

Purpose and Guidelines for Toxicokinetic Studies within the National Toxicology Program

J. Robert Buchanan,¹ Leo T. Burka,² and Ronald L. Melnick²

¹Mathematics Department, Millersville University, Millersville, PA 17551 USA; ²National Institute of Environmental Health Sciences, Research Triangle Park, NC 27709 USA

Toxicokinetic studies undertaken within the National Toxicology Program are intended to aid the design of toxicology and carcinogenicity studies, help interpret the results of toxicology and carcinogenicity studies with respect to the relationship between toxic effects and external exposure, and define the parameters of dose, distribution, metabolism, and elimination that can be used in human risk assessment. Descriptions of two study designs presented here represent the possible extremes in approaches to toxicokinetic studies. The comprehensive approach is geared toward the development of physiologically based pharmacokinetic models that relate external exposure to target organ dosimetry and addresses the questions: Is the chemical absorbed? How is the chemical metabolized? Where are the chemical and/or metabolites distributed in the body? What are the elimination rate and route of the chemical? What is the effect of dose on absorption, distribution, metabolism, and elimination? The minimal study design is more limited in scope than the comprehensive design and addresses primarily the issues of absorption, distribution, and elimination of the parent chemical. Study protocols for most chemicals lie somewhere between these two extreme approaches. An increased understanding of the relationships between external exposure, target organ dosimetry, and adverse effects should provide greater confidence in making low-dose extrapolations of human risk. This paper focuses on the collection of data from animal toxicokinetic studies. The construction of comparable models to characterize target organ dosimetry in exposed humans would certainly require the use of human parameter values obtained from human tissue samples and volunteers. **Key words:** PBPK modeling, study design, toxicokinetics. *Environ Health Perspect* 105:468–471 (1997)

The purpose of National Toxicology Program (NTP) toxicology/carcinogenicity studies is to identify toxic/carcinogenic effects resulting from exposure to a particular chemical or agent and to characterize dose–response relationships. In support of the toxicology/carcinogenicity studies, toxicokinetic (TK) studies are conducted to characterize the absorption, distribution, metabolism, and elimination (ADME) of xenobiotic materials and to quantify the influence of exposure on those properties. Toxicokinetic studies are proposed to

- help in the design (e.g., dose selection, route of exposure, etc.) of meaningful and useful toxicology studies;
- aid in the interpretation of toxicology studies by increasing our understanding of relationships between exposure, time-dependent target organ dosimetry, and adverse effects; and
- define the parameters of dose, distribution, metabolism, and elimination that can be used in human risk assessment.

Toxicokinetic studies also investigate the effects of sex, species, and age on ADME. Data collected during a toxicokinetic study focus on time profiles of parent chemical and metabolite concentrations in plasma and other tissues and can include rates of chemical absorption, metabolism, and excretion, chemical related changes in blood chemistry, bioavailability, protein binding, and depletion of cofactors.

Ultimately, the data gathered during toxicology/carcinogenicity studies are used to estimate the risk to human health from exposure to a chemical. Characterization of relationships between toxic/carcinogenic endpoints and target organ dosimetry aid in the estimation of human risk. A description of a study design well-suited to the support of human risk assessment follows. This comprehensive design aims to collect the data necessary to develop physiologically based pharmacokinetic (PBPK) models incorporating some degree of anatomical realism and including all relevant biochemical and physiological processes. Prior to the determination of toxicological effects, it is impossible to propose a TK study design that is sufficient for all chemicals. The design proposed below contains elements believed to be necessary for the construction of PBPK models. The relative importance of some measurements may change depending on the compound under study.

Two of the most difficult but critical problems in risk assessment are defining equivalent human dose and delineating the dose–response curve in the low-dose region (i.e., below the range of the experimental data). These problems motivate the recommendations to collect data to support the development and validation of PBPK models. If the only information available to the risk assessor is the applied or external

dose, the estimated human equivalency dose will be based on several assumptions, such as equivalent absorption in laboratory animals and humans. However, uncertainties may exist in the predicted low-dose effects in humans if all of the doses that caused the adverse effect in the animal model occurred above the metabolic saturation exposure, if the chemical is absorbed differently in animals than in humans, or if the chemical is metabolized differently (qualitatively or quantitatively) between species. An objective of toxicokinetic studies is to improve our understanding of the relationship between external exposure and the tissue level of the toxic agent at the target site. By increasing our understanding of the relationships between target organ dosimetry and adverse effects, confidence in low-dose extrapolations of risk to humans is increased. Thus, knowledge of internal dose and factors that influence absorption, distribution, metabolism, and elimination in experimental animals and humans provide a greater scientific rationale for estimating low-dose risk than do estimates that are based solely on external exposure. Exposure- and time-dependent changes in target tissue concentrations of a parent compound and metabolites (especially if the adverse effects of exposure are due to particular metabolites) provide an even stronger scientific foundation for estimating equivalent human dose in various risk assessment models. Toxicokinetic/pharmacokinetic models provide the basis for more accurate low-dose extrapolations to humans. PBPK models can be adapted to different routes of exposure and dosage regimens and can accommodate factors that contribute to interindividual variabilities. Thus, the modeler can also provide valuable information

Address correspondence to J.R. Buchanan, Mathematics Department, Millersville University, PO Box 1002, Millersville, PA 17551-0302 USA.

The authors wish to thank the members of the Toxicokinetic Faculty of the National Institute of Environmental Health Sciences (NIEHS) for their comments and suggestions regarding previous versions of this paper. In addition to the authors, the faculty members are Gary Boorman, John Bucher, Brad Collins, Tom Goehl, Joe Haseman, Michael Kohn, Skip Matthews, Mike Moorman, Fred Parham, Chris Portier, and Claudia Thompson.

Received 21 November 1996; accepted 8 January 1997.

relevant to the distribution of risk in exposed populations. The combination of dosimetry models with mechanistic data (e.g., mutagenicity, altered gene expression) can lead to more realistic biological models that would be useful in predicting site-specific dose responses and assessing human risk.

For most of the chemicals studied by the NTP, the mechanisms of toxicity or carcinogenicity are not known. In the absence of knowledge regarding the mechanism of action of a chemical, it is important to measure the effects of dose on chemical fate in test animals. To maximize the sensitivity of NTP bioassays, chemicals are often studied at doses that are far in excess of those that humans are likely to encounter. For some chemicals the effect of the dose on the test species is determined by measuring the plasma levels of the parent compound and metabolites over the dose range of interest. For other chemicals the critical toxic action may occur in the stomach, gastrointestinal tract, or liver prior to entering the systemic circulation. Even when the amount of parent compound absorbed into the systemic circulation is proportional to the dose administered, the metabolites that may actually account for toxicity associated with exposure to the parent compound may not be formed in or cleared from the target tissues in a manner proportional to dose. These parameters can be determined only through detailed studies of the dose-dependent metabolism and disposition of the chemical in question. For chemicals with low toxicity or which were not carcinogenic under the conditions of the bioassay, there may be little need for studies supporting the development of mechanistic models. However, negative toxicology studies and adequate disposition studies may provide valuable information on positive studies of structurally related chemicals. If the metabolic processes associated with a chemical differ between species, detailed modeling and comprehensive studies may be necessary. Chemicals of moderate or unknown toxicity (and/or equivocal carcinogenicity) that are important to public health may warrant comprehensive studies aimed at identifying the mechanisms of toxicity/carcinogenicity and associated pharmacokinetic parameters.

In the next section, we describe a general approach for a comprehensive toxicokinetic study and make suggestions of areas in which the design could be tailored for a specific chemical. Clewell and Andersen (1) describe the utility and advantages of using PBPK models in interpreting the results of toxicology experiments and in assessing

human risk. In the last section, we describe a minimal approach that would be useful for the design of toxicology/carcinogenicity studies. It should be emphasized that the two approaches represent the extreme ends of a continuum of designs. Most studies will follow protocols that fall somewhere between these extremes. Study designers are free to eliminate elements from the comprehensive design that are inappropriate or unattainable for the chemical of interest or to add features to the minimal approach to answer specific questions of interest. The chemistry and toxicity of a specific compound may change the nature of the study conducted. For example, measurement of parent chemical and metabolite levels in the bile via bile duct cannulation may be useful for chemicals known or thought to undergo significant enterohepatic recirculation. Otherwise, the efforts and expense involved in bile duct cannulation may be unwarranted. Customization of study design should always be considered.

A Comprehensive Toxicokinetic Study Design

To facilitate an understanding between exposure and tissue dosimetry, a toxicokinetic study should aim to answer the following five questions for the chemical under study: 1) Is the chemical absorbed (rate and degree)? 2) Where are the chemical and/or metabolites distributed in the body? 3) How is the chemical metabolized (site and rate of metabolism and identification of metabolites)? 4) What are the elimination rate and elimination route(s) of the chemical and possibly its metabolite(s)? 5) What is the effect of dose on absorption, distribution, metabolism, and elimination?

The answers to these questions are helpful in the construction of pharmacokinetic/toxicokinetic models, design of toxicology studies, and interpretation of animal dose response. Conversely, an existing properly constructed pharmacokinetic/toxicokinetic model should be able to predict answers to these questions. The accuracy of prediction will depend on the degree of physiological and biochemical fidelity captured by the model. If the ultimate goal of the animal study is to estimate risks to human health from chemical exposure, higher quality and more reliable risk estimates can be made from mechanistic models such as well-constructed and validated PBPK models. Such models would certainly require the use of human parameter values obtained from human tissue samples and volunteers. This section covers a proposed study design to support the development of a PBPK model. The reader should keep in mind that this study design is proposed without regard to

a specific chemical. There will no doubt be chemicals for which this proposed design will be inappropriate, inadequate, or impractical. For those chemicals, the following approach should be used as a starting point and modified to address the particular constraints and needs of working with those chemicals.

Animals. The same strain and species used in the toxicology studies should be used in the TK studies. Initial studies may involve one sex of each species (e.g., male rats and mice). The use of more than a single species of test animal is relevant to the question of expected differences in biochemical processes between test animals and humans. If a sex difference in toxic response is found, then TK studies will be necessary in the second sex. Young healthy adults (10–12 weeks of age, matched for body weight) should be used, except for studies examining the effect of age on the ADME of a chemical. Sufficient animals should be included in the TK study design to guarantee a minimum of three data samples per time point of analysis.

Chemical. Radiolabeled chemicals should be used if available or if their synthesis is economically and technically feasible. Radiolabeled chemicals have the advantage that the total amount of parent compound plus metabolites in a tissue at the time of sampling are easily and quickly measured. As noted below, analyses of tissue samples should distinguish the parent compound from its metabolites. If it is not possible to synthesize the radiolabeled parent compound, unlabeled material may be used to determine toxicokinetic properties of the chemical if a sensitive analytical method is available. Unlabeled material may also be used if an objective of the study is to evaluate ADME in the same animals that are on study for toxicological evaluations.

Storage and disposal of carcasses, tissue samples, excreta, and excess radiolabeled chemicals will always be a concern and result in greater expense compared to the use of unlabeled chemicals. It is probably impractical to conduct a long-term exposure study using radiolabeled chemicals, especially in dosed feed or drinking water exposures. Thus it is important that analytical methods of sufficient sensitivity be developed for the parent chemical and metabolites of interest so that the use of radiolabeled chemicals in later stages of studies can be minimized or eliminated.

Analyses. For radiolabeled chemicals, excreta and tissue samples should be assayed for total radioactivity. The samples should then be subjected to a procedure that separates and quantitates the amount

of radioactivity in the parent compound and metabolites (e.g., high pressure liquid chromatography, gas chromatography). Positive identifications should be made for major metabolites and those that account for >10% of the absorbed dose. If the mechanism of toxicity/carcinogenicity is known or thought to be due to the production of a specific metabolite, then if possible, concentrations of that metabolite should be measured at least in the target tissue or organ, blood, and the site of metabolite production. This information is of fundamental importance to the translation of external dose to toxic effect. Other metabolites may need to be identified under some circumstances.

Route of exposure. The same route in the toxicology study and/or the most common route of human exposure should be used, plus the intravenous (iv) route, for estimation of absorption. If feed or drinking water will be the exposure route in the toxicology study, gavage exposure should also be included to estimate the absorption parameter(s). Separate studies may be needed to characterize the effect of vehicle, dose rate, and vehicle volume on absorption and bioavailability of the parent compound.

Dose levels. The same doses used or anticipated in the toxicology studies should be used. A minimum of three doses is recommended. If TK studies are conducted in advance of toxicology studies on the chemical, then the doses should approximate 0.1 median lethal dose (LD_{50}), 0.01 LD_{50} , and 0.001 LD_{50} . Three dose levels may be sufficient to demonstrate dose proportionality; however, additional doses, both higher and lower than those mentioned, would provide a better characterization of the linear kinetic range and a more accurate estimation of the dose level at which the onset of nonlinear kinetics occurs.

Samples. For the iv route of exposure, samples should be taken for each of the tissues: blood (or plasma), fat, muscle, liver, and kidney. If a target organ or tissue for the chemical under study is known or anticipated and is not included in the above list, it should be sampled also. Sampling times will depend on the rate at which the chemical is eliminated or metabolized and on the sensitivity of the analytical method. For an iv administration, samples should be collected at several (8–12) time points. For example, if the plasma concentration of a chemical after a specified dose reaches the limit of detection of the analytical method at approximately 24 hr, samples might be collected at 5, 10, 20, 40, and 60 min and 2, 4, 6, 8, and 24 hr after dosing. Excreta should be collected for analysis at several time points over this same interval. For

volatile chemicals or volatile metabolites, expired air should also be collected.

For many chemicals administered via inhalation, the plasma levels of chemical drop rapidly once an animal is removed from exposure. Blood samples can be taken during exposure if the exposure is performed in a nose-only exposure system. Rats can be sampled from the tail vein, implanted jugular cannula, or by cardiac puncture. Mice are limited to sampling by cardiac puncture. Jugular cannulated animals should be allowed sufficient time to recover from surgery and anesthesia before exposure.

For other non-iv routes of exposure, samples of the same tissues listed for the iv route should be taken (plus skin if the non-iv route is dermal). To characterize the extent of absorption and routes of elimination, additional groups of dosed animals should be included for collection and analyses of excreta, carcass, and (if appropriate) expired air. Animals should be fed the night before dosing. The samples should be taken at several time points (10–12) sufficient to characterize the chemical concentration in plasma or other tissues versus time profiles. Sampling interval and times will depend on the rate of absorption and clearance of the chemical. To characterize the absorption rate and maximum concentration (C_{max}), samples are needed at three distinct times prior to the time at which the peak blood concentration is reached. A pilot study may be needed to help select the appropriate sampling times for this phase of the experiment. Special care must be taken to not disturb the feeding habits of the remaining animals on test during the collection of night-time blood samples.

Partition coefficients. For the development of PBPK models, tissue:blood partition coefficients should be obtained for the parent compound and toxic metabolites in fat, muscle, liver, kidney, and additional major metabolizing tissues, target organs, and storage depots (if not included among the specifically listed tissues). Anomalies between tissue hydrophobicity and partitioning characteristics can be evidence of specific binding in tissues. For volatile compounds, the vial head-space technique (2,3) is often used to measure partition coefficients, whereas nonvolatile compounds may be more amenable to the technique of Jepson et al. (4). For nonvolatile compounds that are cleared slowly, tissue:blood partition ratios may be measured at steady state after repeated exposure to a chemical.

Enzymology. The biochemical realism of the PBPK model depends on an understanding and quantitative characterization

of the processes that convert a xenobiotic to possibly one or more metabolites and the processes that detoxify the parent chemical and metabolites. It may be that the toxic/carcinogenic agent is actually a short-lived reactive intermediate that could be difficult to measure in tissue samples. It may be possible to obtain kinetic constants for the production and degradation of reactive intermediates *in vitro*. Although an *in vitro* study of the parent compound and a metabolite may ignore complicating interactions present *in vivo*, it presents an opportunity to measure production and detoxification rates needed for initial construction of the model. The appropriateness of this simplification of the biochemistry will be revealed during validation of the model. The range of substrate concentrations tested *in vitro* should include the range of concentrations expected *in vivo*, provided sufficiently sensitive analytical techniques are available. High substrate concentrations should also be tested *in vitro* to determine maximal rate (V_{max}). The PBPK model may need to include the kinetics of cofactor changes (e.g., glutathione, reduced nicotinamide adenine dinucleotide, reduced nicotinamide adenine dinucleotide phosphate) or competitive binding between metabolites and/or parent compound. Thus, the kinetics of these changes should also be measured *in vitro*. Use of human tissue samples provides an opportunity to compare metabolic activities between laboratory animals and humans. However, one should also be cautious in making quantitative comparisons because of the potential variability in the quality of normal human tissue.

Model validation. The output of the PBPK model should be compared against experimental data that were not used to estimate parameters for the model. For example, model predictions could be compared against an experiment in the same species as that used to construct the model, but using a different route of exposure, dose level, or dose frequency. Comparison between model predictions and experimental data in a species different from that used to construct the model is also a method of validation of the model structure. This comparison will require model parameter values for the second species. Accurate predictions of cofactor changes can be used as evidence of model validation. If practical and ethical, human tissue or human volunteer studies can be used to determine the validity of the PBPK model before it is considered to be suitable for risk assessment. If the agreement between the model output and experimental data is not reasonable, the new data should be used as a basis for improv-

ing parameter estimates or changing the form of the model. The output of the improved model would then be compared against yet another set of experimental data. Model construction and validation are iterative processes.

Multiple dosing. Single-dose experiments are adequate except if changes in metabolic enzyme levels are anticipated (e.g., induction, suicide inhibition) or if the chemical has a long half-life. In these cases, 5 days to 2 weeks of dosing at the same frequency as in the toxicology study may be needed. If enzymatic changes are important in the biotransformation of the parent compound or its toxic metabolites, the toxicity studies should include measurements of the time-dependent changes in these activities as a function of dose.

Effects of age. If multiple dosing studies provide the same parameter values as single-dose studies, the effect of age on the ADME of the chemical can be evaluated in naive animals of different ages (e.g., 3, 6, 12, and 18 months). The PBPK model parameters can be adjusted for the physiology of aged animals (increased body weight, increased percentage of body fat, etc.). Predictions of the PBPK model should be validated against experimental data collected in aged animals. This may reduce or eliminate the need to expose animals to a xenobiotic for extended periods (e.g., more than 2 weeks) or it may justify the need for a chronic exposure TK study.

A Minimal Toxicokinetic Study Design

In the previous section, we described a study protocol design that could be used in the development of a PBPK model. It is at one end of the spectrum of possible study designs; at the opposite end is a minimal study protocol. Toxicologists are not compelled to make a choice between the two extremes of protocol design. The minimal protocol is a subset of the ideal protocol and does not require the use of radiolabeled compound. This study design is likely to answer only questions 1), 4), and 5) from the list presented at the beginning of the previous section and would greatly help in the design and interpretation of the animal toxicity studies. However, the inclusion of some components of the comprehensive protocol in the minimal protocol (e.g., sam-

pling excreta in addition to blood or measuring partition coefficients) would further improve the utility of the latter protocol in interpreting the results of a toxicology study and in risk assessment. Both study designs share the following common features.

Animals. The same strain and species as in the toxicology studies should be used in the TK studies. Initial studies may involve one sex of one species (e.g., male rats). Young healthy adults (10–12 weeks of age, matched for body weight) should be used, except for studies examining the effect of age on the ADME of a chemical. Sufficient animals should be included in the TK study design to guarantee a minimum of three data samples per time point of analysis.

Chemical. Radiolabeled chemical is not required although it can be used if available.

Route of exposure. The same route in the toxicology study or the most common route of human exposure should be used, in addition to the iv route for estimation of absorption.

Dose levels. The same doses used or anticipated in the toxicity studies should be used. A minimum of three doses is recommended. If TK studies are conducted in advance of toxicology studies on the chemical, the doses should approximate 0.1 LD₅₀, 0.01 LD₅₀, and 0.001 LD₅₀. Three dose levels may be sufficient to demonstrate dose proportionality; however, additional doses may be needed to more accurately estimate where nonlinear kinetics occur.

Samples. For an iv administration, samples should be collected at several (8–12) time points. For example, if the plasma concentration of a chemical after a specified dose reaches the limit of detection of the analytical method at approximately 24 hr, samples might be collected at 5, 10, 20, 40, and 60 min and 2, 4, 6, 8, and 24 hr after dosing. Rats can be sampled from the tail vein, implanted jugular cannula, or by cardiac puncture. Mice are limited to cardiac puncture. If bile duct cannulation is performed, animals should be protected from dehydration. After a steady bile flow is established, bile should be collected at intervals separated by at least 30 min.

For non-iv routes of exposure, blood samples (plasma or serum) should be collected at several time points (10–12) sufficient to characterize the chemical concentration in plasma or other tissues versus

time profiles. Sampling interval and times will depend on the rate of absorption and clearance of the chemical.

Analyses. Blood/plasma samples should be assayed for the administered compound. For some chemicals (e.g., esters that are rapidly hydrolyzed), a major metabolite must be measured instead of the parent chemical.

This design is similar to several protocols already in use at the National Institute of Environmental Health Sciences. The minimal design almost completely ignores the issue of metabolism and identification of the toxic agent (if presumed different than the parent chemical). This design is purposely modest in scale. In most cases, parameters obtained from disposition or toxicokinetic studies in a single sex of a single species provide some indication of the processes involved. If differences in LD₅₀ or health effects between species or sexes are known, it may be necessary to include animals of more than one species or of both sexes. In general, a chemical that is readily absorbed by the male rat will be readily absorbed by the female and by mice and by humans, though the rates will not be identical. Additional *a priori* knowledge about a chemical could further change this proposed design. For instance, study of a chemical known to undergo significant enterohepatic recirculation may call for the use of bile duct-cannulated animals.

REFERENCES

1. Clewell HJ III, Andersen ME. Risk assessment extrapolations and physiological modeling. In: Advances in health risk assessment for systemic toxicants and chemical mixtures, Toxicology and industrial health, vol 1 (Stara JF, Erdreich LS, eds). Princeton, NJ:Princeton Scientific Publishers, 1985;111–113.
2. Gargas ML. Chemical-specific constants for physiologically-based pharmacokinetic models. Technical report 11. Research Triangle Park, NC:Chemical Industry Institute of Toxicology, 1991.
3. Sato A, Nakajima T. A vial-equilibration method to evaluate the drug-metabolizing enzyme activity for volatile hydrocarbons. *Toxicol Appl Pharmacol* 47:41–46 (1979).
4. Jepson GW, Hoover DK, Black RK, McCafferty JD, Mahle DA, Gearhart JM. A partition coefficient determination method for nonvolatile chemicals in biological tissues. *Fundam Appl Toxicol* 22:519–524 (1994).