198 (1.13)	1.37 (1.10)	0.196 (1.05)	1.21 (1.11)
VRC[a]	0.0483 (1.13)	1.37 (1.10)	0.0437 (1.10)	1.13 (1.08)
PB[a]	13.7 (1.19)	1.52 (1.14)	18.0 (1.09)	1.21 (1.12)
PLU[b]	0.391 (1.6)	–	0.386 (1.18)	1.37 (1.24)
PL[a]	6.69 (1.32)	1.52 (1.16)	5.81 (1.15)	1.22 (1.15)
PK[b]	1.08 (1.6)	–	1.07 (1.18)	1.37 (1.24)
PF[a]	53 (1.22)	1.50 (1.15)	50.9 (1.15)	1.32 (1.20)
PS[a]	2.7 (1.23)	1.49 (1.15)	1.50 (1.14)	1.22 (1.17)
PR[a]	5.05 (1.28)	1.51 (1.16)	3.67 (1.11)	1.17 (1.10)
POHLU[b]	0.67 (1.6)	–	0.685 (1.18)	1.37 (1.24)
POHL[b]	0.589 (1.6)	–	0.616 (1.18)	1.37 (1.25)
POHK[b]	2.15 (1.6)	–	2.16 (1.18)	1.37 (1.24)
POHB[b]	0.91 (1.6)	–	1.26 (1.10)	1.47 (1.24)
PTCLU[b]	0.47 (1.6)	–	0.472 (1.18)	1.37 (1.24)
PTCL[b]	0.66 (1.6)	–	0.708 (1.21)	1.37 (1.24)
PTCK[b]	0.66 (1.6)	–	0.668 (1.18)	1.37 (1.24)
PTCB[b]	0.519 (1.6)	–	0.601 (1.10)	1.41 (1.23)
VMAXC[a]	43.8 (1.97)	1.70 (1.28)	4.22 (1.19)	1.48 (1.33)
KM[a]	0.542 (2.35)	1.76 (1.35)	0.801 (1.40)	1.96 (1.65)
POH[c]	0.9 (0.1)	–	0.730 (0.041)	0.187 (0.10)
VMTCOC[b]	1 (5)	–	5.26 (1.56)	1.76 (1.72)
KMTCOH[b]	1.57 (5)	–	2.72 (1.12)	1.70 (1.35)
KOCHC[b]	1 (5)	–	8.58 (1.14)	1.51 (1.29)
KTMETC[b]	0.1 (5)	–	0.301 (1.22)	1.98 (1.57)
KRTCC[b]	1 (5)	–	1.15 (1.19)	2.18 (1.56)

[a]Prior distribution based on the posterior distribution obtained previously (*3*). [b]Prior distribution on μ based on Fisher's values (*2*). A vague inverse-gamma (1, 0.22) was used. [c]For this parameter a normal distribution was used, with truncation at 0.6 and 1, so the parameters are mean and SD in natural space.

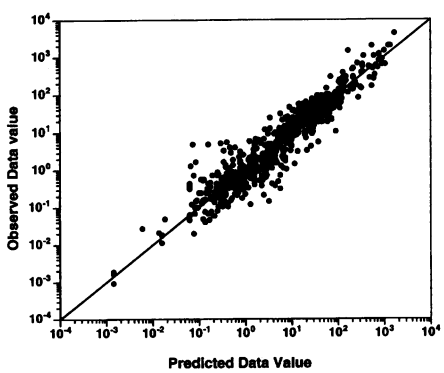

Figure 1. Observed versus predicted mice data values (all concentrations or quantities) for the Monte Carlo iteration of highest posterior probability.

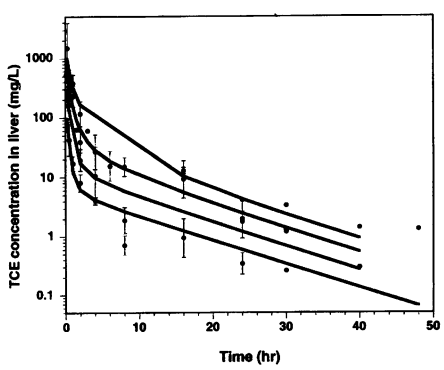

Figure 2. Predicted (solid line) and observed (points) time course of TCE concentration in liver of mice dosed with various quantities of TCE in corn oil. Error bars correspond to ±2 SD.

Results

Mice

For mice, 20,000 iterations were necessary to reach convergence of the sampler. One of every 5 of the last 5,000 simulations of three independent Markov chains were recorded, yielding 3,000 sets of parameter values from which the inferences and predictions presented in the following were made.

Quality of data adjustment. Figure 1 presents the data values observed compared with their predicted counterparts (data values are concentrations or excreted quantities). Predictions were made with the parameter set of highest posterior density. For a perfect fit, all points would fall on the diagonal (equality of predicted and observed values). Such an adjustment is not expected given the analytical measurement errors in the data. The graph is presented on log–log scale, since the errors are assumed to be lognormally distributed and the data span a wide range. Most of the residuals are contained within a factor of 3 along the diagonal, but quite a few reach a factor of 10 or even a factor of 100. Figures 2–7 show all the gavage data and corresponding predictions in the liver as a function of time. Basically the model reproduces correctly the profiles for TCE and TCA. TCOH and TCOG concentrations are less well predicted. CH and DCA concentrations and TCOG excreted are rather poorly fitted. The fact that each data point represents a different group of animals may explain the noise present in the data. However, some differences between model and data appear to be systematic (Figure 3, for example).

Posterior parameter distributions. The joint distribution of all parameters is obtained in output of the MCMC simulations. This allows consideration of marginal distributions (distributions of the parameters considered individually) and also of correlations of any order. Table 1 summarizes the posterior distributions of the mouse parameter values obtained in the last 1,000 iterations of the three runs performed (results of the three runs are pooled, and the distributions are established with 3,000 values). The geometric means can be interpreted as representing the values for an "average" mouse. Note that the columns of geometric standard deviations represent the sum of group variability among mice and experimental error. Overall, the parameters retain physiologically plausible values.

Yet the posterior means for several parameters are quite far from the prior means. This is observed for metabolic parameters or for some partition coefficients. In particular, the value of KM for TCE is 20 times that found with the previous model and experiments. SDs about these geometric means are quite low and hover around 1.1 (corresponding to a 10% CV, which includes uncertainty and variability). The parameter values are therefore quite well identified by the data.

Very high correlations are observed between parameter estimates for mice (up to 0.95 for KFTCC and KFDCC). Any parameterization neglecting to estimate these covariances will produce incorrect predictions, since these parameters cannot be sampled independently without producing highly improbable combinations and hence highly improbable predictions.

Humans

More than 100,000 iterations were necessary to reach convergence of the sampler in the case of human data. One of every 10 of the last 50,000 simulations of three independent Markov chains were recorded, yielding 15,000 sets of parameter values from which

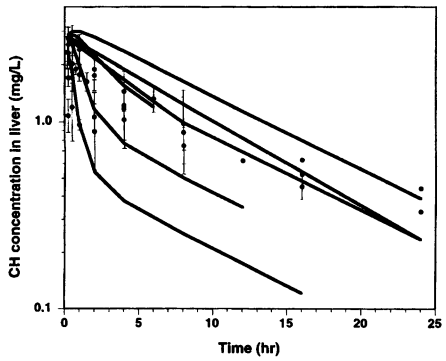

Figure 3. Predicted (solid line) and observed (points) time course of CH concentration in liver of mice dosed with various quantities of TCE in corn oil. Error bars correspond to ±2 SD.

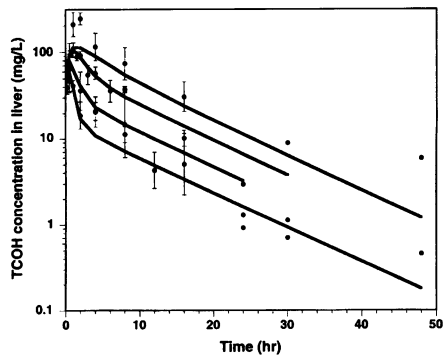

Figure 4. Predicted (solid line) and observed (points) time course of TCOH concentration in liver of mice dosed with various quantities of TCE in corn oil. Error bars correspond to ±2 SD.

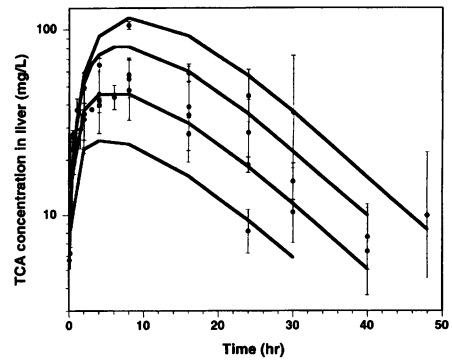

Figure 5. Predicted (solid line) and observed (points) time course of TCA concentration in liver of mice dosed with various quantities of TCE in corn oil. Error bars correspond to ±2 SD.

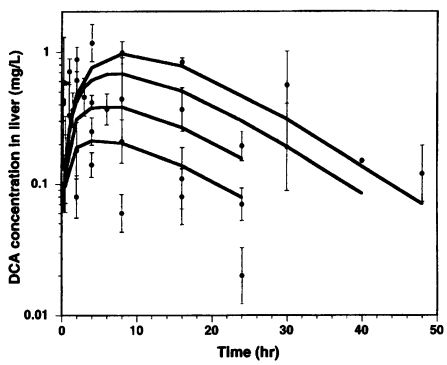

Figure 6. Predicted (solid line) and observed (points) time course of DCA concentration in liver of mice dosed with various quantities of TCE in corn oil. Error bars correspond to ±2 SD.

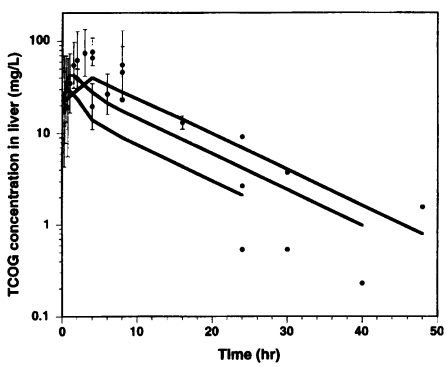

Figure 7. Predicted (solid line) and observed (points) time course of TCOG concentration in liver of mice dosed with various quantities of TCE in corn oil. Error bars correspond to ±2 SD.

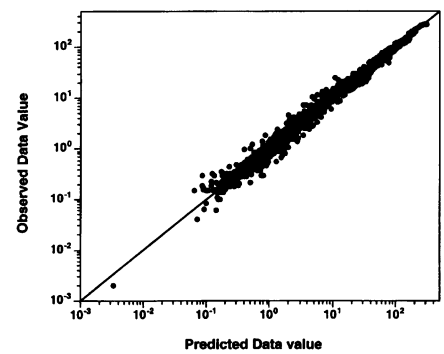

Figure 8. Observed versus predicted human data values (all concentrations or quantities) for the Monte Carlo iteration of highest posterior probability.

the inferences and predictions presented in the following were made.

Quality of data adjustment. Figure 8 gives the observed data values versus their predicted counterparts (data values are concentrations or excreted quantities). Predictions were made with the parameter set of highest posterior density. The graph is presented on log–log scale, since the errors are assumed to be lognormally distributed and the data span a wide range. Most of the residuals are contained within a factor of 2 along the diagonal. The fit here is substantially better than for mice. Figures 9–11 show all the data and corresponding predictions as a function of time. The model was adjusted to each individual's data, using a population toxicokinetic approach. The data are reasonably well fitted, overall, and particularly the TCA data. However, the concentration of TCE in alveolar air and venous blood is not adjusted satisfactorily, given the quality of adjustment expected from a PBPK model: The curves do not "bend" enough (Figure 9). Another area of significant misfit is the prediction of TCOH concentration in venous blood. The predicted terminal half-life is clearly too short (data not shown). Yet, a good adjustment of TCOH can be obtained for the volunteers observed up to 22 hr (Figure 10). Most of the variability observed on the figures is due to differences in inhalation exposure levels (concentrations of 50, 60, or 100 ppm TCE were used). However, a large part of the variability in TCOG elimination to urine seems due to factors other than exposure levels (data not shown). Somewhat troublesome is the disagreement between measured (albeit, approximately) fractions of body weight as fat (VFC) and the corresponding toxicokinetic estimates. The estimates take into account the measured values as well as what can be inferred about the size of the fat compartment from the TCE and other concentration data. The measured values of VFC reach much more extreme values than the estimates (Figure 12).

Posterior parameter distributions. Table 3 summarizes the posterior distributions of the population means and SDs for all human model parameters (the distributions were established with 15,000 values). The population means represent the values for an average person. They are affected by uncertainty and were each assigned a geometric mean and a GSD (note that an approximate CV can quickly be computed, since CV ≈ GSD – 1). The population SDs measure between-subject variability. They are also affected by uncertainty (i.e., the population SD cannot be exactly computed from a finite sample).

The posterior means for several parameters are quite far from their prior estimates. The location of the blood over air partition coefficient, PB, is shifted from 13.7 (±20%) to 18 (±10%). The location of other partition coefficients (slowly perfused muscle over blood and richly perfused liver over blood) is also changed. The largest shifts are observed for metabolic parameters. The estimate of the scaling coefficient of TCE V_{max} in liver (VMAXC) is 4.22 (±20%) instead of 43.8. A large uncertainty existed on that prior estimate, but 4.22 is still outside its prior 95% confidence interval (lower bound at 12.6). Therefore, there appears to be a conflict between the data studied here and the data studied previously. Note also that the posterior mean of the fraction TCE converted to TCOH (POH) is quite lower than a priori estimated. The posterior means of the scaling coefficient of V_{max} for TCOH metabolism (VMTCOC) and of the scaling coefficient of the rate constant from TCOH to TCA (KOCHC) are much increased but not

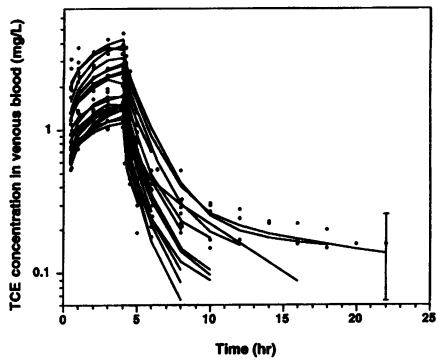

Figure 9. Predicted (solid lines) and observed (points) time course of TCE concentration in venous blood of human volunteers exposed by inhalation to various concentrations of TCE. The error bars presented on a data point correspond to ±2 estimated measurement SD (the size of the error is the same for all points).

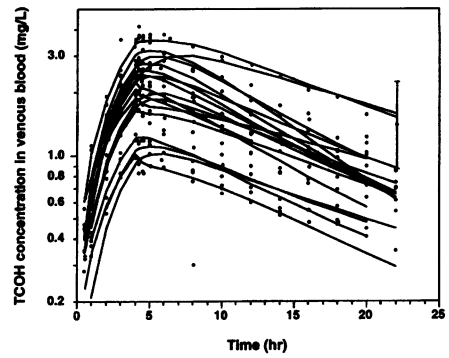

Figure 10. Predicted (solid lines) and observed (points) time course of TCOH concentration in venous blood of human volunteers exposed by inhalation to various concentrations of TCE. These volunteers were followed for up to 22 hr. The error bars presented on a data point correspond to ±2 estimated measurement SD (the size of the error is the same for all points).

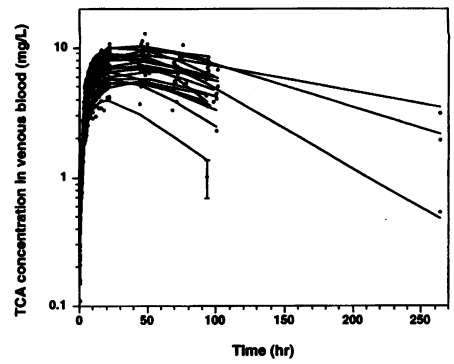

Figure 11. Predicted (solid lines) and observed (points) time course of TCA concentration in venous blood of human volunteers exposed by inhalation to various concentrations of TCE. The error bars presented on a data point correspond to ±2 estimated measurement SD (the size of the error is the same for all points).

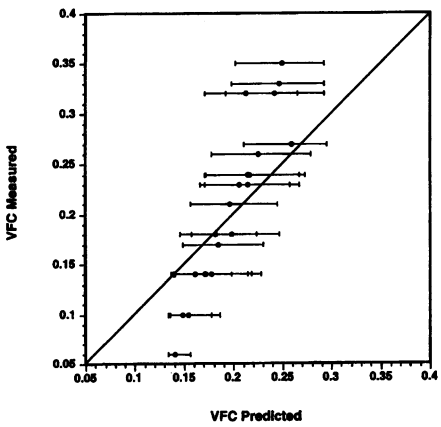

Figure 12. Observed versus predicted fraction of body weight as fat in human volunteers exposed by inhalation to various concentrations of TCE. The error bars span the 95% confidence intervals of the predictions.

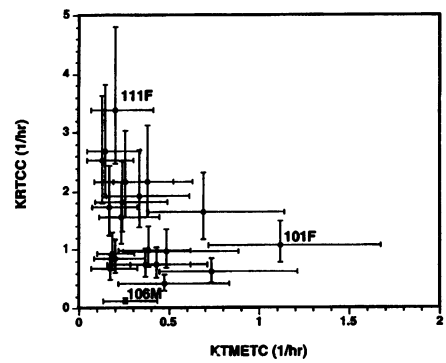

Figure 13. Individual estimates of the rate constant scaling coefficient for TCA metabolism (KTMETC) and of the rate constant of TCA loss to urine (KRTCC). The error bars span the 95% confidence intervals of the estimates. Large individual differences are observed, with some degree of correlation.

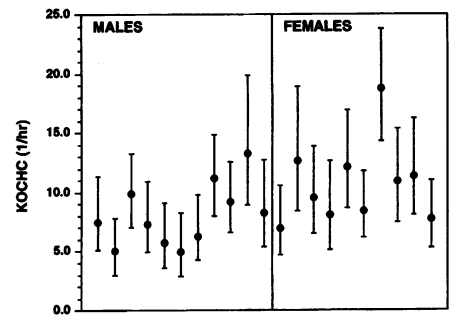

Figure 14. Individual estimates of the rate constant scaling coefficient for TCOH to TCA, grouped by sex. The error bars span the 95% confidence intervals of the estimates.

incompatible with the vague priors used. In terms of uncertainty, SDs about the posterior means are quite low, approximately 1.1 or 1.2 (corresponding to a 10–20% CV). The parameter values are overall quite well identified by the data.

Estimates of population variability are given by the posteriors of the population SDs, Σ. Volumes and flows appear to vary across subject by approximately 20% (CV). The variability is somewhat higher for partition coefficients (between 30 and 40% CVs). It is much higher for the metabolic parameters (SDs representing approximately a factor of 2 difference). For example, the lowest rate constant scaling coefficient for TCA metabolism (KTMETC) is $0.12 \mathbin{\div\!\times} 1.6\ \mathrm{hr}^{-1}$ (subject 110F) and the highest is $1.1 \mathbin{\div\!\times} 1.25\ \mathrm{hr}^{-1}$ (subject 101F, the subject with lowest TCA levels on Figure 11). For the scaling coefficient of the rate constant of TCA loss to urine (KRTCC), the values range from $0.13 \mathbin{\div\!\times} 1.2\ \mathrm{hr}^{-1}$ (subject 106M) to $3.4 \mathbin{\div\!\times} 1.2\ \mathrm{hr}^{-1}$ (subject 111F) (Figure 13). Subject 111F excreted the highest amount of TCA in urine (data not shown). Note that the estimates of variability given are themselves affected by uncertainty (up to 50% CV for the variability estimates of the metabolic parameters). For example, the 95% posterior confidence interval of S for KRTCC is [1.77, 2.85].

Examination of the subjects' parameter values points to peculiar pharmacokinetic behaviors for some of them:

- Subject 106M has the highest fraction fat (VFC), particularly for a male (some female subjects approach his VF, but no males of the group studied approach it). He is also an "outlier" with a TCE fat over blood partition coefficient 3 times that of the others. His maximal rate scaling coefficient for TCE metabolism (VMAXC) and his rate constant of TCA loss to urine (KRTCC) are the lowest. Note that this subject shows the lowest amount of TCA excreted (data not shown), and his TCOH blood concentrations are not well modeled.
- Subject 201M is also outstanding for his highest alveolar ventilation rate scaling coefficient (QPC), highest fractional blood flow to the kidney (QKC), and high TCOH body over blood partition coefficient (POHB), particularly for a male (females tend to have higher POHBs, see below). This subject has the second lowest TCA excretion in urine and one of the lowest TCOG elimination. It seems that this subject can be characterized as "fast" in intake and elimination, at least from a physiological (if not enzymatic) point of view.
- Subject 103M has a Michaelis-Menten constant (KM) about 3 times as high as everyone else. He has one of the highest excretions of TCA in urine.

Splitting the subjects by sex reveals some differences between males and females. The visual examination of the marginal distributions of a given parameter for each subject reveals little pattern. For example, in Figure 14, it appears that female subjects exhibit higher values of KOCHC, the rate constant scaling coefficient for TCOH to TCA. However, a large part of the variability

observed is due to uncertainty about the population mean. When the effect of the overall mean is removed, the difference is highly significant: $P(KOCHC_{female} > KOCHC_{male}) = 0.9998$ (i.e., the female average was superior to the male average 99.98% of the times in the sample of 15,000 parameter vectors obtained by MCMC sampling). Other significant sex differences were found:

- The scaling coefficient of the alveolar ventilation rate is higher for males $P(QPC_{male} > QPC_{female}) = 0.9997$. This implies that the data are sufficiently informative to show such a difference and also that the surface area correction is not sufficient to explain sex differences in alveolar ventilation. The level of physical training might be another covariate to consider for alveolar ventilation in human studies.
- The posterior estimates of the VFC are higher on average for females than for males: $P(VFC_{female} > VFC_{male}) > 0.9999$. That was expected and is influenced by the covariate measurements of adiposity and by the toxicokinetic data.
- Female subjects also seem to have higher TCOH body over blood partition coefficients: $P(POHB_{female} > POHB_{male}) > 0.9999$; they also have a higher V_{max} over K_m ratio (i.e., rate constant at low concentration) for TCOH glucuronidation: $P(KMTCOH_{female} > KMTCOH_{male}) > 0.9999$, and a higher rate constant scaling coefficient for TCA urinary excretion: $P(KRTCC_{female} > KRTCC_{male}) > 0.9943$.
- Males were found to have higher TCA body over blood partition coefficients: $P(PTCB_{male} > PTCB_{female}) = 0.992$.

All other male versus female parameter averages do differ slightly, but the P values are less than 0.95, most of them being around 0.5.

Discussion

The mice data show a rather large interindividual or interlot variability. There is little to do about that, unless the experimental design were to allow the observation of individual animals. The data still permit an extensive calibration of the model. The human data have the advantage of being available for each individual. There appear to be, a posteriori, a few outlying points (for example, Figure 10), which could be checked and eventually removed. These few apparent outliers have little weight and should not affect sensibly the results presented here. There could be a problem with the measurement of fat content; this will be discussed below, in light of the model fit.

The mouse model is quite complicated, and yet some aspects of the data are not so well described by it (CH and DCA concentrations, and TCOG excreted). This is somewhat disappointing, given the complexity of the model, and one wonders what will have to be done to fit such an extensive data set. It is possible that some metabolic reactions do not obey simple Michaelis-Menten reactions or that prior opinions about the model parameters were for some of them overly confident. Note also that even in large models, model uncertainty can be large. The model, for example, does not include hepatic recycling or all possibilities of extrahepatic metabolism. Yet, overall, the model fits well a large part of the data, in particular TCE and TCA distribution. It would be interesting to see the performance of other models [e.g., those presented in (*12,23*)] with this data set, which would allow a formal comparison of the competing models on the basis of a common measure of goodness of fit, such as likelihood ratios.

The posterior parameter distributions obtained for mice are quite narrow (with CVs of about 10 or 20%), indicating that the data are strongly informative for most parts of the model. Indeed, as indicated above, some parameters or processes might need to be added for a better fit and the model is somewhat minimal with respect to the data. However, it should be kept in mind that the fit is not excellent, and that may overconstrain the posterior distributions. Part of the high covariance between parameters may also be due to overconstraining. It is possible to model statistically the lack of fit by including an autocorrelation between data points (*24*). This has not been attempted here and could be a useful improvement. Note that the posterior uncertainty for the metabolic parameters would have been further underestimated if all physiological parameters had been set to predefined values.

The human model, even though complex, also has difficulties in fitting all the data. This is true in particular for TCOH concentrations over a long period of time, and some improvement of the model in that respect may be needed. Similarly, the adiposity of the subjects (a measured covariate) does not fit well with the estimated fraction of body weight as fat. It is possible that the pharmacokinetic compartment "fat" is not well estimated by external adiposity measurements, in particular for extreme values. It also appears that the model may not be able to describe correctly outlying subjects like subject 106M. This could be due to the PBPK model, which lacks some component important for such a subject. The misfit for that subject could also be due to a lack of flexibility of the statistical model adopted here (log-normal distributions of the parameters in the population). A possibility for checking would be to fit the data of only that subject to determine if a good fit could be obtained.

The human posterior parameter distributions agree in general quite well with the corresponding priors, with reduced uncertainty (since information from the data has been gained). SDs about the posterior means are quite low and correspond to a 10–20% CV: The parameter values are overall quite well identified by the data. There is, however, a conflict for the values of VMAXC between the values previously found (*3*) and the ones obtained here. This is an important parameter and a good characterization is important. The difference could be due to conflicts between the data analyzed here and the previous data. For example, extensive TCOH data are available here. It is also possible that the conflict is due to differences in human model structures. A solution to this dilemma would be to take the model of Clewell et al. (*12*) and fit it to the data of Fisher (*2*) to obtain an estimate of VMAXC with the same data set. In any case, it is not obvious that the differences in parameter values would result in notable differences when predicting toxicologically relevant end points, such as internal TCA concentrations. This remains to be checked.

Volumes and flows appear to vary across subject by 20% (CV), approximately. The variability is somewhat higher for partition coefficients (between 30 and 40% CVs) and much higher for the metabolic parameters (SDs representing about a factor of 2 difference). This is similar to what was found for a small group of human volunteers exposed to tetrachloroethylene (*5*). Differences appear to exist between sexes in the toxicokinetics of TCE. There are differences between males and females in alveolar ventilation and adiposity for the population sample studied. This should be true for compounds other than TCE and shows that the model scaling could be improved. More important for TCE kinetics are the findings that females have higher TCOH body over blood partition coefficients, lower TCA body over blood partition coefficients, higher V_{max} over K_m ratios for TCOH glucuronidation, higher rate constant scaling coefficients for TCOH to TCA, and higher rate constant scaling coefficients for TCA urinary excretion. Note that the statistically highly significant differences found for these parameters should be interpreted with some caution. They are conditional on the model structure being correct. At least, at this point, it can be said that there are most certainly differences in the kinetic behavior of TCE between the males and females of the sample studied. It would have been hard to reach that conclusion without the statistical adjustment of a model, given the multiple exposure levels, differences in body weight, nonlinear kinetics, etc. Still, these differences may not be significant in terms of TCE toxicity (i.e., biologically significant). It would be interesting to assess by simulations whether internal metabolite concentrations are much different for males and females for the same TCE exposure.

References and Notes

1. Abbas R, Fisher JW. A physiologically based pharmacokinetic model for trichloroethylene and its metabolites, chloral hydrate, trichloroacetate, dichloroacetate, trichloroethanol, and trichloroethanol glucuronide in B6C3F$_1$ mice. Toxicol Appl Pharmacol 147:15–30 (1997).
2. Fisher JW. Personal communication, 1997.
3. Gelman A, Bois F, Jiang J. Physiological pharmacokinetic analysis using population modeling and informative prior distributions. J Am Stat Assoc 91:1400–1412(1996).
4. Wakefield JC. The Bayesian analysis of population pharmacokinetic models. J Am Stat Assoc 91:62–75(1995).
5. Bois FY, Gelman A, Jiang J, Maszle D, Zeise L, Alexeef G. Population toxicokinetics of tetrachloroethylene. Arch Toxicol 70:347–355 (1996).
6. Bois F, Jackson E, Pekari K, Smith M. Population toxicokinetics of benzene. Environ Health Perspect 104(suppl 6):1405–1411 (1996).
7. Fanning E, Bois FY, Rothman N, Bechtold B, Li G, Hayes R, Smith M. Population toxicokinetics of benzene and its metabolites. In: Society of Toxicology Annual Meeting, Cincinnati, Ohio, 1997.
8. Johanson G, Jonsson F, Bois F. Development of new technique for risk assessment using physiologically based toxicokinetic models. Am J Ind Med (suppl)1:101–03 (1999).
9. Abbas R, Seckel CS, Kidney JK, Fisher JW. Pharmacokinetic analysis of chloral hydrate and its metabolism in B6C3F$_1$ mice. Drug Metab Dispos 24:1340–1346 (1996).
10. Templin MV, Parker JC, Bull RJ. Relative formation of dichloroacetate and trichloroacetate from trichloroethylene in male B6C3F$_1$ mice. Toxicol Appl Pharmacol 123:1–8 (1993).
11. Maszle D, Bois FY. Program MCSim - User Manual. Available: *ftp://rfs63.berkeley.edu/pub/mcsim*, 1993.
12. Clewell HJ III, Gentry PR, Covington TR, Gearhart JM. Development of a physiologically based pharmacokinetic model of trichloroethylene and its metabolites for use in risk assessment. Environ Health Perspect 108(suppl 2):283–305 (2000).
13. Bois FY. Statistical analysis of Clewell et al. PBPK model of trichloroethylene kinetics. Environ Health Perspect 108(suppl 2):307–316 (2000).
14. Bernardo JM, Smith AFM. Bayesian Theory. New York:Wiley, 1994.
15. Gelman A, Carlin JB, Stern HS, Rubin DB. Bayesian Data Analysis. London:Chapman & Hall, 1995.
16. Gelfand AE, Smith AFM. Sampling-based approaches to calculating marginal densities. J Am Stat Assoc 85:398–409 (1990).
17. Gelfand AE, Hills SE, Racine-Poon A, Smith AFM. Illustration of Bayesian inference in normal data models using Gibbs sampling. J Am Stat Assoc 85:972–985 (1990).
18. Gelfand AE, Smith AFM, Lee T-M. Bayesian analysis of constrained parameter and truncated data problems using Gibbs sampling. J Am Stat Assoc 87:523–532 (1992).
19. Gelman A. Iterative and non-iterative simulation algorithms. Comput Sci Stat 24:433–438 (1992).
20. Tanner MA. Tools for Statistical Inference - Observed Data and Data Augmentation Methods. Vol 67. Berlin:Springer-Verlag, 1991.
21. Wakefield JC, Smith AFM, Racine-Poon A, Gelfand AE. Bayesian analysis of linear and non-linear population models using the Gibbs sampler. Appl Stat - J Royal Stat Soc Series C 43:201–221 (1994).
22. Gelman A, Rubin DB. Inference from iterative simulation using multiple sequences (with discussion). Stat Sci 7:457–511 (1992).
23. Stenner RD, Merdink JL, Fisher JW, Bull RJ. Physiologically-based pharmacokinetic model for trichloroethylene considering enterohepatic recirculation of major metabolites. Risk Anal 18:261–269 (1998).
24. Davidian M, Giltinan DM. Nonlinear Models for Repeated Measurement Data. London:Chapman & Hall, 1995.