Journal of Statistical Software

MMMMMM YYYY, Volume VV, Issue II. *http://www.jstatsoft.org/*

httk: R Package for High-Throughput Toxicokinetics

Robert G. Pearce, R. Woodrow Setzer, Cory L. Strope, Nisha S. Sipes, John F. Wambaugh

Abstract

Thousands of chemicals have been profiled by high-throughput screening programs such as ToxCast and Tox21; these chemicals are tested in part because most of them have limited or no data on hazard, exposure, or toxicokinetics. Toxicokinetic models aid in predicting tissue concentrations resulting from chemical exposure, and a "reverse dosimetry" approach can be used to predict exposure doses sufficient to cause tissue concentrations that have been identified as bioactive by high-throughput screening. We have created four toxicokinetic models within a new R software package, **httk**. These models are designed to be parameterized using high-throughput *in vitro* data (plasma protein binding and hepatic clearance), as well as structure-derived physicochemical properties and species-specific physiological data. The package contains tools for Monte Carlo sampling and reverse dosimetry along with functions for the analysis of concentration vs. time simulations. The package can currently use human *in vitro* data to make predictions for 543 chemicals in humans, rats, mice, dogs, and rabbits, including 95 pharmaceuticals and 416 ToxCast chemicals. For 66 of these chemicals, the package includes rat-specific *in vitro* data. This package is structured to be augmented with additional chemical data as they become available. **httk** enables the inclusion of toxicokinetics in the statistical analysis of chemicals undergoing high-throughput screening.

Keywords: high-throughput,Toxcast,httk,toxicokinetics,pharmacokinetics.

1. Introduction

Humans are exposed to thousands of chemicals from the environment and consumer products, most of which have not been tested for toxicity (Park *et al.* 2012; Wambaugh *et al.* 2013b; Egeghy *et al.* 2011; Judson *et al.* 2008). In order to screen for potential bioactivity, *in vitro* data have been generated in the Tox21 (Bucher 2008) and ToxCast (Judson *et al.* 2010) programs using high-throughput screening systems. Over 8500 chemicals have been tested in at least 50 assays (Tox21), and a subset of around 1800 have had nearly 1200 assay endpoints

measured (ToxCast). Recently, high-throughput exposure modeling has provided estimates of daily human exposure for thousands of environmental contaminants (Wambaugh *et al.* 2014). However, linking these hazard and exposure predictions to estimate risk requires development and use of high-throughput toxicokinetics. The terms "pharmacokinetic", "toxicokinetic", and "biokinetic" model have been used somewhat interchangeably in the scientific literature. However, since this package is intended to provide dose context to high-throughput toxicity screening projects, we have selected the term "toxicokinetic" even though we include several compounds with known therapeutic benefits and many others that may not cause adversity for the highest plausible dose.

Toxicokinetics is a field of study for determining the absorption, distribution, metabolism, and excretion of substances in the body (O'Flaherty 1981). The necessary data for toxicokinetics are commonly collected in rats and other animals, but the collection of these data for thousands of chemicals is costly in time, money, and animals (Rovida and Hartung 2009). Creating computational predictive models parameterized with more easily obtained *in vitro* data may help address these problems. Inputting estimated exposures into toxicokinetic models yields information about the steady state and time course concentrations in various parts of the body. These concentrations can then be compared to concentrations that cause biological activity in *in vitro* assays. The models can also be used in a reverse manner, known as reverse toxicokinetics, by predicting the dose needed to produce a specific concentration of interest, such as the *in vitro* AC_{50} or other levels of biological activity as done in Wetmore *et al.* (2012) and Wetmore (2015). Thus chemicals can be ranked based on the ratio of the predicted exposure dose to the back-calculated bioactive dose (Thomas *et al.* 2013), which, due to the linearity of these models, is equal to the ratio of the predicted steady state concentration to the *in vitro* bioactive concentration.

Many basic toxicokinetic models (Wetmore *et al.* 2012; Wetmore 2015; Pelekis *et al.* 1997) only predict steady state plasma concentrations (C_{ss}), assuming a dose rate that is both continuous and constant (e.g., infusion dose). With a more dynamic model, such as a physiologically based toxicokinetic (pbtk) model, we can simulate discrete doses to reach steady state, which we observe to oscillate around the infusion dose prediction. Our pbtk models include multiple compartments with partition coefficients. These models are expressed as a set of mass balance differential equations describing the rate of change of the amount of a substance in each compartment. Chemical-specific physicochemical data and species-specific *in vitro* and physiological data are used in calculating the partition coefficients, clearance, tissue volumes, and blood flows. These *in vitro* data consist of the intrinsic hepatic clearance, Cl_{int}, and the plasma protein binding, f_{ub}. **httk** provides tools for Monte Carlo sampling and reverse dosimetry (Tan *et al.* 2006) along with functions that solve for concentration vs. time curves, steady state concentrations, the number of days to steady state, and other toxicokinetic summary statistics for chemicals as shown in Table 4 with the corresponding abbreviations in Table 1. With this R package we provide data, models, and examples to allow the inclusion of toxicokinetics in statistical analysis of chemical exposure and toxicity for 543 chemicals. The package is structured to be modular and expandable to allow new modeling approaches and chemical data to be added as they become available.

Variable	Name
1compartment	One compartment model (O'Flaherty 1981), shown in Figure 1
3compartment	Three compartment model (Jamei *et al.* 2009), shown in Figure 1
3compartmentss	Three compartment steady state model (Wetmore *et al.* 2012; Wetmore 2015)
BW	Body weight
C_{ss}	Average plasma concentration of a chemical at steady state
Cl_{int}	*In vitro* intrinsic hepatic clearance
$Cl_{metabolism}$	Whole liver hepatic clearance, scaled from Cl_{int}
$Cl_{well-stirred}$	Hepatic clearance modeled with well-stirred approximation using $Cl_{metabolism}$
f_{ub}	Fraction unbound, *in vitro* ratio of unbound to total concentration in plasma
httk	High-throughput toxicokinetics
k_{elim}	Elimination rate
k_{gutabs}	Gut absorption rate, default of 1 h^{-1}
logP	Logarithm (base 10) of octanol to water partition coefficient
pbtk	Physiologically based toxicokinetic model, shown in Figure 1
$Q_{cardiac}$	Cardiac output, blood flow through the heart and lungs
Q_{gfr}	Glomerular filtration rate
Q_{rest}	The difference between $Q_{cardiac}$ and the flow to the liver, kidney, and gut
Q_{tissue}	Blood flow to a tissue
QSAR	Quantitative structure-activity relationship
$R_{blood2plasma}$	Ratio of the blood concentration of a chemical to the plasma concentration
SBML	Systems biology markup language
SMILES	Simplified molecular-input line-entry system
V_{dist}	Volume of distribution, the weighted sum of all partition coefficients

Table 1: List of abbreviations.

2. Methods

Version 1.4 of **httk** is used in this manuscript.

2.1. Models included

The four models in **httk** include: "pbtk", "3compartment", "3compartmentss", and "1compartment"; the predictions and parameters of these models are compared in Table 2. All models currently use only oral and intravenous (i.v.) dosing. The models pbtk and 3compartment, shown in Figure 1, use tissue to unbound plasma partition coefficients calculated with Schmitt's method (Schmitt 2008b) (using f_{ub}, octanol-water partitioning, membrane affinity, acid/base dissociation constants, and tissue compositions) to simulate chemical concentrations over time for multiple tissue compartments. The model pbtk contains separate tissue compartments for the gut, liver, lungs, arteries, veins, and kidneys while the model 3compartment only contains compartments for the liver and gut and is essentially a condensed form of the model pbtk. The tissues contained in `tissue.data` that are unused in each of these models are aggregated into a single compartment termed "rest", whose partition coefficient is calculated by averaging the remaining partition coefficients, weighted by their

species-specific tissue volumes. Absorption from the gut lumen into gut tissue is modeled as a first order process with an arbitrary "fast" absorption rate of 1 h^{-1}. The fraction of the dose absorbed into the system through the gut wall is set to 1 when measured data are unavailable. The gut blood flows directly into the liver, where the hepatic clearance, $Cl_{metabolism}$, is calculated with a unit conversion of Cl_{int} using the density of hepatocytes in the liver (1.1 x 10^8 hepatocytes per gram of liver from Birnbaum *et al.* (1994) and a liver density of 1.05 g/mL from Ito and Houston (2004)). Both models also feature renal elimination by passive glomerular filtration through the kidneys. We assume perfusion-limited tissue (i.e., tissue, red blood cells, and plasma come to equilibrium rapidly with respect to the flow of blood), and a constant $R_{blood2plasma}$ is used throughout the body and is predicted using the hematocrit and the predicted partitioning between red blood cells and plasma.

The models 3compartmentss and 1compartment both contain only plasma without separate compartments for blood and tissue (and thus no partition coefficients). The model 3compartmentss, "ss" standing for steady state, is a single equation for the C_{ss} of the rest-of-body compartment in the model 3compartment resulting from i.v. dosing with $R_{blood2plasma} = 1$. This is the same equation used for determining C_{ss} in previous work (Rotroff *et al.* 2010; Wetmore *et al.* 2012; Wetmore 2015; Wilkinson and Shand 1975) and is equal to the steady state concentration of the model 1compartment resulting from infusion dosing. The model 1compartment features an absorption compartment and a total clearance equal to the sum of the metabolism of the parent compound in the liver, modeled with the "well-stirred" approximation (Wilkinson and Shand 1975; Houston and Carlile 1997), and the renal clearance by passive glomerular filtration. The elimination rate, k_e, is equal to the total clearance divided by the volume of distribution, V_{dist}. V_{dist} is used as the volume of the compartment and is calculated by summing the plasma volume and the products of each tissue to unbound plasma partition coefficient, its corresponding volume , and f_{ub} (Schmitt 2008b).

Among the four models in the package, the simplest model, 3compartmentss, is applicable to the largest number of chemicals, specifically those which are missing information needed to parameterize the other models. It is the only model that does not use partition coefficients and thus does not require logP, and when f_{ub} is below the limit of detection, the model can be used with Monte Carlo to simulate C_{ss} distributions. Thus, f_{ub} below the limit of detection (set to zero in `chem.physical_and_invitro.data` and 0.005 in default parameter lists) and Cl_{int} are the minimum data requirements for running a model. The model 1compartment is included to compare our predictions with *in vivo* experiments which are often characterized by one compartment model parameters (V_{dist} and k_{elim}). We note that fully understanding the kinetics of a given chemical might require additional data on features currently not accessible with high-throughput *in vitro* approaches, such as bioavailability, transporters, protein-binding kinetics, and extra-hepatic or strongly saturable metabolism (Rotroff *et al.* 2010).

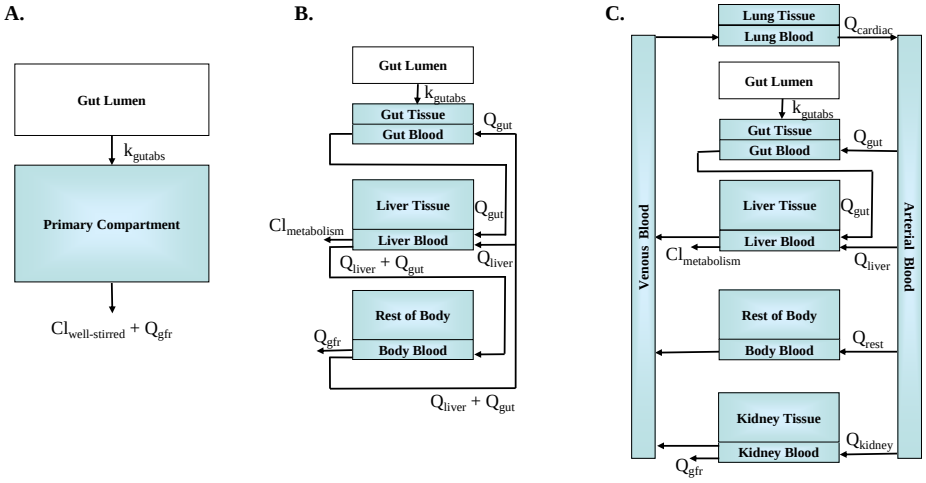

Figure 1: Models(A) 1compartment, (B) 3compartment, and (C) pbtk. In order to preserve mass-balance, Q_{rest} is defined as the difference between $Q_{cardiac}$ and the flow to the liver, kidney, and gut. Variable names are defined in Table 1.

Model	Hepatic clearance	Partition coefficients	Fraction unbound	Hematocrit	Molecular weight	Ratio of blood to plasma	Elimination rate*	Volume of distribution*	Dynamic prediction	Steady state prediction
pbtk	Yes	Yes	Yes	Yes	Yes	Yes	No	No	Yes	Yes
1compartment	No	No	No	No	Yes	No	Yes	Yes	Yes	Yes
3compartment	Yes	Yes	Yes	Yes	Yes	Yes	No	No	Yes	Yes
3compartmentss	Yes	No	Yes	No	Yes	No	No	No	No	Yes

Table 2: Model parameter and prediction comparison. *Partition coefficients are needed in calculating V_{dist}. Clearances and f_{ub} are needed in calculating k_{elim}.

2.2. Model equations

The differential equations below describe changes in the concentrations or amounts of a substance within the model compartments. Although the models are written as changes in tissue concentrations, excluding the gut lumen, the equations actually express changes in the amount of substance in the blood of each tissue divided by the tissue volume with the blood and tissue concentrations related through the out-flowing concentration in blood. A blood flow, Q

(L/day), multiplied by a concentration, C (mol/L), is equal to the amount of the substance entering or leaving a compartment through the blood, QC (mol/day). We define partition coefficients as the ratio of the concentration in a tissue to the unbound concentration in plasma of that tissue; $K_{tissue2pu} = C_{tissue}/(f_{ub}C_{plasma})$. Thus dividing C_{tissue} by $K_{tissue2pu}$ and f_{ub} and then multiplying by $R_{blood2plasma}$ yields the blood concentration of the compartment at equilibrium, $C_{tissue}R_{blood2plasma}/(f_{ub}K_{tissue2pu})$. Assuming perfusion-limited tissue, we substitute this term for the out-flowing blood concentrations (Campbell *et al.* 2012) and assume negligible blood volume fractions in all tissues to justify dividing by the tissue volume without a blood volume fraction and partition coefficient dependency. The flow to the rest-of-body, Q_{rest}, is calculated by subtracting the sum of all the other tissue flows (i.e., gut, liver, and kidney) from the total cardiac output. The glomerular filtration rate, Q_{gfr}, and the hepatic clearance, $Cl_{metabolism}$, (both in L/day) are both multiplied by the unbound plasma concentrations, $C_{tissue}/K_{tissue2pu}$, in the kidney and liver to express the amount of the substance leaving the system. Note that although the units of the clearances, flows, and absorption rate are in days, being consistent with the model outputs, they are initially entered in units of hours. The model 3comparmentss assumes a constant dose rate, k_{dose} (mg/kg BW/day), and the other models use discrete changes in the amount in the gutlumen or venous concentration, depending on which type of dose is specified. The function that is part of the gut lumen equation, $g(t)$ describes the oral dosing schedule. MCSim (Bois and Maszle 1997) was used for converting the model equations into C code, which is used with **deSolve** (Soetaert *et al.* 2010a) in solving each system of equations.

pbtk equations

$$\frac{d}{dt}A_{gutlumen} = -k_{gutabs}A_{gutlumen} + g(t)$$

$$\frac{d}{dt}C_{gut} = \left(k_{gutabs}A_{gutlumen} + Q_{gut}\left(C_{art} - \frac{R_{blood2plasma}}{K_{gut2pu}f_{ub}}C_{gut}\right)\right)\Big/V_{gut}$$

$$\frac{d}{dt}C_{liver} = \left(Q_{liver}C_{art} + \frac{Q_{gut}R_{blood2plasma}}{K_{gut2pu}f_{ub}}C_{gut} - \frac{(Q_{liver}+Q_{gut})R_{blood2plasma}}{K_{liver2pu}f_{ub}}C_{liver} - \frac{Cl_{metabolism}}{K_{liver2pu}}C_{liver}\right)\Big/V_{liver}$$

$$\frac{d}{dt}C_{ven} = \left(\left(\frac{Q_{liver}+Q_{gut}}{K_{liver2pu}}C_{liver} + \frac{Q_{kidney}}{K_{kidney2pu}}C_{kidney} + \frac{Q_{rest}}{K_{rest2pu}}C_{rest}\right)\frac{R_{blood2plasma}}{f_{ub}} - Q_{cardiac}C_{ven}\right)\Big/V_{ven}$$

$$\frac{d}{dt}C_{lung} = Q_{cardiac}\left(C_{ven} - \frac{R_{blood2plasma}}{K_{lung2pu}f_{ub}}C_{lung}\right)\Big/V_{lung}$$

$$\frac{d}{dt}C_{art} = Q_{cardiac}\left(\frac{R_{blood2plasma}}{K_{lung2pu}f_{ub}}C_{lung} - C_{art}\right)\Big/V_{art}$$

$$\frac{d}{dt}C_{rest} = Q_{rest}\left(C_{art} - \frac{R_{blood2plasma}}{K_{rest2pu}f_{ub}}C_{rest}\right)\Big/V_{rest}$$

$$\frac{d}{dt}C_{kidney} = \left(Q_{kidney}C_{art} - \frac{Q_{Kidney}R_{blood2plasma}}{K_{kidney2pu}f_{ub}}C_{kidney} - \frac{Q_{gfr}}{K_{kidney2pu}}C_{kidney}\right)\Big/V_{kidney}$$

3compartment equations

$$\frac{d}{dt}A_{gutlumen} = -k_{gutabs}A_{gutlumen} + g(t)$$

$$\frac{d}{dt}C_{gut} = \left(k_{gutabs}A_{gutlumen} + \frac{Q_{gut}R_{blood2plasma}}{K_{rest2pu}f_{ub}}C_{rest} - \frac{Q_{gut}R_{blood2plasma}}{K_{gut2pu}f_{ub}}C_{gut}\right)\Big/V_{gut}$$

$$\frac{d}{dt}C_{liver} = \left(\frac{Q_{gut}R_{blood2plasma}}{K_{gut2pu}f_{ub}}C_{gut} + \frac{Q_{liver}R_{blood2plasma}}{K_{rest2pu}f_{ub}}C_{rest} - \left(\frac{(Q_{liver}+Q_{gut})R_{blood2plasma}}{f_{ub}K_{liver2pu}} + \frac{Cl_{metabolism}}{K_{liver2pu}}\right)C_{liver}\right)\Big/V_{liver}$$

$$\frac{d}{dt}C_{rest} = \left(\frac{(Q_{gut}+Q_{liver})R_{blood2plasma}}{f_{ub}K_{liver2pu}}C_{liver} - \left(\frac{R_{blood2plasma}}{f_{ub}}\Big(Q_{gut}+Q_{liver}\Big) + Q_{gfr}\right)\frac{C_{rest}}{K_{rest2pu}}\right)\Big/V_{rest}$$

1compartment equations

$$\frac{d}{dt}A_{gutlumen} = -k_{gutabs}A_{gutlumen} + g(t)$$

$$\frac{d}{dt}C_{rest} = \frac{k_{gutabs}}{V_{dist}}A_{gutlumen} - k_{elim}C_{rest}$$

3compartmentss

$$C_{ss} = k_{dose}\Big/\left(f_{ub}Q_{gfr} + \frac{(Q_{liver}+Q_{gut})f_{ub}Cl_{metabolism}}{(Q_{liver}+Q_{gut}) + f_{ub}Cl_{metabolism}}\right)$$

Data table	Description
chem.invivo.PK.data	This data set includes time and dose specific measurements of chemical concentrations in tissues taken from animals administered control doses of the chemicals either orally or intravenously. These plasma concentration-time data are from rat experiments reported in public sources. Toxicokinetic data were retrieved from those studies by the Netherlands Organisation for Applied Scientific Research (TNO) using curve stripping (TechDig v2). These data are provided for statistical analysis as in Wambaugh *et al.* (2015).
chem.invivo.PK.summary.data	This data set summarizes the time course data in the chem.invivo.PK.data table. Maximum concentration (Cmax), time integrated plasma concentration for the duration of treatment (AUC.treatment) and extrapolated to zero concentration (AUC.infinity) as well as half-life are calculated. Summary values are given for each study and dosage.
chem.physical_and_invitro.data	This data set contains the necessary information to make basic, high-throughput toxicokinetic predictions for compounds, including f_{ub}, Cl_{int}, molecular weight, logP, logMA (membrane affinity), and pKa.
tissue.data	This data set contains values from Schmitt (2008a) describing the composition of specific tissues and from Birnbaum *et al.* (1994) describing volumes of and blood flows to those tissues, allowing parameterization of toxicokinetic models for human, mouse, rat, dog, or rabbit.
physiology.data	This data set contains additional physiological values necessary to paramaterize a toxicokinetic model for human, mouse, rat, dog, or rabbit.
Wetmore.data	This data set gives the chemical-specific predictions for serum concentration at steady state resulting from infusion exposure at a constant rate, as published in a series of papers from Barbara Wetmore's group (Wetmore *et al.* 2012, 2013; Wetmore 2015) at the Hamner Institutes for Life Sciences. Predictions include the median and 90% interval in μM and mg/L. Calculations were made using the 1 and 10 μM *in vitro* measured clearances.

Table 3: List of data tables in the package.

2.3. In vitro chemical data

In vitro experiments provide empirical data for two model parameters. The first, Cl_{int}, the intrinsic hepatic clearance of the parent compound by primary hepatocytes (substrate depletion approach), was measured in a well on a multi-compound plate (Shibata *et al.* 2002). This was determined by dividing the *in vitro* clearance of the unbound parent chemical by the fraction of chemical unbound in the hepatocyte intrinsic clearance assay provided in the parameter lists, which was estimated using a distribution coefficient calculated from pKa and the method of Kilford *et al.* (2008). The second *in vitro* measurement is f_{ub}, assessed using rapid equilibrium dialysis (RED) in which two wells are separated by a membrane that is permeable by smaller molecules but prevents the plasma protein added to one well from

migrating to the other well (the relative chemical concentration in the two linked wells gives the free fraction of chemical) (Waters *et al.* 2008).

For non-pharmaceutical chemicals, *in vitro* experimental data were obtained primarily from Wetmore *et al.* (2012), Wetmore (2015), and Tonnelier *et al.* (2012) for humans and Wetmore *et al.* (2013) for rats. For pharmaceutical compounds these values are compiled from Obach (1999), Jones *et al.* (2002), Naritomi *et al.* (2003), Ito and Houston (2004), Riley *et al.* (2005), Schmitt (2008a), and Obach *et al.* (2008). These data are contained in `chem.physical_and_invitro.data`.

2.4. Physicochemical properties

Physicochemical properties were collated from various sources: molecular weight and structure are determined from the DSStox database (http:// www.epa.gov/ncct/dsstox), and octanol to water partitioning is predicted for most compounds with EPA's Estimation Program Interface (EPI) Suite (http://www.epa.gov/tsca-screening-tools/epi-suitetm-estimation-program-interface). EPI Suite quantitative structure activity relationships (QSARs) were used to estimate octanol to water partitioning (logP) if Simplified Molecular Input Line Entry System (SMILES) descriptions of chemical structure were available and the QSARs did not fail for that structure. In addition to QSAR model estimates, EPI Suite contains a database of experimentally obtained octanol to water partition coefficients that were used in place of estimated values when available. Where available, ionization association/dissociation equilibrium constants (pKa) were curated from the literature; otherwise predictions were made from structure using the SPARC (SPARC Performs Automated Reasoning in Chemistry) model (Hilal *et al.* 1995). Experimental values for membrane affinities (i.e., lipid-bilayer to water concentration ratios) were taken from Endo *et al.* (2011) or predicted using a regression of the data (Endo *et al.* 2011) when unavailable based on octanol to water partitioning and temperature in the calculation of partition coefficients. These data are contained in `chem.physical_and_invitro.data`.

2.5. Physiological and tissue data

The tissue data needed for calculating partition coefficients, taken from the corrected table in Schmitt (2008a), include: cellular and water fractions of total volume, lipid and protein fractions of cellular volume, lipid fractions of the total lipid volume, and the pH of each tissue. A default plasma pH of 7.4 is taken from Schmitt (2008b) in calculating ionization. The partition coefficient for the mass and volume of the body unaccounted for by the tissues included in Schmitt (2008b) is calculated with the averages of the fractional volumes and pH of these tissues, excluding red blood cells. Temperature data used in the regression for membrane affinity in the calculation of partition coefficients are taken from Robertshaw and Reece (2004), Stammers (1926), and Gordon (1993). Tissue flows, volumes, liver density, hematocrit, and glomerular filtration are taken from Birnbaum *et al.* (1994). Tissue volumes are scaled linearly to body weight while the flows, including glomerular filtration, are scaled by body weight to the 3/4 power (Campbell *et al.* 2012). The fractional volume of protein in plasma is calculated by dividing the protein concentration in plasma from Gardner and Scott (1980) by the density of plasma (calculated with the specific gravity of plasma from Trudnowski and Rico (1974) and the density of water from Weast and Astle (1982)). The available data for the fraction of a dose absorbed into the gutlumen are taken from Naritomi

et al. (2003). The remaining data are taken from Davies and Morris (1993). These data are included in the `physiology.data` and `tissue.data` tables that are accessible in the package.

2.6. Determination of steady state

Although the discrete dosing in our models produces an oscillating steady state, we use the steady state resulting from equivalent oral infusion dosing at a constant rate, calculated by analytically solving the differential equations at steady state, to determine when steady state is reached. The day a chemical reaches steady state is found by determining when the average concentration of the numerically solved solution for a given day falls within a specified percent of the analytic solution from oral infusion dosing, `calc_css` defaulting to 1%.

2.7. Monte Carlo sampler

The package contains a Monte Carlo sampler, `monte_carlo`, used in `calc_mc_css` with the model 3compartmentss for probabilistically simulating biological variability and measurement uncertainty in parameters determining C_{ss} (Thomas *et al.* 1996). Normal distributions, truncated to ensure positive values are used in the sampler with mean values equal to the model parameters and default coefficients of variation of 0.3. Body weight, liver volume and blood flow, cellular density in the liver, Q_{gfr}, and $Cl_{metabolism}$ are all varied in this manner. f_{ub} is drawn from a censored distribution with identical properties to the other distributions, where values are sampled from a uniform distribution between 0% and the limit of detection (default of 1% unbound) at a rate proportional to the number of samples from the truncated normal distribution below the limit of detection. f_{ub} below the limit of detection (set to zero in `chem.physical_and_invitro.data`) is set to a default value of 0.005 in the model parameters. For each chemical, a default of 1000 different combinations of parameters are used to determine C_{ss}. These concentrations are determined with doses of 1 mg/kg BW/day but, given the linear concentration response of the models, can be extrapolated to other doses with `calc_mc_css`. Using `calc_mc_oral_equiv`, we can, in a reverse manner, back calculate the dose for a given concentration and quantile. The functions `get_wetmore_css` and `get_wetmore_oral_equiv` perform the same operations on doses and concentrations using the published C_{ss} results from Wetmore *et al.* (2012), Wetmore *et al.* (2013), and Wetmore (2015), contained in the `Wetmore.data` table, which used the same *in vitro* data as contained in the package. However, these data only contain the 5%, median, and 95% quantiles for humans and the median for rats. These results were obtained with the SimCYP population simulator (Jamei *et al.* 2009) in a manner identical to the default simulation in `calc_mc_css` with two exceptions. The Wetmore data assumed $f_{ub} = 0.005$ for chemicals with f_{ub} below the limit of detection instead of sampling the value from a censored distribution, and the $Cl_{metabolism}$ values were accepted as nonzero if the p-value was less than 0.1 instead of 0.05 as used in our sampler.

Function	Description
add_chemtable	Adds a table of chemical data to the data tables contained in the package.
calc_analytic_css	Calculates C_{ss} and blood concentrations for the four models used in the package from infusion dosing at a constant rate.
calc_css	Calculates the max and average steady state concentrations along with the day steady state is reached from the numerical solution.
calc_elimination_rate	Calculates k_{elim} for a one compartment model due to the liver and kidneys, dividing the total clearance by V_{dist}.
calc_hepatic_clearance	Calculates the hepatic clearance for a well-stirred model or other type if specified. (Ito and Houston 2004)
calc_mc_css	Monte Carlo simulation of the model 3compartmentss.
calc_mc_oral_equiv	Converts C_{ss} to an oral equivalent dose using a concentration obtained from `calc_mc_css`.
calc_rblood2plasma	Calculates the ratio of chemical concentration in blood to plasma.
calc_stats	Calculates the area under the curve, mean, and peak values for the blood or plasma concentration of either a specified chemical or all chemicals for a given simulation.
calc_total_clearance	Calculates the total clearance rate for a one compartment model where clearance is equal to the sum of the well-stirred metabolism by the liver and glomerular filtration in the kidneys.
calc_vdist	Calculates the volume of distribution for a one compartment model. (Schmitt 2008b)
export_pbtk_jarnac	Exports the model pbtk to Jarnac. (Sauro and Fell 2000)
export_pbtk_sbml	Exports the model pbtk to SBML. (Hucka *et al.* 2003)
get_cheminfo	Provides a list of CAS numbers along with compound names, logP, pKa, molecular weight, Cl_{int} and its p-value, and f_{ub} if specified for chemicals with sufficient data for a given model.
get_wetmore_cheminfo	Provides the names and CAS numbers of chemicals with information from Wetmore *et al.* (2012), Wetmore *et al.* (2013), and Wetmore (2015).
get_wetmore_css	Retrieves C_{ss} as a result of oral infusion dosing from Wetmore *et al.* (2012), Wetmore *et al.* (2013), and Wetmore (2015).
get_wetmore_oral_equiv	Converts C_{ss} to an oral equivalent dose using the values from Wetmore *et al.* (2012), Wetmore *et al.* (2013), and Wetmore (2015).
lump_tissues	Lumps tissue flows, volumes, and input partition coefficients based on specified grouping.
monte_carlo	Runs a monte carlo simulation of a given model.
parameterize_1comp	Parameterizes the model 1compartment.
parameterize_3comp	Parameterizes the model 3compartment.
parameterize_pbtk	Parameterizes the model pbtk.
parameterize_schmitt	Parameterizes `predict_partitioning_schmitt`.
parameterize_steady-state	Parameterizes the model 3compartmentss, used in Wetmore *et al.* (2012) and Wetmore (2015).
predict_partitioning_schmitt	Predicts partition coefficients using Schmitt's method (Schmitt 2008b).
solve_1comp	Solves the model 1compartment.
solve_3comp	Solves the model 3compartment.
solve_pbtk	Solves the model pbtk.

Table 4: List of functions in the package. Models are described in Table 2. Parameters are defined in Table 1. Jarnac and SBML are external languages for systems biology models.

3. Examples

The following examples are run with version 1.4 and may not generate the same outputs as other versions. If using **httk** for regulatory purposes, archive a copy of the version used. To check if the version installed is 1.4:

```
R> sessionInfo()$otherPkgs$httk$Version == "1.4"
```

3.1. Accessing and changing model parameters

httk allows the user to access and change the parameters used in each of the models. The models each contain their own `parameterize` function that generates a list of the parameters required by the model. For example, to get a list of parameters for the pbtk model of triclosan in a rat:

```
R> parameters <- parameterize_pbtk(chem.name = "triclosan", species = "rat")
```

To see the effect a change in parameters has on the model, we can modify the desired entries in the list and use the new parameter list as an input for the parameters argument of a function that uses that model. For example, to change the f_{ub} in the previous parameters list to 0.1 from the default of 0.005 (noting the warning that f_{ub} is below the limit of detection) and use it in a simulation of the pbtk model for a single dose of 1 mg/kg BW of triclosan in a rat:

```
R> parameters["Funbound.plasma"] <- 0.1
R> out <- solve_pbtk(parameters=parameters)
```

Individual parameters such as the $R_{blood2plasma}$, total clearance, V_{dist}, metabolic clearance, and k_{elim} can also be calculated using the functions with the prefix `calc` followed by the parameter name and the same arguments as the above `parameterize` function.

3.2. Making data frames and tables

In order to compare predictions or models, we can construct tables or data frames. Suppose we want to look at how C_{ss} at 1 mg/kg BW/day compares for the model pbtk, the median of the Monte Carlo simulation, and the Wetmore data. We can construct a data frame (used with **ggplot2** (Wickham 2009) in the following examples) containing these data with a for loop. The intersection of `get_wetmore_cheminfo` and `get_cheminfo` contains all the CAS numbers that will work for all three functions. In the example below, setting `model` to "pbtk" in `get_cheminfo` removes the chemicals from the list with f_{ub} below the limit of detection. This is the same as setting `exclude.fub.zero` to true. However, we could include these chemicals by using the default model option of "3compartmentss", and f_{ub} would then automatically be set to 0.005.

```
R> table <- NULL
R> for(this.cas in intersect(get_cheminfo(model = "pbtk"),
+  get_wetmore_cheminfo())){
+  this.row <- as.data.frame(this.cas)
+  this.row <- cbind(this.row, as.data.frame(calc_analytic_css(
```

```
+  chem.cas = this.cas, model = "pbtk", output.units = "mg/L")))
+  this.row <- cbind(this.row, as.data.frame(get_wetmore_css(
+  chem.cas = this.cas, which.quantile = .50)))
+  this.row <- cbind(this.row, as.data.frame(calc_mc_css(
+  chem.cas = this.cas, which.quantile = .50)))
+  table <- rbind(table, this.row)
+  }
R> colnames(table) <- c("CAS", "PBTK", "Wetmore", "MC")
```

3.3. Plotting

Concentration vs time

The function `solve_pbtk` has the option of returning plots for the compartment concentrations vs. time, but to see how C_{ss} resulting from discrete dosing deviates from the average steady state concentration, we can make a plot with **ggplot2** that includes a horizontal line through the y axis at the predicted C_{ss} for oral infusion dosing (Figure 2). We calculate the analytic C_{ss} and enter it into `geom_hline` as the y intercept and add all the other options to our ggplot object.

```
R> library("ggplot2")
R> out <- solve_pbtk(chem.name = "Bisphenol A", days = 50, doses.per.day = 3)
R> plot.data <- as.data.frame(out)
R> css <- calc_analytic_css(chem.name = "Bisphenol A")
R> c.vs.t <- ggplot(plot.data,aes(time, Cplasma)) + geom_line() +
+  geom_hline(yintercept = css) + ylab("Plasma Concentration (uM)") +
+  xlab("Day") + theme(axis.text = element_text(size = 16), axis.title =
+  element_text(size = 16), plot.title = element_text(size = 17)) +
+  ggtitle("Bisphenol A")
R> print(c.vs.t)
```

This example plots the concentration vs. time of 1 mg/kg BW/day of Bisphenol A broken into three doses per day. The same plots can be made for the other models by substituting one of the other two solve functions, `solve_3comp` or `solve_1comp`, for `solve_pbtk` and setting the `model` argument of `calc_analytic_css` to the corresponding model. These three functions also have the option of simulating a single oral or i.v. dose and setting the initial values of each compartment with units matching the specified output units (default is μM).

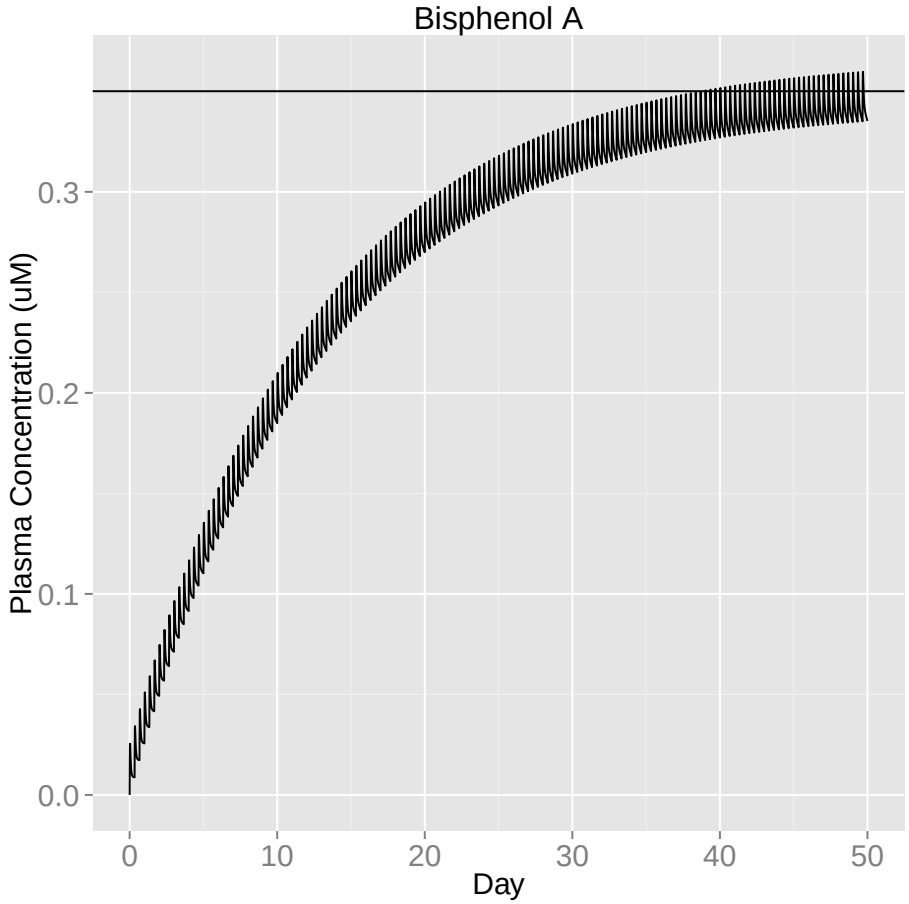

Figure 2: C_{ss} at 3 doses per day, 1 mg/kg BW/day.

Days to steady state histogram

Creating histograms can allow us to visualize how a given value varies across all the chemicals contained within the package. To create a histogram using **ggplot2** of the number of days to steady state, we must first set up a for loop with `get_cheminfo` and `calc_css` to generate a vector containing the data. Vectors containing the average and maximum concentrations at steady state are also generated in this example, avg and max. The data contained in the days vector are then plotted as a histogram (Figure 3). We can just as easily create a histogram containing the average or maximum steady state concentrations by substituting avg or max for days.

```
R> library("ggplot2")
R> days <- NULL
R> avg <- NULL
R> max <- NULL
R> for(this.cas in get_cheminfo()){
+  css.info <- calc_css(chem.cas = this.cas, doses.per.day = 1,
```

```
+  suppress.messages=T)
+  days[[this.cas]] <- css.info[["the.day"]]
+  avg[[this.cas]] <- css.info[["avg"]]
+  max[[this.cas]] <- css.info[["max"]]
+  }
R> days.data <- as.data.frame(days)
R> hist <- ggplot(days.data, aes(days)) +
+  geom_histogram(fill = "blue", binwidth = 1/6) + scale_x_log10() +
+  ylab("Number of Chemicals") + xlab("Days") + theme(axis.text =
+  element_text(size = 16), axis.title = element_text(size = 16))
R> print(hist)
```

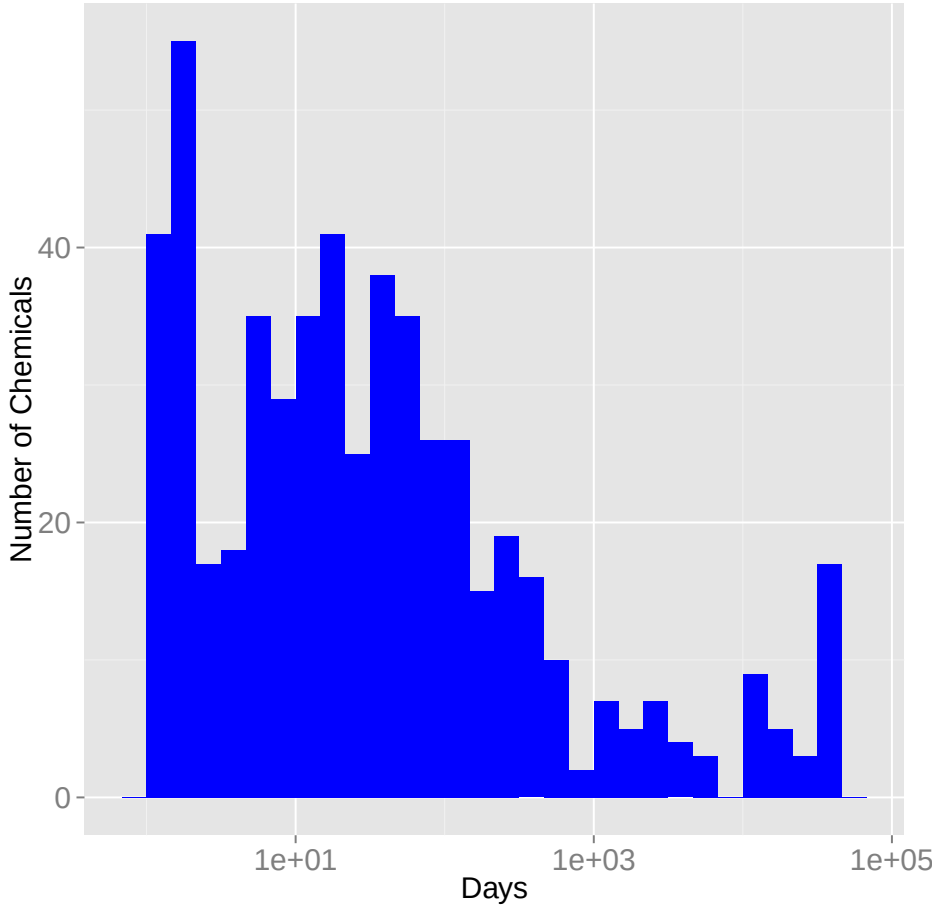

Figure 3: Days to steady state histogram.

Average vs maximum concentration

We can compare the average and maximum concentrations at steady state using the average and maximum concentration at steady state vectors, avg and max, from the previous example.

The vectors are bound into a data frame and plotted with a line through the origin with a slope of 1 (Figure 4).

```
R> library("ggplot2")
R> avg.max.data <- as.data.frame(cbind(avg, max))
R> avg.vs.max <- ggplot(avg.max.data, aes(avg, max)) + geom_point() +
+  geom_abline() + scale_x_log10() + scale_y_log10() +
+  xlab("Average Concentration at Steady State (uM)") +
+  ylab("Max Concentration at Steady State (uM)") +
+  theme(axis.text = element_text(size = 16),
+  axis.title = element_text(size = 16))
R> print(avg.vs.max)
```

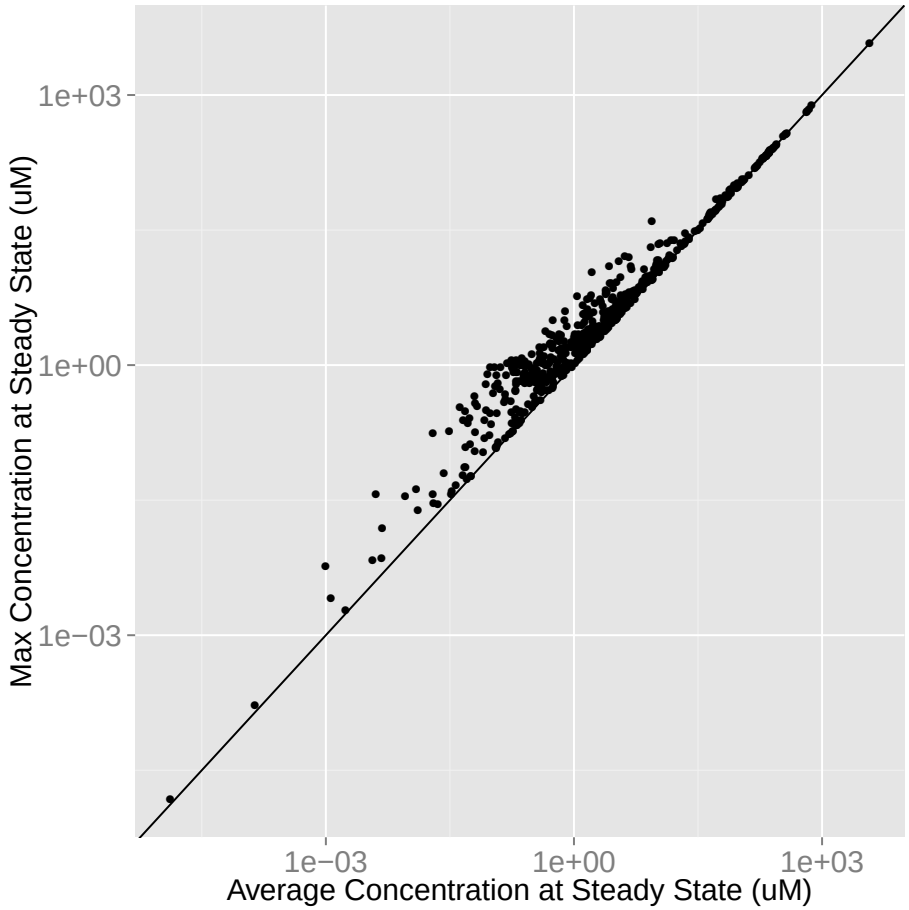

Figure 4: Average vs. maximum concentration at steady state for 1 dose per day, 1 mg/kg BW/day.

3.4. Calculating AUC, peak, and mean values

The function `calc_stats` calculates the area under the curve (AUC), peak, and mean concen-

trations of any of the `solve` functions. If a chemical name or CAS number is specified, it will calculate the specified statistics for that chemical, and if not, it will calculate the values for all chemicals with sufficient data. To calculate the peak statistics for all chemicals simulated for 10 days at 1 mg/kg BW/day with 3 doses per day and a list containing the AUC, peak, and mean for a single 1 mg dose of triclosan over 10 days, we have:

```
R> all.peak.stats <- calc_stats(days=10, doses.per.day = 3, stats = "peak")
R> triclosan.stats <- calc_stats(days=10, chem.name = "triclosan")
```

3.5. Monte Carlo sampler

The functions `calc_mc_css` and `get_wetmore_css` generate C_{ss} vectors of Monte Carlo samples and their quantiles. While `calc_mc_css` generates new values using the sampler, `get_wetmore_css` retrieves literature values from `Wetmore.data`. Below are examples of these two functions, comparing the medians of the Wetmore data in humans for 1 mg/kg BW/day of Bisphenol A with the `calc_mc_css` simulation with probability distributions containing a third of the standard deviation, half the limit of detection for f_{ub}, and double the number of samples of the parameters used in Wetmore *et al.* (2012) and Wetmore (2015).

```
R> get_wetmore_css(chem.cas="80-05-7", daily.dose = 1,
+  which.quantile = 0.5, output.units = "uM")

R> calc_mc_css(chem.cas = "80-05-7", daily.dose = 1, which.quantile = 0.5,
+  censored.params = list(Funbound.plasma = list(cv = 0.1, lod = 0.005)),
+  vary.params = list(BW = 0.15, Vliverc = 0.15, Qgfrc = 0.15,
+  Qtotal.liverc = 0.15, million.cells.per.gliver = 0.15, Clint = 0.15),
+  output.units = "uM", samples = 2000)
```

The oral equivalent functions convert C_{ss} into a dose. Below is an example of a 50 μM C_{ss} of Bisphenol A converted to an oral equivalent dose using the Wetmore data for the 95^{th} quantile of human C_{ss}. We can call `calc_mc_oral_equiv` in the same manner, passing additional arguments to `calc_mc_css` within the function and specifying any quantile we want.

```
R> get_wetmore_oral_equiv(50, chem.cas = "80-05-7")
```

We can also use the Monte Carlo sampler used in `calc_mc_css`, `monte_carlo`, to perform the same simulations using another model. Setting the `return.samples` argument of the `calc_mc` or `monte_carlo` functions to true, we can generate the sampling distribution for the Monte Carlo simulation from which the quantiles are calculated. To perform a Monte Carlo simulation on zoxamide (Figure 5) with the model pbtk with the same limit of detection and coefficients of variation of two thirds the size of those used in `calc_mc_css`, we have:

```
R> vary.params <- NULL
R> params <- parameterize_pbtk(chem.name = "Zoxamide")
R> for(this.param in names(subset(params,
+  names(params) != "Funbound.plasma"))) vary.params[this.param] <- .2
```

```
R> censored.params <- list(Funbound.plasma = list(cv = 0.2, lod = 0.01))
R> set.seed(1)
R> out <- monte_carlo(params, cv.params = vary.params,
+  censored.params = censored.params, return.samples = T,
+  model = "pbtk", suppress.messages = T)
R> zoxamide <- ggplot(as.data.frame(out), aes(out)) +
+  geom_histogram(fill="blue", binwidth=1/6) + scale_x_log10() +
+  ylab("Number of Samples") +  xlab("Steady State Concentration (uM)") +
+  theme(axis.text = element_text(size = 16),
+  axis.title = element_text(size = 16))
R> print(zoxamide)
```

The `out` vector is then plotted in a similar way to the `days` vector in the previous histogram example.

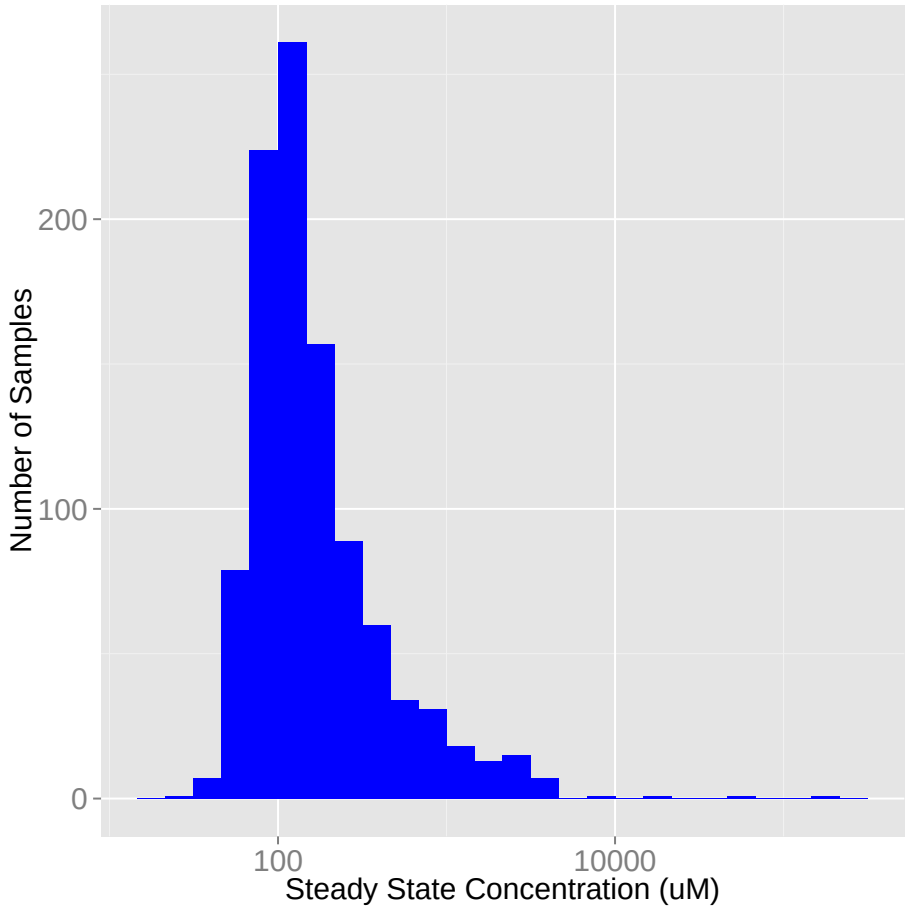

Figure 5: Sampling distribution of zoxomide C_{ss} from the model pbtk.

3.6. Adding a tissue

The fractional volumes and pH values from Schmitt (2008a) needed to calculate the partition

coefficients are contained in `tissue.data`. New tissues can be added to this table to generate their partition coefficients. We can add thyroid to the tissue data by making a row containing its data, subtracting the volumes and flows from the rest-of-body, and binding the row to `tissue.data`. Here we assume it contains the same partition coefficient data as the spleen and a tenth of the volume and blood flow:

```
R> new.tissue <- tissue.data[tissue.data$Tissue == "spleen",]
R> new.tissue[, "Tissue"] <- "thyroid"
R> new.tissue[, 11:20] <- new.tissue[, 11:20] / 10
R> tissue.data[tissue.data$Tissue == "rest", 11:20] <-
R> tissue.data[tissue.data$Tissue == "rest", 11:20] - new.tissue[, 11:20]
R> tissue.data <- rbind(tissue.data, new.tissue)
```

We can also choose what tissues we want lumped together or in the rest-of-body compartment. The `tissuelist` argument in `parameterize_pbtk` contains a list of the desired compartment names, each containing a vector of the names of the tissues in `tissue.data` to be lumped together in that compartment. All unspecified tissues in `tissue.data` are lumped together in the rest-of-body. Lumped flows and volumes are calculated through addition of the individual component flows and volumes while the lumped partition coefficients are calculated through dividing the sum of the products of the partition coefficients and their corresponding compartment volumes by the new lumped volume. To generate the parameters for a model with kidneys, thyroid, a liver compartment combining the liver and gut, and a rest-of-body compartment:

```
R> compartments <- list(liver = c("liver", "gut"),
+  kidney = c("kidney"), thyroid = c("thyroid"))
R> parameterize_pbtk(chem.name="Nicotine", tissuelist = compartments)
```

No matter which compartments we specify, the liver volume as well as the gut, liver, and kidney flows are returned for the calculation of clearance and metabolism.

3.7. Export functions

Jarnac (Sauro and Fell 2000) and SBML (Hucka *et al.* 2003) are commonly used languages for systems biology models of cellular and physiological processes. In the event that a modeler wishes to couple such a model to a toxicokinetic model, we provide functions to export model equations and chemical-specific parameters to these languages. The two functions, `export_pbtk_sbml` and `export_pbtk_jarnac`, have the same arguments and only differ in the file extension names (.xml and .jan) entered into the `filename` argument. Both use liters as the units for volume, but the amounts are unitless and to be determined by the user. If we suppose that we enter an initial amount of 1 mg in the gut lumen, then all the other compartments will contain amounts in mg. Below is a call of an export function for a dose of 1 given to a rat.

```
R> export_pbtk_sbml(chem.name = "Bisphenol A", species = "Rat",
+  initial.amounts = list(Agutlumen = 1), filename = "PBTKmodel.xml")
```

Concluding remarks

The R software platform is increasingly being used for the statistical analysis of mathematical models (Wambaugh *et al.* 2015; Gelman *et al.* 2013). With the launch of the package **odesolve** (Setzer 2001), which was expanded and replaced by **deSolve** (Soetaert *et al.* 2010b), R can be used to solve models consisting of systems of differential equations. R further allows organization and handling of large data sets, making it especially suitable for analyzing the results from high-throughput experiments (Judson *et al.* 2010). Given the reliance on mathematical models and *in vitro* testing to prioritize investigation of the large number of relatively untested environmental chemicals (Wetmore *et al.* 2012; Wetmore 2015), software platforms such as R allow the systematic statistical evaluation of the performance of these technologies (Wambaugh *et al.* 2015). **httk** allows the simulation of four toxicokinetic models for 543 chemicals (including 416 ToxCast chemicals and 95 pharmaceuticals) in humans, rats, mice, dogs, and rabbits.

All models in **httk** are parameterized using data on key determinants of toxicokinetics that can be measured *in vitro* using relatively high-throughput methods (Wetmore *et al.* 2012; Wetmore 2015). The package includes toxicokinetic models ranging from a one compartment model to a PBTK model, but even the PBTK model is relatively spare, with most tissues lumped into a rest-of-body compartment. Rowland (2004) argued that the "best" model is the one that most reliably answers the question at hand (Rowland *et al.* 2004). Thus, the most parsimonious model — that is, the simplest, most easily understood model allowing useful predictions — should be preferred (Chiu and White 2006). However, since PBTK models allow the incorporation of additional, physiological information, we may expect our model pbtk to be the most accurate on average. We hope these models provide predictions of chemical-specific toxicokinetics as informed by *in vitro* data and physicochemical properties without introducing errors from unnecessary assumptions (Rowland *et al.* 2004; Chiu and White 2006). The `parameterize_pbtk` function provides parameter estimates for more complex models as needed (Yang and Lu 2007), though our solvers are currently limited to the four model structures. In future versions, we expect to have the ability to add new compartments using these parameters and simulate dermal and inhalation exposure.

The **httk** package provides functions for the application of Monte Carlo methods, *in vitro-in vivo* extrapolation, and reverse dosimetry. **httk** links exposure scenarios, including constant oral infusion, a single dose, or multiple discrete doses, to predicted tissue and plasma concentrations. Standard toxicokinetic statistics including peak concentration and time-integrated plasma concentration (area under the curve or AUC) can be predicted, facilitating dosimetric anchoring (Wambaugh *et al.* 2013a) for comparing *in vivo* toxicity studies where toxicokinetic data were not collected (Wetmore *et al.* 2013). Important aspects of the steady-state behavior of the chemicals can be predicted for use in analysis of biomonitoring data, including the time to steady-state and C_{ss} (Wetmore *et al.* 2012; Wetmore 2015; Aylward and Hays 2011; Wambaugh *et al.* 2013b, 2014). Finally, as ongoing *in vitro* experiments allow parameterization of the models for additional chemicals, these new data can be easily distributed as updates to the package on the CRAN repository.

Acknowledgments

The authors appreciate editorial comments from Nicole Kleinstreuer and Kevin M. Crofton,

help with the Schmitt algorithm from Jimena Davis, help with Jarnac from James Sluka, and chemical properties and *in vivo* data provided by the Netherlands Organization for Applied Scientific Research (TNO) through Sieto Bosgra.

Disclaimer:

The views expressed in this publication are those of the authors and do not necessarily represent the views for policies of the U.S. Environmental Protection Agency. Reference to commercial products or services does not constitute endorsement.

References

Aylward LL, Hays SM (2011). "Consideration of Dosimetry in Evaluation of ToxCast Data." *Journal of Applied Toxicology*, **31**(8), 741–751.

Birnbaum L, Brown R, Bischoff K, Foran J, Blancato J, Clewell H, Dedrick R (1994). "Physiological Parameter Values for PBPK Models." *International Life Sciences Institute, Risk Science Institute, Washington, DC*.

Bois F, Maszle D (1997). "MCSim: A Monte Carlo Simulation Program." *Journal of Statistical Software*, **2**(1), 1–60. ISSN 1548-7660. doi:10.18637/jss.v002.i09. URL http://www.jstatsoft.org/index.php/jss/article/view/v002i09.

Bucher JR (2008). "Guest Editorial: NTP: New Initiatives, New Alignment." *Environ Health Perspect*, **116**(1). URL http://dx.doi.org/10.1289%2Fehp.11100.

Campbell JerryL J, Clewell R, Gentry P, Andersen M, Clewell HarveyJ I (2012). "Physiologically Based Pharmacokinetic/Toxicokinetic Modeling." In B Reisfeld, AN Mayeno (eds.), *Computational Toxicology*, volume 929 of *Methods in Molecular Biology*, pp. 439–499. Humana Press. ISBN 978-1-62703-049-6. doi:10.1007/978-1-62703-050-2_18. URL http://dx.doi.org/10.1007/978-1-62703-050-2_18.

Chiu WA, White P (2006). "Steady-State Solutions to PBPK Models and Their Applications to Risk Assessment I: Route-to-Route Extrapolation of Volatile Chemicals." *Risk analysis*, **26**(3), 769–780.

Davies B, Morris T (1993). "Physiological Parameters in Laboratory Animals and Humans." *Pharmaceutical Research*, **10**(7), 1093–1095. ISSN 0724-8741. doi:10.1023/a:1018943613122. URL http://dx.doi.org/10.1023/A:1018943613122.

Egeghy PP, Vallero DA, Cohen Hubal EA (2011). "Exposure-Based Prioritization of Chemicals for Risk Assessment." *Environmental Science & Policy*, **14**(8), 950–964. ISSN 1462-9011. doi:http://dx.doi.org/10.1016/j.envsci.2011.07.010. URL http://www.sciencedirect.com/science/article/pii/S1462901111001304.

Endo S, Escher BI, Goss KU (2011). "Capacities of Membrane Lipids to Accumulate Neutral Organic Chemicals." *Environmental Science & Technology*, **45**(14), 5912–5921.

Gardner M, Scott R (1980). "Age and Sex-Related Reference Ranges for Eight Plasma Constituents Derived from Randomly Selected Adults in a Scottish New Town." *Journal of Clinical Pathology*, **33**(4), 380–385.

Gelman A, Carlin JB, Stern HS, Dunson DB, Vehtari A, Rubin DB (2013). *Bayesian Data Analysis*. CRC press.

Gordon CJ (1993). *Temperature Regulation in Laboratory Rodents*. Cambridge University Press.

Hilal SH, Karickhoff SW, Carreira LA (1995). "A Rigorous Test for SPARC's Chemical Reactivity Models: Estimation of More Than 4300 Ionization pKas." *Quantitative Structure-Activity Relationships*, **14**(4), 348–355. ISSN 1521-3838. doi:10.1002/qsar.19950140405. URL http://dx.doi.org/10.1002/qsar.19950140405.

Houston JB, Carlile DJ (1997). "Prediction of Hepatic Clearance from Microsomes, Hepatocytes, and Liver Slices." *Drug Metab Rev*, **29**(4), 891–922. ISSN 0360-2532 (Print) 0360-2532. doi:10.3109/03602539709002237. Houston, J B Carlile, D J Comparative Study Journal Article Review United states Drug Metab Rev. 1997 Nov;29(4):891-922.

Hucka M, Finney A, Sauro HM, Bolouri H, Doyle JC, Kitano H, Arkin AP, Bornstein BJ, Bray D, Cornish-Bowden A, *et al.* (2003). "The Systems Biology Markup Language (SBML): a Medium for Representation and Exchange of Biochemical Network Models." *Bioinformatics*, **19**(4), 524–531.

Ito K, Houston JB (2004). "Comparison of the Use of Liver Models for Predicting Drug Clearance Using in Vitro Kinetic Data from Hepatic Microsomes and Isolated Hepatocytes." *Pharm Res*, **21**(5), 785–92. ISSN 0724-8741 (Print) 0724-8741. Ito, Kiyomi Houston, J Brian Comparative Study Journal Article United States Pharm Res. 2004 May;21(5):785-92.

Jamei M, Marciniak S, Feng K, Barnett A, Tucker G, Rostami-Hodjegan A (2009). "The Simcyp Population-Based ADME Simulator." *Expert Opin Drug Metab Toxicol*, **5**(2), 211–23. ISSN 1742-5255. doi:10.1517/17425250802691074.

Jones OA, Voulvoulis N, Lester JN (2002). "Aquatic Environmental Assessment of the Top 25 English Prescription Pharmaceuticals." *Water Res*, **36**(20), 5013–22. ISSN 0043-1354 (Print) 0043-1354. Jones, O A H Voulvoulis, N Lester, J N Journal Article England Water Res. 2002 Dec;36(20):5013-22.

Judson R, Richard A, Dix DJ, Houck K, Martin M, Kavlock R, Dellarco V, Henry T, Holderman T, Sayre P, Tan S, Carpenter T, Smith E (2008). "The Toxicity Data Landscape for Environmental Chemicals." *Environ Health Perspect*, **117**(5), 685–695. URL http://dx.doi.org/10.1289%2Fehp.0800168.

Judson RS, Houck KA, Kavlock RJ, Knudsen TB, Martin MT, Mortensen HM, Reif DM, Rotroff DM, Shah I, Richard AM, Dix DJ (2010). "In Vitro Screening of Environmental Chemicals for Targeted Testing Prioritization: The ToxCast Project." *Environmental Health Perspectives*, **118**(4), 485–492. URL http://dx.doi.org/10.1289%2Fehp.0901392.

Kilford PJ, Gertz M, Houston JB, Galetin A (2008). "Hepatocellular Binding of Drugs: Correction for Unbound Fraction in Hepatocyte Incubations Using Microsomal Binding or Drug Lipophilicity Data." *Drug Metabolism and Disposition*, **36**(7), 1194–1197.

Naritomi Y, Terashita S, Kagayama A, Sugiyama Y (2003). "Utility of Hepatocytes in Predicting Drug Metabolism: Comparison of Hepatic Intrinsic Clearance in Rats and Humans

in Vivo and in Vitro." *Drug Metabolism and Disposition*, **31**(5), 580–588. doi:10.1124/dmd.31.5.580. URL http://dmd.aspetjournals.org/content/31/5/580.abstract.

Obach RS (1999). "Prediction of Human Clearance of Twenty-Nine Drugs from Hepatic Microsomal Intrinsic Clearance Data: An Examination of in Vitro Half-Life Approach and Nonspecific Binding to Microsomes." *Drug Metab Dispos*, **27**(11), 1350–9. ISSN 0090-9556 (Print) 0090-9556. Obach, R S Journal Article United states Drug Metab Dispos. 1999 Nov;27(11):1350-9.

Obach RS, Lombardo F, Waters NJ (2008). "Trend Analysis of a Database of Intravenous Pharmacokinetic Parameters in Humans for 670 Drug Compounds." *Drug Metab Dispos*, **36**(7), 1385–405. ISSN 0090-9556. doi:10.1124/dmd.108.020479. 1521-009x Obach, R Scott Lombardo, Franco Waters, Nigel J Journal Article United States Drug Metab Dispos. 2008 Jul;36(7):1385-405. doi: 10.1124/dmd.108.020479. Epub 2008 Apr 21.

O'Flaherty EJ (1981). *Toxicants and Drugs: Kinetics and Dynamics*. John Wiley & Sons Inc, New York, NY.

Park YH, Lee K, Soltow QA, Strobel FH, Brigham KL, Parker RE, Wilson ME, Sutliff RL, Mansfield KG, Wachtman LM, Ziegler TR, Jones DP (2012). "High-Performance Metabolic Profiling of Plasma from Seven Mammalian Species for Simultaneous Environmental Chemical Surveillance and Bioeffect Monitoring." *Toxicology*, **295**(1-3), 47–55. ISSN 0300-483X. doi:10.1016/j.tox.2012.02.007. URL http://www.sciencedirect.com/science/article/pii/S0300483X12000492.

Pelekis D, Krewski K, Krishnan M (1997). "Physiologically Based Algebraic for Predicting Steady-State Toxicokinetics of Inhaled Vapors." *Toxicology Mechanisms and Methods*, **7**(3), 205–226.

Riley RJ, McGinnity DF, Austin RP (2005). "A Unified Model for Predicting Human Hepatic, Metabolic Clearance from in Vitro Intrinsic Clearance Data in Hepatocytes and Microsomes." *Drug Metab Dispos*, **33**(9), 1304–11. ISSN 0090-9556 (Print) 0090-9556. doi:10.1124/dmd.105.004259.

Robertshaw D, Reece W (2004). "Temperature Regulation and the Thermal Environment." *Dukes' Physiology of Domestic Animals*, **1**(12), 962–973.

Rotroff DM, Wetmore BA, Dix DJ, Ferguson SS, Clewell HJ, Houck KA, Lecluyse EL, Andersen ME, Judson RS, Smith CM, Sochaski MA, Kavlock RJ, Boellmann F, Martin MT, Reif DM, Wambaugh JF, Thomas RS (2010). "Incorporating Human Dosimetry and Exposure into High-Throughput in Vitro Toxicity Screening." *Toxicol Sci*, **117**(2), 348–58. ISSN 1096-0929 (Electronic) 1096-0929 (Linking). doi:10.1093/toxsci/kfq220. URL http://www.ncbi.nlm.nih.gov/pubmed/20639261.

Rovida C, Hartung T (2009). "Re-evaluation of Animal Numbers and Costs for in Vivo Tests to Accomplish REACH Legislation Requirements for Chemicals-a Report by the Transatlantic Think Tank for Toxicology (t (4))." *Altex*, **26**(3), 187–208.

Rowland M, Balant L, Peck C (2004). "Physiologically Based Pharmacokinetics in Drug Development and Regulatory Science: a Workshop Report (Georgetown University, Washington, DC, May 29–30, 2002)." *AAPS PharmSci*, **6**(1), 56–67.

Sauro HM, Fell D (2000). "Jarnac: a System for Interactive Metabolic Analysis." In *Animating the Cellular Map: Proceedings of the 9th International Meeting on BioThermoKinetics*, pp. 221–228. Stellenbosch University Press.

Schmitt W (2008a). "Corrigendum to: General Approach for the Calculation of Tissue to Plasma Partition Coefficients [Toxicology in Vitro 22 (2008) 457âĂŞ467]." *Toxicology in Vitro*, **22**(6), 1666 –. ISSN 0887-2333. doi:http://dx.doi.org/10.1016/j.tiv.2008.04.020. URL http://www.sciencedirect.com/science/article/pii/S0887233308001276.

Schmitt W (2008b). "General approach for the calculation of tissue to plasma partition coefficients." *Toxicology in Vitro*, **22**(2), 457 – 467. ISSN 0887-2333. doi:http://dx.doi.org/10.1016/j.tiv.2007.09.010. URL http://www.sciencedirect.com/science/article/pii/S0887233307002573.

Setzer R (2001). "The odesolve Package: Solvers for Ordinary Differential Equations." *R package version 0.1-1*.

Shibata Y, Takahashi H, Chiba M, Ishii Y (2002). "Prediction of Hepatic Clearance and Availability by Cryopreserved Human Hepatocytes: An Application of Serum Incubation Method." *Drug Metabolism and Disposition*, **30**(8), 892–896. doi:10.1124/dmd.30.8.892. URL http://dmd.aspetjournals.org/content/30/8/892.abstract.

Soetaert K, Petzoldt T, Setzer RW (2010a). "Package deSolve: Solving Initial Value Differential Equations in R." *CRAN R Project Documentation*.

Soetaert K, Petzoldt T, Setzer RW (2010b). "Solving Differential Equations in R: Package deSolve." *Journal of Statistical Software*, **33**(9), 1–25.

Stammers AD (1926). "The Blood Count and Body Temperature in Normal Rats." *The Journal of Physiology*, **61**(3), 329–336.

Tan YM, Liao KH, Clewell Harvey J I (2006). "Reverse Dosimetry: Interpreting Trihalomethanes Biomonitoring Data Using Physiologically Based Pharmacokinetic Modeling." *J Expos Sci Environ Epidemiol*, **17**(7), 591–603. ISSN 1559-0631. URL http://dx.doi.org/10.1038/sj.jes.7500540.

Thomas RS, Lytle WE, Keefe TJ, Constan AA, Yang RSH (1996). "Incorporating Monte Carlo Simulation into Physiologically Based Pharmacokinetic Models Using Advanced Continuous Simulation Language (ACSL): A Computational Method." *Toxicological Sciences*, **31**(1), 19–28. doi:10.1093/toxsci/31.1.19. http://toxsci.oxfordjournals.org/content/31/1/19.full.pdf+html, URL http://toxsci.oxfordjournals.org/content/31/1/19.abstract.

Thomas RS, Philbert MA, Auerbach SS, Wetmore BA, Devito MJ, Cote I, Rowlands JC, Whelan MP, Hays SM, Andersen ME, *et al.* (2013). "Incorporating New Technologies into Toxicity Testing and Risk Assessment: Moving from 21st Century Vision to a Data-Driven Framework." *toxicological sciences*, p. kft178.

Tonnelier A, Coecke S, Zaldivar JM (2012). "Screening of Chemicals for Human Bioaccumulative Potential with a Physiologically Based Toxicokinetic Model." *Archives of Toxicology*, **86**(3), 393–403. ISSN 0340-5761. doi:10.1007/s00204-011-0768-0. URL http://dx.doi.org/10.1007/s00204-011-0768-0.

Trudnowski RJ, Rico RC (1974). "Specific Gravity of Blood and Plasma at 4 and 37 C." *Clinical Chemistry*, **20**(5), 615–616.

Wambaugh JF, Setzer RW, Pitruzzello AM, Liu J, Reif DM, Kleinstreuer NC, Wang NCY, Sipes N, Martin M, Das K, *et al.* (2013a). "Dosimetric Anchoring of In Vivo and In Vitro Studies for Perfluorooctanoate and Perfluorooctanesulfonate." *Toxicological Sciences*, **136**(2), 308–327.

Wambaugh JF, Setzer RW, Reif DM, Gangwal S, Mitchell-Blackwood J, Arnot JA, Joliet O, Frame A, Rabinowitz J, Knudsen TB, *et al.* (2013b). "High-Throughput Models for Exposure-Based Chemical Prioritization in the ExpoCast Project." *Environmental Science & Technology*, **47**(15), 8479–8488.

Wambaugh JF, Wang A, Dionisio KL, Frame A, Egeghy P, Judson R, Setzer RW (2014). "High throughput heuristics for prioritizing human exposure to environmental chemicals." *Environmental science & technology*, **48**(21), 12760–12767.

Wambaugh JF, Wetmore BA, Pearce R, Strope C, Goldsmith R, Sluka JP, Sedykh A, Tropsha A, Bosgra S, Shah I, *et al.* (2015). "Toxicokinetic triage for environmental chemicals." *Toxicological Sciences*, p. kfv118.

Waters NJ, Jones R, Williams G, Sohal B (2008). "Validation of a Rapid Equilibrium Dialysis Approach for the Measurement of Plasma Protein Binding." *Journal of Pharmaceutical Sciences*, **97**(10), 4586–4595. ISSN 1520-6017. doi:10.1002/jps.21317. URL http://dx.doi.org/10.1002/jps.21317.

Weast RC, Astle M (1982). "CRC Handbook of Physics and Chemistry." *CRC Press, Boca Raton, FL (1988 1989)*.

Wetmore BA (2015). "Quantitative in vitro-to-in vivo extrapolation in a high-throughput environment." *Toxicology*, **332**, 94–101.

Wetmore BA, Wambaugh JF, Ferguson SS, Li L, Clewell HJ, Judson RS, Freeman K, Bao W, Sochaski MA, Chu TM, Black MB, Healy E, Allen B, Andersen ME, Wolfinger RD, Thomas RS (2013). "Relative Impact of Incorporating Pharmacokinetics on Predicting In Vivo Hazard and Mode of Action from High-Throughput In Vitro Toxicity Assays." *Toxicological Sciences*, **132**(2), 327–346. doi:10.1093/toxsci/kft012. URL http://toxsci.oxfordjournals.org/content/132/2/327.abstract.

Wetmore BA, Wambaugh JF, Ferguson SS, Sochaski MA, Rotroff DM, Freeman K, Clewell H J r, Dix DJ, Andersen ME, Houck KA, Allen B, Judson RS, Singh R, Kavlock RJ, Richard AM, Thomas RS (2012). "Integration of Dosimetry, Exposure, and High-Throughput Screening Data in Chemical Toxicity Assessment." *Toxicol Sci*, **125**(1), 157–74. ISSN 1096-0929. doi:10.1093/toxsci/kfr254.

Wickham H (2009). *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag.

Wilkinson GR, Shand DG (1975). "Commentary: A Physiological Approach to Hepatic Drug Clearance." *Clin Pharmacol Ther*, **18**(4), 377–90. ISSN 0009-9236 (Print) 0009-9236. Wilkinson, G R Shand, D G Journal Article Research Support, U.S. Gov't, P.H.S. United states Clin Pharmacol Ther. 1975 Oct;18(4):377-90.

Yang RS, Lu Y (2007). "The Application of Physiologically Based Pharmacokinetic (PBPK) Modeling to Risk Assessment." *Risk Assessment for Environmental Health*, p. 85.

Affiliation:

John F. Wambaugh
U.S. Environmental Protection Agency
109 T.W. Alexander Dr.
Mail Code D143-02
Research Triangle Park, NC 27711
wambaugh.john@epa.gov
http://www.epa.gov/ncct/

Robert G. Pearce
U.S. Environmental Protection Agency
109 T.W. Alexander Dr.
Mail Code D143-02
Research Triangle Park, NC 27711
http://www.epa.gov/ncct/
http://orcid.org/0000-0003-3168-4049

R. Woodrow Setzer
U.S. Environmental Protection Agency
109 T.W. Alexander Dr.
Mail Code D143-02
Research Triangle Park, NC 27711
http://www.epa.gov/ncct/

Nisha S. Sipes
Division of the National Toxicology Program
National Institute of Environmental Health Sciences
111 T.W. Alexander Dr., ML: K2-17
Research Triangle Park, NC 27709
http://www.niehs.nih.gov/research/atniehs/labs/bmsb/

Cory L. Strope
U.S. Environmental Protection Agency
109 T.W. Alexander Dr.
Mail Code D143-02
Research Triangle Park, NC 27711
http://www.epa.gov/ncct/

Journal of Statistical Software http://www.jstatsoft.org/
published by the American Statistical Association http://www.amstat.org/
Volume VV, Issue II *Submitted:* yyyy-mm-dd
MMMMMM YYYY *Accepted:* yyyy-mm-dd