Performance evaluation of a beta version of USEPA's SHEDS-HT chemical exposure model

Jessica L. Levasseur

A thesis

submitted in partial fulfillment of the

requirements for the degree of

Master of Science

University of Washington

2015

Committee:

John C. Kissel

Jeffry H. Shirai

Richard A. Fenske

Program Authorized to Offer Degree:

School of Public Health –

Department of Environmental and Occupational Health Sciences

University of Washington

Abstract

Performance evaluation of a beta version of USEPA's SHEDS-HT chemical exposure model

Jessica L. Levasseur

Chair of the Supervisory Committee:
Professor John C. Kissel, PhD
Department of Environmental and Occupational Health Sciences

The United States Environmental Protection Agency (USEPA) is in the process of developing a high throughput (HT) model for exposure-based prioritization to inform toxicity testing and chemical risk assessment within its ExpoCast program. This mechanistic modeling approach is adapted from the prior Stochastic Human Exposure and Dose Simulation (SHEDS) framework, which was implemented in SAS. SHEDS-HT, written in R, reduces user burden and input demands by linking chemicals to particular exposure scenarios based on consumer product categories or food groups. SHEDS-HT rapidly generates probabilistic population distribution estimates of pathway-specific exposures, overall exposures, and intake doses. These predictions are based on specific exposure scenarios, a fugacity-based indoor environmental media model, and information from human activity databases. The predictive ability of SHEDS-HT is best evaluated by comparison to available biomarker measurements or to alternative predictions from prior measurement-intensive exposure studies. A key example of the latter is the USEPA's Children's Total Exposure to Persistent Pesticides and Other Persistent Organic Pollutants (CTEPP) study, which was conducted in Ohio and North Carolina. CTEPP provides both environmental and corresponding urinary biomarker data for a relatively large sample of commonly encountered commercial chemicals. The National Health and Nutrition Examination

Survey (NHANES) also provides representative urinary biomarker data for some of the same compounds on a whole US population basis. However, biomarker data can reflect exposure to both a parent compound and its metabolite(s). Absence of environmental measurements for most biomarkers greatly reduces the number of compounds for which complete accounting of inputs and outputs can be conducted. In this evaluation of SHEDS-HT, focus is placed on 2,4-dichlorophenoxyacetic acid (2,4-D) and three parent compounds of the urinary biomarker 3,5,6-trichloro-2-pyridinol (TCPy), chlorpyrifos, chlorpyrifos methyl, and triclopyr.

Acknowledgements

Completing this thesis was a wild adventure and a total whirlwind. I would genuinely not have made it through this process without many people:

John Kissel, thank you so much for helping me finish this degree in two years after our first project fell apart. I am so fortunate to have been able to work in your lab and learn from you. Thank you for being an absolutely fantastic teacher and mentor. I am forever grateful for the knowledge I have gained from you.

Jeff Shirai, thank you for enduring the endless streams of questions with which I bombarded you. Your continued presence in the Kissel lab is an absolute gift. Thank you for literally always being there for all of us, and for helping us each navigate the sometimes challenging and always rewarding world of the Kissel Lab.

Rich Fenske, thank you for being on my committee. You helped make this thesis process (relatively) painless.

My fellow Fugacity-Competent Kisselites, especially Trevor and Julie, thank you for your endless encouragement and for keeping me laughing.

My friends in and out of DEOHS: your companionship and optimism undoubtedly facilitated the successful completion of this thesis.

Mom, Dad, and Justin: Your unwavering belief in me and your unending support were nothing short of incredible. Thank you all for never doubting that I could succeed, even when I questioned myself. Thank you for all of the long phone calls and the care packages. Thank you for also remembering the 3 hour time difference when you called on weekend mornings.

For my friends at Community Dinners,

who showed me that no effort to make the world better is too small.

And for my family: Mom, Dad, and Justin –

Thanks for sending me pictures of our cute dog when I was stressed about this thesis.

TABLE OF CONTENTS

Page

List of Figures

Page

LIST OF TABLES

Page

Introduction

Humans are exposed to a variety of consumer products that release potentially hazardous chemicals into our environment, exposing them through direct and indirect sources. The number of chemicals in household products is continually expanding, causing consumer product use of many chemicals to possibly be one of the most influential pathways of exposure.[1] Americans spend an average of 22 hours per day indoors.[2] The U.S. Environmental Protection Agency (USEPA) determined, through field studies conducted during the Total Exposure Assessment Methodology (TEAM) study in the early 1980's,[3] that the risk of indoor sources of toxic chemicals outnumbered outdoor sources by 80-100%. Furthermore, these indoor exposures were of greater intensity due to closer proximity of sources and limited opportunity for dilution indoors.[4,5] More recently, the highest chemical exposures were most often associated with consumer product use, as determined through reverse pharmacokinetic evaluations of biomarker data.[6]

In addition to the concerns over consumer product chemical exposures, the European Union's Registration, Evaluation, Authorisation, and Restriction of Chemical (REACH) regulations have introduced a need for high throughput (HT) exposure predictions for consumer products. This requires accurate estimation of multichemical signatures of exposure, uptake, and body burden.[7-9] In addition, it is also necessary to develop effective HT approaches for adequate Lifecycle Assessments (LCA) due to the ever increasing variety of chemicals in the consumer product universe. Both the ubiquity of chemicals and the rate at which new products enter the product line make it infeasible to study each individual chemical in a rapid manner. With new databases that contain information regarding the chemical composition of consumer products,[10,11] HT screening of new chemicals has become increasingly attainable. Though some

simple and HT exposure models have been attempted,[1,6,12-16] USEPA's SHEDS-HT is the latest exposure model based upon the SHEDS framework for high throughput, exposure-based prioritization of chemicals.

SHEDS-HT is the Stochastic Human Exposure and Dose Simulation Model for High Throughput assessment of exposure potential, developed by the USEPA. Past versions of the SHEDS framework, (SHEDS Multimedia, or SHEDS-MM) were developed in SAS programming language, often requiring a large number of chemical-specific inputs by the user. This resulted in relatively slow and burdensome evaluations, which is not ideal for evaluating a large number of chemicals quickly. SHEDS-HT is the first EPA SHEDS framework model that can be run with quick execution speed and relatively few inputs. SHEDS-HT allows for quantification of daily exposures across the US population and in many population subsets. With more precise exposure quantification, more accurate health assessments to these exposures can be investigated. SHEDS-HT code is written in freely-available R, a statistical computing software, though R Studio may also be used as an interface. SHEDS-HT was developed to generate quantitative exposure estimates rapidly and efficiently with minimal input information required. The ultimate purpose of SHEDS-HT is to assist the USEPA in developing the capacity for exposure-based prioritization of chemicals under the ExpoCast program. ExpoCast is an exposure forecasting research effort by the USEPA. This program aims to develop innovative, high-throughput methods for estimating chemical exposures. These better estimations can be used in conjunction with the USEPA's Toxicity Forecaster (ToxCast) to evaluate potential human health and environmental risks of chemical exposures. Through the SHEDS-HT modeling approach, chemical evaluation based on biologically relevant human exposures can better inform toxicity testing and prioritization for chemical risk assessments.[17]

SHEDS-HT produces population distributions of daily exposure levels and intake doses for a variety of chemicals found in residential environments, food, and drinking water. This predictive exposure model for chemicals utilizes chemical property, human activity patterns, and microenvironment information from a variety of databases. Chemicals are linked to particular direct exposure scenarios based on consumer product categories or food groups, determined through activity and dietary databases. Human activity information is from the USEPA's Consolidated Human Activity Database (CHAD),[18] while dietary information is pulled both from the National Health and Nutrition Examination Survey - What We Eat in America (NHANES-WWEIA)[19] study and the Pesticide Data Project (PDP).[20] NHANES-WWEIA provides estimates the types of food eaten by different population groups, while the PDP quantifies residues on food products. To accurately assess the personal care product ingredients, information is aggregated from two sources: the Consumer Product Chemical Profile Database (CPCPdb) produced by the USEPA[11] and the Household Products Database (HPDB) from the National Library of Medicine.[21] In addition, SHEDS-HT incorporates a fugacity-based indoor fate and transport model for indirect exposure quantification. These databases will be discussed further in the methods section. Currently, SHEDS-HT has been applied to over 2500 consumer product ingredients and agricultural pesticides.[17]

To evaluate how successful SHEDS-HT is at estimating exposures and doses in the non-occupational environment, output from SHEDS-HT will be compared to data from the USEPA's Children's Total Exposure to Persistent Pesticides and Other Persistent Organic Pollutants (CTEPP) study. CTEPP was conducted from 2000-2001, with the purpose of addressing the data gap in the area of chronic, low-level exposures of children in their home and school environments through multiple media. Total exposures of 257 children, aged 1.5 – 5 years, and

their primary adult caregivers to contaminants in their daily environments were assessed through biomonitoring and environmental monitoring. The study populations were in North Carolina and Ohio. Compounds of interest in CTEPP were several pesticides, phenols, polychlorinated biphenyls, polycyclic aromatic hydrocarbons, and phthalate esters.[22] SHEDS-HT prediction output and CTEPP data will also be compared to the National Health and Nutrition Examination Survey (NHANES) biomarker data, which is considered to be a benchmark of exposure levels in the United States (US).

The CTEPP study is of particular interest because it collected both biomonitoring and environmental measurements of the chemicals of interest, allowing for a coinciding comparison of environmental media measurements and a subject's exposure biomarkers, on a relatively large scale. This means that the CTEPP study provides an interesting comment on how well exposures can be estimated when only biomonitoring or environmental monitoring is conducted. If potential mismatches of predictions and measured observations occur, it may be necessary to address other potential sources of exposure that may have been missed or reassess how exposure estimates have been made. Furthermore, CTEPP provides environmental measurements of some chemical biomarkers. Typically, it is assumed that biomarker data can reflect both exposure to a parent compound and direct exposure to its biomarker.[23,24] Absence of environmental measurements for most biomarkers greatly reduces the number of compounds for which complete accounting of inputs and outputs may be conducted. As measuring both biomarkers and environmental media can be difficult and expensive, successful implementation of a high-throughput exposure estimation system, such as SHEDS-HT, can be a very valuable tool in estimating the daily environmental exposures experienced by a population. This cost-effective

and timely utility could be instrumental in policy decisions if validated for pathway-specific exposures and total exposures to chemicals encountered in non-occupational settings.

Beta testing of SHEDS-HT in the current study was conducted by comparing compounds investigated in the CTEPP study for which there is adequate information to perform a chemical mass balance and which are also currently accessible for estimation through SHEDS-HT. This total exposure should be confirmed by biomonitoring results. Therefore, the current study will examine SHEDS-HT exposure predictions to chlorpyrifos and its biomarker, 3,5,6-trichloro-2-pyridinol (TCPy), along with 2,4-Dichlorophenoxyacetic acid (2,4-D) in children and adults. In addition, exposure to chlorpyrifos methyl and trichlopyr will be estimated using SHEDS-HT, as these chemicals are also metabolized to TCPy.[25]

In order to examine the ability of SHEDS-HT to meet its purpose and long-term utility goals, SHEDS-HT will be examined through the null hypothesis that SHEDS-HT estimates will not be different than the CTEPP investigation output. The following specific research aims of this study will be used to test this null hypothesis:

1. Generate pathway-specific (dermal, ingestion, inhalation) and total exposure outputs from SHEDS-HT for chlorpyrifos, chlorpyrifos methyl, trichlopyr, and 2,4-D; compare SHEDS-HT estimates of the median and 95th percentile of total absorbed dose in children to those measured by CTEPP.
2. Compare total absorbed doses from SHEDS-HT and both CTEPP predictions and CTEPP biomonitoring results to corresponding National Health and Nutrition Examination Survey (NHANES) values.
3. Identify and characterize any differences between SHEDS-HT exposure estimates and CTEPP observed exposures.

Methods

SHEDS-HT currently parameterizes chemicals in a high-throughput manner by pulling together data from multiple activity databases and quantitative structure-activity relationship (QSAR)-based tools such as the USEPA's Estimation Program Interface (EPI) Suite. The constant emission scenario (discussed briefly later as an indirect exposure scenario defined by a source of constant emission) is still under development, while the down-the-drain scenario can only be used with user input values. These two scenarios will not be investigated in this study.

SHEDS-HT incorporates a reduced version of a complex indoor fugacity model, defined by Bennett and Furtaw in 2004,[26] as a source-to-concentration module for predicting indoor environmental concentrations through near- and far-field scenarios. Near-field scenarios are defined by SHEDS-HT as exposures that result from the use of mostly indoor consumer products. This near-field classification includes direct sources, such as those exposures that release immediately on or near the body (including direct dermal, direct inhalation, and direct or incidental ingestion), and indirect sources (including those exposures that occur when releases are within the residential microenvironment) are included and combined with dietary ingestion through food and drinking water pathways to produce aggregate exposures. The model combines data from multiple databases to parameterize the probabilistic exposure model as seen in Figure 1.

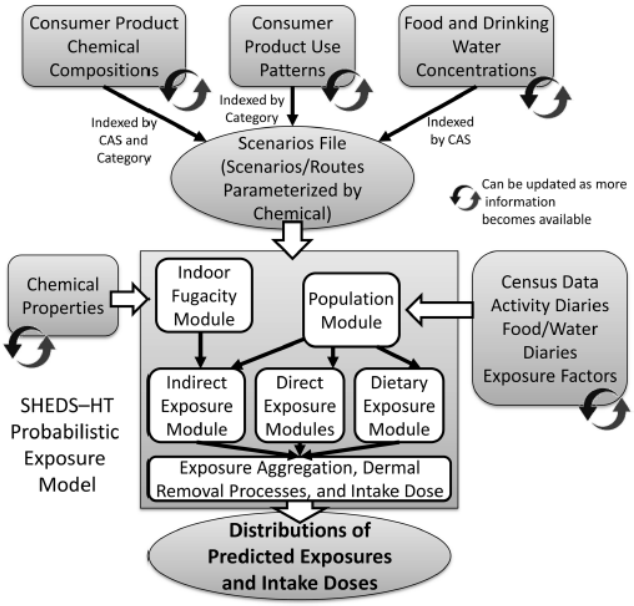

Figure 1. Schematic of SHEDS-HT input data streams and modules involved in exposure estimation[17]

Near field pathways are parameterized through consumer product category associated with the chemical of interest in the input database. Each chemical and its category pair(s) has/have assigned chemical composition information, and each category has assigned usage patterns. Coupled with dietary exposures and the indoor fugacity model for parameterizing fate and transport, an active exposure scenario for each chemical can be predicted.

SHEDS-HT output is comprised of multiple modules, including those for population, indoor fugacity, indirect exposure, direct exposure, dietary exposure, and exposure aggregation. The population module uses US Census-based input data to generate a model population representative of the entire US. Each person simulated is then assigned relevant exposure factors and diaries related to their age and gender cohort. SHEDS-HT default activity diaries summarize minute-level population activity diaries from CHAD. Food diaries are based on NHANES-WWEIA 1999-2006, calculating the mass of food consumed from 41 crop groups (Table 1) defined by the USEPA to set pesticide tolerances. The indoor fugacity model, based on the

model initially presented by Bennett and Furtaw for pesticides,[26] has been expanded to account for other chemicals with indoor sources.[27,28] The indirect exposure model predicts exposures from breathing indoor air or touching contaminated surfaces, and utilizes predictions from the indoor fugacity model (see Figure 1), calculating media concentrations in residences as a function of time based on consumer product use patterns. Chemicals found on surfaces, such as bulk chemical or chemicals found in dust are accounted for in this indirect exposure model, as are inhalation, dermal, hand-to-mouth, and object-to-mouth exposure pathways. The direct exposure model probabilistically predicts exposure for all exposure routes (inhalation, ingestion, dermal) based on category-specific use frequencies, population prevalences, masses, and compositions. Dietary exposures are determined through calculations of the total daily mass of chemical intake for each simulated person per each relevant food group. The sum of each of these exposure models (population, indoor fugacity, indirect exposure, direct exposure, and dietary exposure) is aggregated per chemical across scenarios and pathways, taking into account dermal removal processes to determine the final intake dose. Absorption fractions for inhalation and ingestion are currently the same for all chemicals: fractional absorption from inhalation is assumed to be 0.16, while fractional absorption from direct ingestion is assumed to follow a uniform distribution (0.4, 0.8).[17]

SHEDS-HT: How to Run and How It Works

To run SHEDS-HT, a separate folder should be designated for downloading and installation of the major files necessary. R or R Studio should also be loaded before attempting to open any of the following files. The working directory for SHEDS-HT should have three subfolders: “inputs”, “output”, and “R”. The five source code files (SHEDS_HT.R, ReadData.R, Fugacity.R,

Utility.R, and PostProcess.R) should be copied in to the "R" folder. Input files, all with a .csv extension, should be saved in the "inputs" folder. There are ten total input files.

File contents were not edited unless otherwise noted, though specific scenarios of interest can be requested by utilizing the PostProcess.R program. Edited or otherwise altered input files will result in missing data. As an exception, both the "Scenario" and "Run" files can be changed to reflect specifics chosen through PostProcess.R. All of the input files provided with SHEDS-HT can be appropriately used as defaults for SHEDS-HT runs. The two files most likely to be changed are the Scenarios file and the Run file.

Input files

- The Activities Diaries file summarizes activity diaries from CHAD. Each entry has one 24-hour person-day of information, including indicator variables for age, gender, season, and weekday. Minutes spent in each of the SHEDS-HT microenvironments, along with the mean metabolic equivalent (METS, a multiple of basal metabolic rate) value in each microenvironment, is provided based on the CHAD activity codes. This file includes time spent in the bath or shower, if available.
- The Chemical Properties file contains chemical properties for over 10,000 chemicals, all derived from EPI-Suite.[29]
- The Diet Diaries file summarizes food consumption. Data are drawn from NHANES-WWEIA. The mass of food consumed for each person involved in NHANES-WWEIA is determined for each of the 41 default food groups defined in SHEDS-HT (Table 1).

Table 1. SHEDS-HT default food (crop) groups[17]

Crop Group Label in SHEDS	Crop Group Description
AS	Asparagus
BA	Banana
BF	Beef
BR	Brassica (Cole) Leafy Vegetables
BS	Berry and Small Fuit
BV	Bulb Vegetables
CF	Citrus Fruits
CG	Cereal Grains
CH	Chicken
CV	Cucurbit Vegetables
DP	Dairy Products
DW	Drinking Water (Direct, Indirect)
EF	Edible Fungi
EG	Egg
FI	Fish
FV	Fruiting Vegetables (except cucurbits)
GO	Goat
HM	Human milk
HO	Honey
HS	Herbs and Spices
LE	Leafy Vegetables (except Brassica vegetables)
LR	Leaves of Root and Tuber Vegetables
LV	Legume Vegetables (Succulent or dried)
MA	Mango
MG	Meat, game
NA	Nongrass Animal Feeds (Forage, Fodder, Straw, and Hay)*
OS	Oilseeds
OT	Other
PA	Papaya
OE	Peanut
PF	Pome Fruits
PI	Pineapple
PO	Poultry, Other
PR	Pork
RA	Rabbit
RS	Raisins
RT	Root and Tuber Vegetables
SF	Stone Fruit
SH	Sheep
TN	Tree Nuts
TU	Turkey

* Alfalfa is the only food in this group

- The Exposure Factors file contains default exposure factor distributions defined by SHEDS-HT. The sources for these data and the variables used can be found in the main Exposure SAP White Paper.[17]
- The Fugacity file contains default input parameters for the SHEDS-HT fugacity model (Table 2).

Table 2. Final input variables for the reduced indoor fugacity module of SHEDS-HT and default values utilized in case study provided[17]

Variable Name	Description	Units	Default Value in SHEDS-HT†
AER_o	Air exchange rate for rooms with outdoors	1/d	Lognormal(11.9, 1.7); Assumed
MA_{sf}	Chemical mass applied to surfaces at t=0	g/m^2	Determined from category-specific use patterns
MA_a	Chemical mass applied to air at t=0	$\mu g/m^3$	Assumed to be 0 for indirect pathways
H	Height of walls	m	Uniform(2.44,3); Assumed
Af	Total floor area of the house	m^2	Lognormal(130, 1.8); Assumed
C_{out}	Chemical concentration in outdoor air	$\mu g/m^3$	Assumed to be 0 as default
$C_{I,a}$	Prior chemical concentration in air	$\mu g/m^3$	Assumed to be 0 as default
$C_{I,sf}$	Prior chemical concentration on surfaces	$\mu g/cm^2$	Assumed to be 0 as default
CS_a	chemical emission rate (source strength) in air	$\mu g/m^3$/day	User-supplied; not used in current case study
CS_{sf}	Chemical emission rate (source strength) in surfaces	$\mu g/m^2$/day	User-supplied; not used in current case study
D_a	Chemical degradation rate constant in air	1/d	Chemical-specific; estimated using EPI Suite
D_{sf}	Chemical degradation rate constant on surfaces	1/d	Chemical-specific; estimated using EPI Suite
DC_a	Diffusion coefficient in air	m^2/d	Chemical-specific; estimated using EPI Suite
K_{ow}	Octanol-water partition coefficient	-	Chemical-specific; estimated using EPI Suite
VP	Vapor pressure	Pa	Chemical-specific; estimated using EPI Suite
Sol	Solubility	mol/m^3	Chemical-specific; estimated using EPI Suite
MW	Molecular weight	g/mol	Chemical-specific; estimated using EPI Suite
CF_s	Organic carbon fraction for small particles	-	Normal (0.3,0.03)*
CF_l	Organic carbon fraction for large particles	-	Normal (0.15,0.01)*
$CL_{a,s}$	Cleaning removal rate for small particles in air	1/d	Uniform (0.018,0.22)*
$CL_{a,l}$	Cleaning removal rate for large particles in air	1/d	Uniform (0.03, 0.5)*
$CL_{sf,s}$	Cleaning removal rate for small particles on surfaces	1/d	Uniform (0.035, 0.045)*
$CL_{sf,l}$	Cleaning removal rate for large particles on surfaces	1/d	Uniform (0.035, 0.045)*
$PD_{a,s}$	Air-to-floor small particle deposition rate	m/d	Normal (11,1)*
$PD_{a,l}$	Air-to-floor large particle deposition rate	m/d	Normal (387,20)*
T	Indoor temperature	K	Normal (296,2) ; Assumed
TH_b	Boundary layer thickness over surfaces	m	Uniform (0.025,0.0275); Assumed
TH_{sf}	Effective thickness of surfaces	m	Normal (0.0098,0.002); Assumed
$R_{sf,s}$	Surface-to-air small particle resuspension rate	1/d	Uniform (0.00072, 0.00082)*
$R_{sf,l}$	Surface-to-air large particle resuspension rate	1/d	Uniform (0.0015, 0.0017)*
$L_{a,s}$	Loading of small particles in air	$\mu g/m^3$	Uniform (15, 25)*
$L_{a,l}$	Loading of large particles in air	$\mu g/m^3$	Uniform (2.2, 2.5)*
$L_{sf,s}$	Loading of small particles on surfaces	$\mu g/cm^2$	Uniform (6, 14.5)*
$L_{sf,l}$	Loading of large particles on surfaces	$\mu g/cm^2$	Uniform (11.5, 28)*

†Normal distributions reported as (mean, SD); lognormal as (GM, GSD); Uniform as (min, max)
* Mean estimated from values reported in Bennett and Furtaw (2004) for 3 smaller particle sizes, variability assumed

- The Media file contains contact probability for each media depending on the microenvironment (which must be compatible with those listed in the Activity Diaries file). Media may be defined as air, surfaces, or pets, for example. Contact probability is defined as the fraction of time spent in that microenvironment when one is also in contact with the corresponding media. These values are all between zero and one.
- The Physiological Parameters file contains regression data for physiological information, such as body weight, height, age, and metabolic rate as a function of age and gender. This was developed for SHEDS-MM from NHANES.[30]
- The Population file contains counts of males and females aged 0 to 99 as found in the 2000 US Census. These data generate probabilities for each combination of age and gender possible. The user may alter this file only when a distinctive population subgroup is being investigated.[31]
- The Run file lists the parameters for a given SHEDS-HT run. This contains 12 settings, further described below, and the other 9 input file names. Any of these settings can be altered, but not deleted completely. These include the number of person-days to simulate, the minimum and maximum age (0 – 99), gender, season, and other variables specific to ensuring SHEDS-HT code is read correctly and output is saved, such as file and folder names for input and output.
- The Scenarios file is the main SHEDS-HT input file. It generates the aggregate exposure estimates. Included in this file are the exposure scenarios to be modeled, organized by chemical, and the input parameters required to parameterize the corresponding exposure equations. Eight major exposure scenarios are included: Arts and Crafts, Auto Products, Home Maintenance, Inside the Home, Landscape/Yard, Personal Care, Pesticides, and

Pet Care. This information is derived from consumer product use database and consumer product ingredient databases, CPCPdb and HPDB.[10,11] The scenarios file used in this report investigation was a 'case study' file with information on 2507 chemicals from the 10,000 defined in the chemical properties file.

The Scenarios file, as described above, contains the definitions for parameter distributions for each available scenario for each chemical in a run. In certain scenarios, information may be pulled from the Chemical Properties file. Also included is information on the units, mean probability distribution and CV for the parameter, minimum and maximum age to which to apply the distribution (this may also vary by age), and gender to which to apply the distribution. This file defines 'scenario' multiple ways: it may include an exposure scenario, defined as one of eight major pathways of exposure; or it may refer to the category of exposure, organizing exposure scenarios based on activity groupings. Equations that are utilized in each of these pathways of exposure can be found in the Supplemental Information of Isaacs et al., 2014.[17] The eight pathways of exposure within the Scenarios File are:

- Direct Dermal
- Direct Inhalation of Vapor
- Direct Inhalation of Aerosol Mass
- Direct Incidental Ingestion
- Indirect Exposure: Fugacity
    - Calls the fugacity-based indoor fate and transport model. Non chemical-specific parameters for the fugacity module are included in the fugacity input file. Defined as post-application indirect exposure to a bolus

application of chemical in a residence; exposure defined as inhaling contaminated air or touching contaminated surfaces.

- Indirect Exposure: Constant Emission
    - Parameterized by steady state equilibrium concentration.
- Dietary
    - Ingestion of residues through food and drinking water consumption. If a simulated person is determined to have not received contaminated food, their exposure for that food group is Limit of Detection (LOD)/2 (LOD is specific to food group). Chemicals or food groups for which there is no chemical concentration available have been assigned a daily dietary exposure of 0.
- Down the Drain
    - Scenario describing chemical going down the drain in a house based on use. Not included in this assessment.

Using the Run and Scenario files, the other input files listed above are called upon to make exposure and dose predictions for chemicals of interest. SHEDS-HT utilizes these input files to assign pathway-specific exposure information and subsequent dose estimates using the indoor fugacity model, dietary assumptions, product use assumptions, population models, and indirect exposure model. Each output file contains the results for a separate chemical, labeled by CAS number. Results are presented as population percentage distributions of exposure and subsequent absorption, as well as total aggregate dose, by cohort: Total population, Males, Females, Females of Reproductive age (16-49), 0-5 year olds, 6-11 year olds, 12-19 year olds, 20-65 year olds, and 66+ year olds.

SHEDS-HT Runs Performed in the Current Study

Run and Scenario files were edited to only run chlorpyrifos, chlorpyrifos methyl, trichlopyr, and 2,4-D for all scenarios for which they were coded except for 'Down-the-drain', which requires user input data. No specific product use scenario exclusions were made. No specifications were made in regards to population or temporality. Data regarding route-specific exposure values and route-specific absorption as well as overall absorbed dose were tabulated for 0-5 year olds, 6-11 year olds, and 20-65 year olds for all four compounds of interest.

SHEDS-HT output contains estimates of dose which are calculated from total exposure estimates generated by the program in $mg/kg_{bw}/day$. To adjust these values to nanograms of compound per milliliter of urine (ng/mL), a child body weight of 18.6 kg was assumed,[32] along with a urinary output assumption of 22 $mL/kg_{bw}/day$.[33-35]

TCPy, although known to be found in the environment,[36] is not a compound for which exposure can be calculated in SHEDS-HT. In the case of TCPy exposure and dose estimates, exposure and dose estimates for three of its parent compounds (chlorpyrifos, chlorpyrifos methyl, and trichlopyr) were all made in SHEDS-HT.

Probabilistic Analysis

A probabilistic, two-dimensional (2D) analysis using Oracle Crystal Ball™ release 11.1.2.3.000 (32-bit) (henceforth referred to as Crystal Ball™), was performed only on dietary consumption pathways of 2,4-D as measured by Morgan et al.[37] Chlorpyrifos and TCPy were not included in the probabilistic analyses due to uncertainties regarding the exposure ratio of chlorpyrifos to TCPy in the environment that results in overall TCPy urinary biomarker values.

Using the Crystal Ball™ software available for use in Microsoft Excel, this investigation assigned distributions to each variable in Equation 1 to stochastically predict the subsequent dose children in NC and OH receive from dietary ingestion of 2,4-D.

$$\boldsymbol{Dose}\ (\boldsymbol{ng/kg_{bw}/day}) = \left\{\left(\frac{\boldsymbol{Exposure}\left(^{\boldsymbol{ng}}/_{\boldsymbol{day}}\right)}{\boldsymbol{Body\ Weight}\ (\boldsymbol{kg_{bw}})}\right) \times \boldsymbol{Food\ Precision}\right\} \times \boldsymbol{Absorption} \times \boldsymbol{Bias\ Factor} \quad [1]$$

For successful application of the two-dimension Monte Carlo analyses, variability and uncertainty must be defined appropriately for each variable distribution. Variability is measured or known heterogeneity, while uncertainty is usually the result of lack of knowledge.[38] The uncertainty loop consists of 59 repetitions, with 401 variability repetitions nested within each uncertainty repetition. The significance of the numbers 59 and 401 in the uncertainty and variability loop, respectively, is solely for ease in identifying three curves of particular significance in our simulation: the lower tolerance limit (LTL), median, and upper tolerance limit (UTL). The UTL and LTL each represent the upper 95% confidence limits of the 95th and 5th percentiles, respectively, of the population being modeled. This means that there is a 95% confidence that the UTL does not underestimate the 95th percentile of the true population, while there is 95% confidence that the LTL does not overestimate the 5th percentile of the true population.

This probabilistic analysis investigated the differences realized through use of a constant ingestion absorption factor of 0.5, as was utilized by Morgan et al.[36] across all chemicals, and a distributed absorption factor specific to 2,4-D from the literature.[39] The absorption factor expresses the percentage of contaminant that is taken up by the gastrointestinal tract upon consumption. These two probabilistic analyses were then compared to biomarker dose calculations from the CTEPP study for children in NC and OH. In this study, data specific to

preschool aged children and adults in both NC and OH was used to perform the 2D analysis of 2,4-D dose from dietary ingestion. Dose is generally defined as mass per body weight per time (ng/kg_{bw}/day). A breakdown of parameters and assumptions of the input variables frfrom Equation 1 is shown in Table 3, delineating which of the variables are assigned to the uncertainty or variability loop. Distributions utilized are shown in Table 4.

Table 3. Uncertainty and variability assumptions used in Crystal Ball™

Uncertainty	Variability
NC weights – scale variability	NC child weight (kg)
OH weights – scale variability	OH child weight (kg)
Child Toggling by state	Exposure NC preschool (ng/day)
	Exposure OH preschool (ng/day)
Absorption rate from literature	Absorption lower limit variability
Bias Factor	
Food Precision – NC	
Food Precision - OH	

Table 4. Assumption distributions as defined in Crystal Ball™

Variable	Distribution	Values
NC Exposure (ng/day)[40]	Lognormal	AM = 271 SD = 303
OH Exposure (ng/day)[40]	Lognormal	AM = 150 SD = 759
NC Weight (kg)[40]	Normal	Mean = 17.2 SD = 4.3
OH Weight (kg)[40]	Normal	Mean = 17.7 SD = 4.0
Toggle (NC, OH)	Discrete Uniform	Min = 1 Max = 2
Absorption rate from literature[39]	Uniform	Min = 0.79 Max = 1.0
Lower absorption rate from literature[39]	Normal	Mean = 0.79 SD = 0.03
Scale variability (kg)[41]	Uniform	±0.453592 each state's mean
Bias Factor[42]	Normal	Mean = 1.12 SD = 0.18
Food Precision – NC[36]	Normal	Mean = 0.86 SD = 0.15
Food Precision – OH[36]	Normal	Mean = 0.84 SD = 0.096

As noted above, the CTEPP assumed absorption factor was 50% for all chemical contaminants, including 2,4-D. The governing equation for this investigation is shown in Equation 1. Dose modifications that are specific to this investigation include the Bias Factor and Food Precision. Both are correction factors. The Bias Factor takes into account the difference between reported diary estimates of food and energy intake and laboratory-confirmed food and energy intake values.[42] The Food Precision factor is specific to the state, NC or OH, and reflects the laboratory instrumentation accuracy in detecting 2,4-D in samples.[36] Each of these factors, like many of the other variables in the equation, are defined as distributions. The Food Precision factors also assume that solid and liquid food are taken in at approximately a 50/50 ratio. This

assumption is intended to be a starting point for further analyses as more detailed consumption ratio information could not be found.

Note that the literature absorption factor distribution is defined as uncertain, while it has a variable distribution built in to the lower-end of its distribution. This is because there is uncertainty in the true ingestion absorption rate of 2,4-D. As described in Table 4, this rate may be as high as 100%, but has been observed at 79% 96 hours after ingestion.[39] The aforementioned study reported a measured absorption factor mean and standard deviation from six humans, and thus a variable distribution can be defined on the lower end of this ingestion rate range. The full or complete absorption of 100% is generally accepted as implied but has not yet been proved, hence rendering this range of absorption uncertain overall, with variability built-in. A correlation of 50% was assumed for body weight and exposure, specific to measurements from each state. This is because it is likely that those children that weigh more also eat more, and thus are exposed to 2,4-D residuals at a higher amount than those children who eat less.

As dietary ingestion is demonstrated to dominate the 2,4-D exposure pathways based on the CTEPP data,[37] using a better absorption factor may address some of the shortfall reported between environmental estimates and biomarker measurements in the CTEPP data. Probabilistic analyses of dietary ingestion dose using two different ingestion absorption factors, utilizing distributions on variables of interest rather than point estimates from aggregated data, should also produce estimates of 2,4-D dietary ingestion exposures and subsequent dose predictions for children in NC and OH that are more realistic to what was observed in biomarker data. In addition, probabilistic estimates and biomarker results from Morgan et al.[37] will both be compared to SHEDS-HT ingestion dose predictions for 0-5 year olds. However, the SHEDS-HT ingestion dose is not broken down by dietary or incidental ingestion pathways. Since this is the

only available data from SHEDS-HT predictions for comparison to dietary-specific ingestion exposure and dose values from Morgan et al,[37] SHEDS-HT overall ingestion dose predictions will be used as a surrogate for dietary ingestion alone. In addition, the SHEDS-HT predictions, the biomonitoring-based results from CTEPP, and the results of the two probabilistic analyses (each using the different ingestion absorption factors for 2,4-D described earlier) will be compared to NHANES data on 2,4-D population exposure for children 6-11 years old, which is the youngest age group for which data is available in NHANES.[43]

Results

Urinary data for 2,4-D and TCPy at the 50th and 95th population percentiles from NHANES and CTEPP are tabulated below in Table 5. Data from NHANES and the CTEPP study are measured values of urine concentrations that do not account for creatinine adjustments. Values from the CTEPP study combine home and daycare child data for ease of comparison in this study. NHANES data are available for different age groups, but data shown in Table 5 are only for those age groups comparable with the CTEPP study's age groups. Child levels of urinary 2,4-D are higher than those concentrations of 2,4-D in adults. This general trend also holds true in SHEDS-HT exposure predictions. Adult ages for each state are averages of the corresponding adult caregivers for the children involved in the CTEPP study. TCPy urinary measurements were not collected for adults in either state.

Table 5 shows estimates of TCPy output from its corresponding parent compounds, as produced by SHEDS-HT. However, the prediction totals do not account for the potential contribution of exposures to TCPy found in the environment. As such, the SHEDS-HT predictions could be underestimates which is of note as the current SHEDS-HT prediction for 0-5 year olds already overestimate the child urinary levels seen in NHANES and CTEPP. With respect to the TCPy parent compounds considered and showcased in Table 5, chlorpyrifos clearly dominates. In light of this, all additional SHEDS-HT analyses involving TCPy parent compounds focused solely on chlorpyrifos.

Table 5. Comparison of measured TCPy and 2,4-D urinary levels from NHANES and CTEPP, and predicted TCPy estimates based on metabolism from parent compounds from SHEDS-HT (ng/mL)

		50th percentile		95th percentile	
	Age	TCPy	2,4-D	TCPy	2,4-D
NHANES[44]					
	6-11	3.1	0.29	15.3	1.9
	20-59	--	0.22	--	1.6
CTEPP[40]					
NC	1.5 - 5	5.3	0.7	15.5	2.0
OH	1.5 - 5	5.1	1.0	12.3	3.6
NC	32.8	--	0.6	--	3.0
OH	33.1	--	0.7	--	2.8
SHEDS-HT					
Total	0-5	13.3	4.5	548	154
	6-11	5.1	2.0	240	65
	20-65	--	0.9	--	22
Chlorpyrifos	0-5	13.1	--	541	--
	6-11	4.9	--	233	--
Chlorpyrifos methyl	0-5	0.2	--	7.2	--
	6-11	0.2	--	6.3	--
Trichlopyr	0-5	4.8E-03	--	0.09	--
	6-11	2.8E-03	--	0.05	--

After noting the differences in urinary exposure measurements at the 50th population percentile, a breakdown of dose estimates by pathway were made, as seen in Table 8. The NHANES value of 50th population percentile aggregate dose for 2,4-D was made by converting NHANES urinary output values to $ng/kg_{bw}/day$ using the CTEPP assumed urinary output of 22 $mL/kg_{bw}/day$,[36] as discussed earlier. This conversion was not conducted for TCPy to chlorpyrifos, however, because it is uncertain how much TCPy is due to specifically chlorpyrifos exposure and not environmentally-present TCPy. Figure 2 from Morgan et al.[45] demonstrates that environmental TCPy may be a major player in the urinary excretion levels of TCPy. Dissimilar to chlorpyrifos, 2,4-D is predominantly excreted as 2,4-D from a single parent compound, and thus fewer complexities and uncertainties exist in interpreting the biomarker-based data from the CTEPP study and NHANES.

NHANES does not provide estimates of pathway-specific exposure metrics, and thus dose could not be estimated on a pathway-specific basis. The CTEPP study provides environmental measurements of 2,4-D that do permit estimation of pathway-specific and aggregate dose estimates. Total ingestion dose was calculated through summing indirect ingestion and dietary ingestion. Total ingestion dose provided an estimate that could be better compared to SHEDS-HT ingestion-specific dose estimates, as SHEDS-HT ingestion estimates do not provide a breakdown of how much ingestion dose comes specifically from dietary or indirect sources. In both SHEDS-HT and the CTEPP study, ingestion exposure was demonstrated to dominate the exposure pathways, specifically through dietary sources as calculated by Morgan et al.[40]

Table 6 shows 2,4-D doses per age group and exposure pathway. Biomarker-based estimates of dose are back-calculations from urinary biomarker concentrations, using an assumed urinary output of 22 mL/kg_{bw}/day. CTEPP environmental estimates of exposure and subsequent dose did not calculate a total ingestion, but indirect and dietary ingestion estimates were summed to approximate total ingestion at the appropriate percentile. SHEDS-HT predictions of aggregate dose are reported as mass per body weight per day. However, SHEDS-HT output for individual exposure routes required dividing by the appropriate body weight to estimate daily dose attributable. Body weights of 13.8, 18.6, and 80 kg were assumed for 0-5, 6-11, and 20-65 year old age groups, respectively.

Comparing median dose estimates in Table 6, total ingestion appears to be the dominant route of exposure compared to the other routes estimated in CTEPP and SHEDS-HT. It's important to note that CTEPP and SHEDS-HT are not comprehensive in the exposure pathways considered (e.g., contact of surfaces or objects, dermal), and questions remain about the assumptions made for certain pathways (e.g., dermal). Estimations of dose for children and

adults in the CTEPP study are numerically less than those age-specific calculations of dose from NHANES, comparing each at the 50th percentile. Both CTEPP estimates and NHANES biomarker values vary by age group at the median, with children having numerically higher doses than adults. Median values for adult and child 2,4-D doses are similar, according to NHANES data, though CTEPP estimates put child doses as four times higher than those of adults. SHEDS-HT dose estimates varied accordingly by age, with younger age corresponding to higher dose. The SHEDS-HT estimates for 0-5 year olds and 6-11 year olds were numerically higher than either CTEPP or NHANES data, though the SHEDS-HT estimates for 20-65 year olds was similar to the doses observed through 2,4-D urinary biomarker-based estimates from the CTEPP study from the adult age group in both NC and OH. Dose data from the 95th population percentile were also examined for 2,4-D biomarker levels across NHANES, CTEPP biomarker-based and environmental-based estimates, and SHEDS-HT output predictions. Trends observed at the 95th percentile were generally similar to those described for the 50th percentile data (Table 6).

Table 6. Comparison of 2,4-D percentile dose estimates (ng/kg$_{bw}$/day) for children and adults.

		Dermal		Ingestion (diet)		Ingestion (indirect)		Ingestion (total)		Inhalation		Aggregate dose	
	Age	*50th*	*95th*	*50th*	*95th*	*50th*	*95th*	*50th*	*95th*	*50th*	*95th*	*50th*	*95th*
NHANES biomarker-based estimate[41]													
	6-11	--	--	--	--	--	--	--	--	--	--	*6.4*	*41*
	20-59	--	--	--	--	--	--	--	--	--	--	*4.8*	*34*
CTEPP environmental data-based estimates[39]													
NC	1.5 - 5	--	--	4.8	24	0.04	1.4	*4.9*	*26*	1.0	0.6	4.9	23
OH	1.5 - 5	--	--	3.6	24	0.2	1.8	*3.7*	*26*	0.05	0.4	4.1	40
NC	32.8	--	--	1.0	7.1	0.01	0.2	*0.98*	*7.3*	0.02	0.2	1.1	6.9
OH	33.1	--	--	<MDL	8.1	0.02	0.4	*0.02*	*8.5*	0.02	0.1	1.0	8.4
CTEPP biomarker-based estimates[39]													
NC	1.5 - 5	--	--	--	--	--	--	--	--	--	--	*15*	*43*
OH	1.5 - 5	--	--	--	--	--	--	--	--	--	--	*22*	*79*
NC	32.8	--	--	--	--	--	--	--	--	--	--	*13*	*66*
OH	33.1	--	--	--	--	--	--	--	--	--	--	*15*	*62*
SHEDS-HT predictions													
	0-5	2.7	*105*	--	--	--	--	*68*	*3108*	*0.08*	*1.6*	98	3378
	6-11	1.9	*80*	--	--	--	--	*58*	*1544*	*0.05*	*0.65*	44	1438
	20-65	2.1	*92*	--	--	--	--	*68*	*391*	*0.06*	*0.28*	20	489

Italicized numbers represent values calculated in this study

Probabilistic Analysis Results

Absorbed 2,4-D dose population distribution curves from dietary ingestion using a distributed (Table 4) and a constant (50%) GI absorption factor are seen in Figure 2. Using a distributed GI absorption factor for 2,4-D from dietary sources increased the absorbed dose from dietary ingestion. These two probabilistic curves are plotted against the overall dose values from NC and OH that were calculated by dividing the measured urinary concentrations of 2,4-D (ng/mL) by 22 mL/kg$_{bw}$/day, a constant used during the CTEPP investigation to address urinary output across bodyweights. However, it is important to keep in mind that these state-specific curves reflect total 2,4-D exposure and not only exposure through the dietary ingestion route.

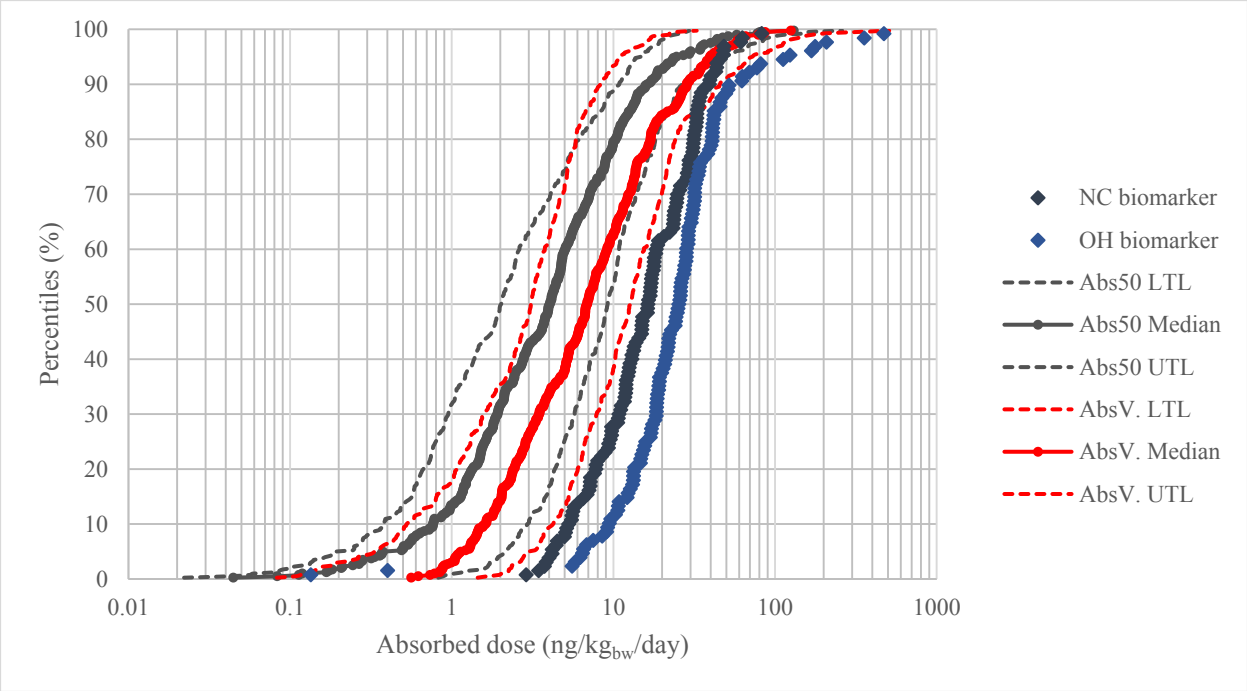

Figure 2. Distributed absorbed dose (AbsV.) and constant absorbed dose (Abs50) from 2,4-D dietary exposure plotted with overall, biomarker-derived dose of 2,4-D in NC and OH children

Selected values shown graphically in Figure 2 are summarized in Table 7. UTL values at each tabulated percentile for both Crystal Ball™ output curves fall short of the respective overall biomarker measurement percentile values, except in the case of the OH exposures at the fifth percentile.

Table 7. Probabilistic estimates of 2,4-D ingestion absorbed dose using a constant 50% ingestion absorption factor and an absorption distribution compared to overall 2,4-D dose estimates from CTEPP (ng/kg$_{bw}$/day)

	Percentile:	**5th**	**50th**	**95th**
Crystal Ball Output				
50% Absorption	Median	0.1	4.1	25
	UTL	0.9	9.2	44
Absorption Distribution	Median	0.6	6.9	40
	UTL	1.7	13	73
Overall Biomarker Measurements				
North Carolina		2.9	16	48
Ohio		0.1	25	124

The Crystal Ball™ output plots from CTEPP environmentally-based data using variable and constant ingestion absorption factors for 2,4-D are plotted against the predicted SHEDS-HT ingestion doses and the NHANES dose measurements in Figure 3. The SHEDS-HT population distribution curve does not delineate between dietary ingestion and indirect ingestion, as mentioned earlier. This should be kept in mind when it is compared to the previously discussed NC and OH biomarker data, and Crystal Ball™ output curves utilizing the two absorption assumptions, each of which represent total exposure from all pathways. At the 50th percentile, the SHEDS-HT prediction is more than an order of magnitude higher than either of the median curves produced from the two different absorption assumptions. This overestimation drops slightly when comparing the two UTLs, but there is still approximately a fivefold difference. NHANES biomarker measurement-based values for total 2,4-D exposure for 6-11 year olds have been converted to total 2,4-D dose received by 6-11 year olds (the youngest available age group for which data is provided). NC and OH total dose estimates from CTEPP biomarker-based measurements are larger than those reported by NHANES. The median curve using a distribution for ingestion absorption seems to align well with the NHANES numbers, but this median curve represents only dose from dietary ingestion, not overall received dose of 2,4-D. The SHEDS-HT curve for 2,4-D dose received by ingestion only is well above any of the other curves.

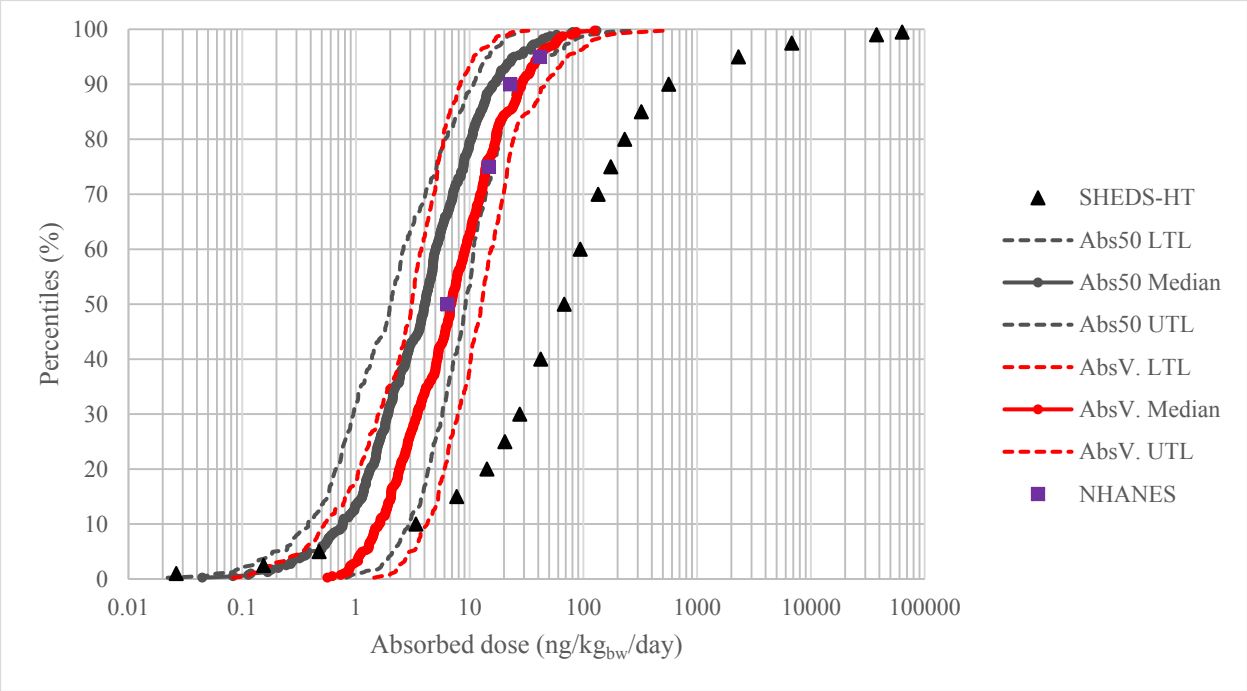

Figure 3. Constant ingestion absorption (Abs50) and distributed ingestion absorption (AbsV.) curves of dietary 2,4-D doses plotted against SHEDS-HT dose predictions and NHANES biomarker measurement-based dose of 2,4-D from all ingestion exposures

Figure 4 illustrates a similar trend when comparing the SHEDS-HT predictions to the overall dose of 2,4-D observed through measured urinary biomarker-based data from the CTEPP NC and OH child populations. Again though, it is important to remember that the state-specific biomarker data represents aggregate exposure of 2,4-D, while the SHEDS-HT prediction represents only 2,4-D dose due to the ingestion exposure pathways.

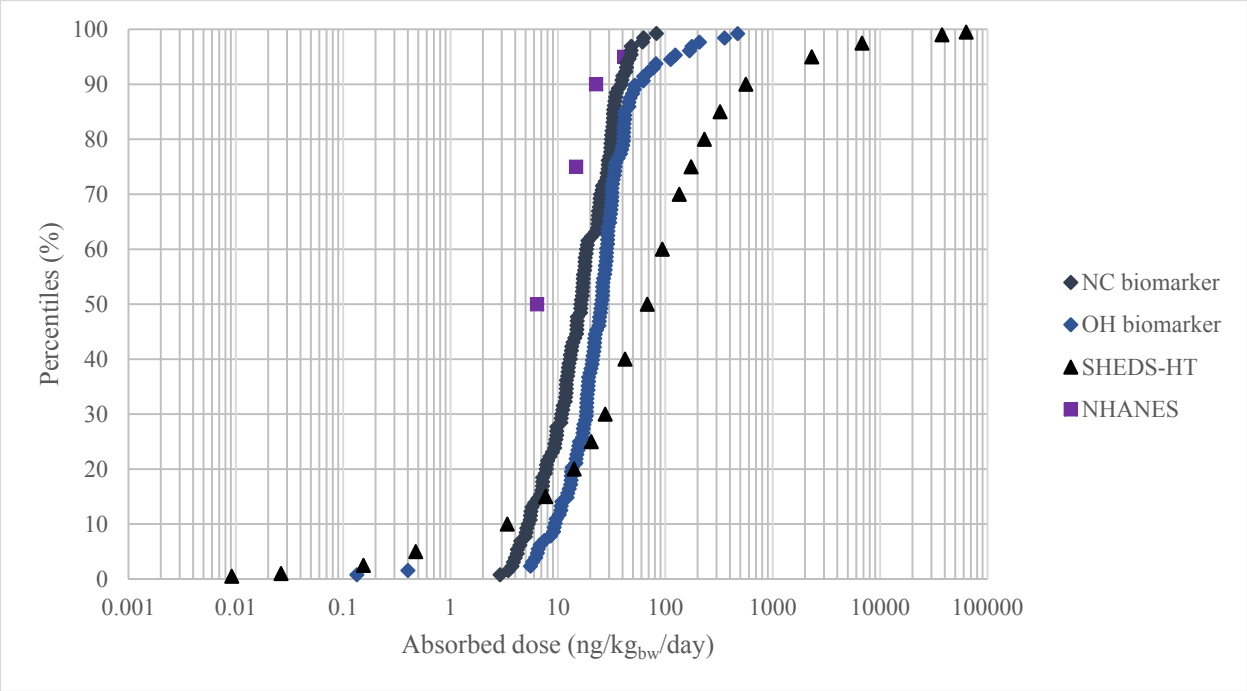

Figure 4. Biomarker measurement-based population distribution curves of overall 2,4-D dose for children in NC and OH and biomarker measurement-based NHANES data plotted against SHEDS-HT predictions of population distribution for dose of 2,4-D from all ingestion exposures

Discussion

SHEDS-HT Over-Predicts Dose as Compared to Biomarker-based Studies

SHEDS-HT predictions were found to be largely dissimilar to CTEPP and NHANES measurements for chlorpyrifos in children and 2,4-D in children and adults at both the 50th and 95th population percentiles. NHANES and CTEPP biomarker-based estimates are more similar to each other than to SHEDS-HT predictions. It remains unclear why the SHEDS-HT predictions are over-predicting the TCPy and 2,4-D seen in measured biomarker levels in children. Further investigation into the inner workings of SHEDS-HT and the assumptions and equations utilized in the generation of its predictions is warranted. A better understanding of the route-specific predictions of exposure and dose is needed, as demonstrated by the SHEDS-HT output. It is possible that the generally accepted assumptions of the different exposure pathways may be incorrect and should be revisited.

This observation is particularly notable given the evidence that standard USEPA exposure protocols under predict dermal exposures. For the compounds evaluated in this study, Morgan et al.[45] observed a shortfall between measured urinary biomarker levels and estimated values based on environmental media concentrations. Figure 5 illustrates this discrepancy for the chlorpyrifos and TCPy case.

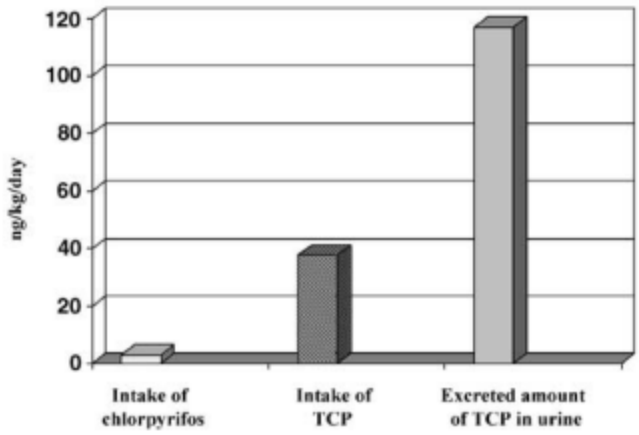

Figure 5. Median potential aggregate absorbed dose of chlorpyrifos and TCPy from the environment as compared to excreted median amounts of TCPy measured in child urine in the CTEPP study [45]

The fact that TCPy is a metabolite derivable from multiple parent compounds, in conjunction with the lack of information available about how parent compounds and TCPy in the environment contribute to observed TCPy urinary levels, makes it a difficult case study. Further, there is a limited understanding about if and how exposures to TCPy in the environment result in TCPy in urine, although two studies[23,24] suggest that this mechanism can occur in rats and sheep. Without the ability to predict TCPy exposure using SHEDS-HT, it is difficult to predict how much of a dose from environmental TCPy is received as a source of comparison for what Morgan et al.[45] measured in the environment during the CTEPP study. The same difficulties regarding parent compound to metabolite ratios are not of concern when examining 2,4-D, since the parent compound and metabolite are the same.

The shortfall observed by Morgan et al.[40] may be explained by mischaracterizations of particular exposure pathways. As stated in Morgan et al.,[45] it is necessary to both develop better methods for obtaining more accurate predictions of exposure and subsequent dose and further characterize pathways that are not yet fully understood, such as dermal and indirect inhalation. Additional tools are being developed that can aid in the process of more accurate exposure and

subsequent dose predictions. The use of these tools can better inform predictions of exposure, and can be used in models such as SHEDS-HT for increased affordability and rapid screening of chemicals.

Crystal Ball™ was used to further investigate if the shortfall observed by Morgan et al. for 2,4-D[37] could be explained by treating the absorption factor as a probabilistic parameter. Using a 2D analysis of different dietary absorption factors, it was demonstrated that increasing the GI absorption factor to nearly 100% does not explain the shortfall observed by biomarker-based measurements of 2,4-D. As Morgan et al.[37] demonstrated for 2,4-D, the difference between the environmentally-based predictions of dietary ingestion dose and overall dose was relatively small (4.8 ng/kg_{bw}/day and 4.9 ng/kg_{bw}/day at the median, respectively). However, this overall dose estimate was many times smaller than the observed urinary levels of 2,4-D. By increasing the ingestion absorption factor used to calculate dietary ingestion exposure and subsequent dose, it was expected that some of this difference between estimated dose and observed dose would be accounted for. However, even assuming 100% absorption for dietary ingestion of 2,4-D, only a two-fold increase in the resulting 2,4-D ingestion dose can be expected because the default assumption used in the CTEPP study for the absorption factor was 50%. Therefore, increasing the absorption factor and assigning it a distribution that allowed for higher dietary absorption could not account for the shortfall observed in the CTEPP study between dose estimates and the observed urinary concentrations of 2,4-D. Because of this, it is worth considering that exposure pathways other than dietary ingestion may contribute non-negligibly to residential aggregate exposures of 2,4-D. As noted by Morgan et al.,[45] leaving the dermal exposure pathway unaccounted for may explain why some of the shortfall between metabolite urinary

concentrations and environmentally-based estimates of exposure. More investigation of non-dietary exposure pathways common to non-occupational exposures is required.

Some of the limitations inherent in this Crystal Ball™ investigation of how dietary ingestion dose influences overall dose to a chemical like 2,4-D involve uncertainty. It is very difficult to characterize uncertainty or evaluate the amount of uncertainty in a study. This is due to a variety of factors. Dietary investigations were conducted over a 48-hour period in homes or daycare facilities,[36] and therefore it is unlikely that food ingested over this period can be extrapolated as an average of each child's diet over the course of their young life. A 50% correlation between body weight and exposure for children measured in the CTEPP study is only an educated estimate, and should be confirmed through further testing. It is also likely that the assumption that the solid to liquid food ratio was not 50/50. This assumption can be investigated further using child-specific data. Ultimately, the uncertainty in this study is uncertain.

SHEDS-HT Utility and Functionality Evaluation

SHEDS-HT is user friendly. The movement from SAS to R has been beneficial to the user, as is the requirement of very limited user input. However, it appears as though SHEDS-HT trades utility for specificity. It is difficult to ascertain exactly what is occurring in SHEDS-HT calculations because internal performance is not transparent. In particular, assumptions incorporated in the indoor fugacity model are generalizations across all chemicals meant to increase the ease with which SHEDS-HT can be used. However, these are not clearly described in currently available documentation.

Some of the current limitations of SHEDS-HT involve the manner in which calculations are reported to the user. It appears that ingestion exposure only accounts for chemical exposure due to direct (dietary) ingestion, while ingestion absorption incorporates direct and indirect (non-

dietary) sources of exposure. Without a clear breakdown of these two very different types of exposure, it is impossible to assess on how each pathway contributes to the amount absorbed and where errors may be occurring in calculating the ingestion exposure and subsequent dose of a chemical. It is difficult to understand exactly what is occurring in SHEDS-HT that causes the ingestion route of exposure to dominate to such a large degree. Although this pathway is generally shown to have the largest impact on overall exposure in residential settings, it is unclear through which exposure pathway in SHEDS-HT this high exposure is occurring. Due to a lack of transparency regarding exposure per individual pathway, it is impossible to assess which internal factors are contributing to the overall high ingestion dose values in the SHEDS-HT output. Furthermore, absorption fractions of inhalation and ingestion are currently constant for all chemicals: fractional absorption from inhalation is assumed to be 0.16 (Morgan et al. utilized a 0.5 inhalation absorption fraction[36]), while fractional absorption from direct ingestion is assumed to follow a uniform distribution (0.4, 0.8).[17] This is a noted shortcoming of the SHEDS-HT model, and chemical-specific algorithms for fractional absorption per exposure pathway are currently being developed.

Another limitation involves the quantification of dermal exposures to chemicals. Chlorpyrifos methyl and trichlopyr were considered to have no dermal exposure in the case study files used in this evaluation of SHEDS-HT. This is because these two chemicals were presumed to be used only outdoors. It is likely that additional chemicals are coded in SHEDS-HT as having no indoor use and therefore no indoor presence, but it is unclear what fraction of the chemicals in the database may be affected by this assumption. Primarily outdoor use does not negate exposure possibilities, as these chemicals can have outdoor residential exposure and can be tracked into indoor environments resulting in subsequent exposures. In addition, SHEDS-HT knowingly

utilizes a fractional absorption model for dermal exposure instead of a load-dependent flux model.[46] This decision was made in an apparent effort to simplify algorithmic choices, but is acknowledged to contribute additional uncertainty.[17] Surface contact is assumed to not occur while subjects are sleeping,[17] though this is unlikely given the presence of various household chemicals that occur in fabrics, foams, and personal care products.

After observing the high predictive doses from SHEDS-HT, exposures were investigated in the SHEDS-HT output for peculiarities. SHEDS-HT reports pathway-specific exposure and absorption values, which are then summed to determine aggregate dose. SHEDS-HT reported rather high gross dermal exposures, though ultimate dermal absorption dose values are much lower for both chlorpyrifos and 2,4-D. Table 8 tabulates these irregularities for 2,4-D. SHEDS-HT literature states that dermal dose should represent 1-5% of dermal exposure, yet the predictive output from SHEDS-HT reported much lower percentages of dermal exposure absorbed as subsequent dose.

Table 8. Dermal-specific exposures and subsequent absorbed doses of 2,4-D at the 50th and 95th percentile as derived from SHEDS-HT predictions (ng/day)

		Dermal Pathway		Dose Ratio (absorbed:exposure)	
	Age	**50th**	**95th**	**50th**	**95th**
Exposure	0-5	3.3E+05	9.9E+06	1.5E-04	2.0E-04
Absorbed dose	0-5	49	1945		
Exposure	6-11	2.7E+05	8.7E+06	1.3E-04	1.7E-04
Absorbed dose	6-11	35	1481		
Exposure	20-65	2.7E+05	9.0E+06	1.4E-04	1.9E-04
Absorbed dose	20-65	38	1714		

Dermal exposure, considered to be total daily loading of a chemical, and the subsequent absorption are influenced by five assumed removal processes: bathing, brush-off, dermal absorption, hand-to-mouth transfer, and hand washing. Bathing and hand washing remove the

largest fraction of the daily exposure, on average, in SHEDS-HT. Each individual is assumed to end the day with zero dermal loading. This assumes that the population showers immediately before bedtime and has no further contact with any chemicals for the duration of the day, which is unlikely for the same reasons as those stated regarding the assumption that surface contact does not occur while sleeping. It is likely that these assumptions reflect data gaps in the literature.

SHEDS-HT also pulls from and combines multiple databases that acknowledge their inherent uncertainty in assessing the US population's activities and exposures. Combining these databases results in compounding uncertainty.[17] However, it is unlikely that the resources will ever be available to fully describe the exposure-relevant behaviors of the entire US population. As a result, this is perhaps a necessary shortfall in the SHEDS-HT process. These databases can be improved, and SHEDS-HT predictions will likely become more accurate as the uncertainties inherent in the databases decrease.

Errors Identified and Corrected

In working closely with the USEPA, some correctable errors were identified and addressed through the beta testing completed in this study:

- Increased usability was addressed through a modified Postprocess.R file, that allowed for particular scenarios and/or particular chemicals to be addressed, rather than running the database in full each time. This cut the run time from approximately 10 hours (for running all 2500+ chemicals) to approximately 10 minutes per chemical or per scenario(s).
- A calculation bug regarding mean exposure for males only was identified in the Utility.R file and corrected.

- Ingestion and inhalation calculations are now normalized by bodyweight, with an additional output column now available for increased transparency on how overall absorbed dose is calculated across percentiles.

Additional minor issues were identified and addressed accordingly. The USEPA has been eager to improve their product and has been extremely amenable to addressing situations as they have arisen.

Conclusions

This evaluation is a specific yet limited investigation into the usability of SHEDS-HT and the accuracy of its predictions for exposure-based prioritization of chemicals in the USEPA's ExpoCast program. Many variables make up the SHEDS-HT model, and this report does not begin to investigate many of the factors that can be changed in the model. For a comprehensive evaluation of exactly what is working and what needs improvement, an extensive sensitivity analysis investigating each variable would likely be necessary. This report only altered chemical type, but left all other variables at their default settings.

This investigative report is unique, however, in that it compares SHEDS-HT predictions to real world biomarker-based data for insight into the accuracy of the model predictions. SHEDS-HT will only be successful if it can predict realistic exposure scenarios, subsequently calculating realistic dose values. Without accurate dose predictions, informative assessments of human health risks cannot be made. SHEDS-HT predictions for 2,4-D and chlorpyrifos are too high to be considered realistic, and thus cannot be reliably used in exposure-based prioritization. Although SHEDS-HT predictions seem to be human-health protective at this point in model development, increased accuracy is required for utility. This utility and resulting reliability may be achieved through better understanding of exposure pathways of non-occupational relevance. In particular, proper characterization of the dermal pathway is necessary to fully understand residential doses of ubiquitous chemical exposures. Reliability may also be enhanced through increased transparency of the assumptions and equations utilized by SHEDS-HT to generate exposure and dose estimates.

However, SHEDS-HT is over predicting children's exposures to both 2,4-D and chlorpyrifos when compared to biomarker-based data from the CTEPP study and NHANES. Increased

accuracy of the dermal pathway would be expected to further increase SHEDS-HT predictions, as these are likely under predicted using the current model structure. Furthermore, using an inhalation absorption assumption fraction of 0.16 may also be low for different classes of chemicals. Since dermal exposure and possibly inhalation are under predictions of exposure, it seems that the ingestion exposure predictions in SHEDS-HT may be too high. It is unclear exactly which exposure pathway is contributing to the over prediction in SHEDS-HT output, but increased transparency of the exposures per pathway should elucidate this discrepancy.

SHEDS-HT is an easy to use exposure-based prediction model that is a promising tool for future use. However, further development is necessary before this exposure-based prioritization tool can be accurately and successfully used in regulatory decisions.

References

1. Jayjock MA, Chaisson CF, Franklin CA, Arnold S, Price PS. Using publicly available information to create exposure and risk-based ranking of chemicals used in the workplace and consumer products. *J. Exposure Sci. Environ. Epidemiol.* 2009;19:515-524.

2. Bernstein JA, Alexis N, Bacchus H, et al. The health effects of nonindustrial indoor air pollution. *Journal of Allergy and Clinical Immunologies.* 2008;121(3):585-591.

3. Wallace L. The Total Exposure Assessment Methodology (TEAM) Study. *USEPA Research and Development.* 1987.

4. Wallace LA. Comparison of risks from outdoor and indoor exposure to toxic chemicals. *Environ. Health Perspect.* 1991(95):7-13.

5. Ott WR. Total human exposure: basic concepts, EPA field studies, and future research needs. *J. Air Waste Manage. Assoc.* 1991(40):966-975.

6. Wambaugh JF, Setzer RW, Reif DM, et al. High-throughput models for exposure-based chemical prioritization in the ExpoCast project. *Environ. Sci. Technol.* 2013;47:8479-8488.

7. Fryer M, Collins CD, Ferrier H, Colvile RN, Nieuwenhuijsen MJ. Human exposure modelling for chemical risk assessment: a review of current approaches and research and policy implications. *Environ. Sci. Policy.* 2006(9):261-274.

8. Hall B, Tozer S, Safford B, et al. European consumer exposure to cosmetic products, a framework for conducting population exposure assessments. *Food Chem. Toxicol.* 2007(45):2097-2108.

9. Hall B, Steiling W, Safford B, et al. European consumer exposure to cosmetic products, a framework for conducting population exposure assessments Part 2. *Food Chem. Toxicol.* 2011(49):408-422.

10. The Household Products Database. http://householdproducts.nlm.nih.gov. Accessed August 30, 2014.

11. Goldsmith MR, Grulke CM, Brooks RD, et al. Development of a consumer product ingredient database for chemical exposure screening and prioritization. *Food Chem. Toxicol.* 2014(65):269-279.

12. Egeghy PP, Vallero DA, Cohen-Hubal EA. Exposure-based prioritization of chemicals for risk assessment. *Environ. Sci. Policy.* 2011;14:950-964.

13. Arnot JA, Brown TN, Wania F, Breivik K, McLachlan MS. Prioritizing chemicals and data requirements for screening-level exposure and risk assessement. *Environ. Health Perspect.* 2012;46:1565-1570.

14. Little JC, Weschler CJ, Nazaroff WW, Liu Z, Cohen-Hubal EA. Rapid methods to estimate potential exposure to semivolatile organic compounds in the indoor environment. *Environ. Sci. Technol.* 2012:11171-11178.

15. Vermeire TG, van der Poel P, van de Laar RT, Roelfzema H. Estimation of consumer exposure to chemicals: application of simple models. *Sci. Total Environ.* 1993(136):155-176.

16. Hauschild M, Huijbregts M, Jolliet O, et al. Building a model based on scientific consensus for life cycle impact assessment of chemicals: the search for harmony and parsimony. *Environmental Science and Technology.* 2008;42:7032-7037.

17. Isaacs KK, Glen WG, Egeghy P, et al. SHEDS-HT: An Integrated Probabilistic Exposure Model for Prioritizing Exposures to Chemicals with Near-Field and Dietary Sources *Environmental Science & Technology.* 2014;48:12750−12759.

18. USEPA. The Consolidated Human Activity Database. http://www.epa.gov/heasd/chad.html. 2014.

19. USDA. What We Eat in America, NHANES 1999-2006: http://www.cdc.gov/nchs/nhanes/wweia/htm. 2014.

20. USDA. Pesticide Data Program: http://www.ams.usda.gov/AMSv1.0/pdp. 2014.

21. National Library of Medicine. The Household Products Database Home Page. http://householdproducts.nlm.nih.gov/. 2014.

22. Wilson NK, Chuang JC, Iachan R, et al. Design and sampling methodology for a large study of preschool children's aggregate exposures to persistent organic pollutants in their everyday environments *Journal of Exposure Analysis and Environmental Epidemiology.* 2004;14:260-274.

23. Timchalk C, Busby A, Campbell J, Needham L, Barr D. Comparative pharmacokinetics of the organophosphorus insecticide chlorpyrifos and its major metabolites diethylphosphate, diethylthiophosphate and 3,5,6-trichloro-2-pyridinol in the rat. *Toxicology.* 2007;237:145-157.

24. Bakke J, Price C. Metabolism of O,O-dimethyl-O-(3,5,6-trichloro-2-pyridyl) phosphorothioate in sheep and rats and of 2,5,6-trichloro-2-pyridinol in sheep. *Journal of Environmental Science and Health, Part B: Pesticides, Food Contaminants, and Agricultural Wastes.* 1976;11(1):9-22.

25. Needham LL, Patterson DGJ, Barr DB, Grainger J, Calafat AM. Uses of speciation techniques in biomonitoring for assessing human exposure to organic environmental chemicals *Analytical and Bioanalytical Chemistry.* 2005;381:397-404.

26. Bennett DH, Furtaw EJ, Jr. Fugacity-based indoor residential pesticide fate model. *Environ. Sci. Technol.* 2004;38(7):2142-2152.

27. Shin HM, McKone TE, Bennett DH. Intake fraction for the indoor environment: a tool for prioritizing indoor chemical sources. *Environ. Sci. Technol.* 2012;46(10063-10072).

28. Zhang X, Diamon ML, Ibarra C, Harrad S. Multimedia modeling of polybrominated diphenyl ether emissions and fate indoors. *Environ. Sci. Technol.* 2009;43:2845-2850.

29. USEPA. Estimation Programs Interface Suite™ for Microsoft Windos, v.4.11. http://www.epa.gov/oppt/exposure/pubs/episuite.htm. 2012.

30. Glen G, Zartarian V, Smith L, Xue J. *The Stochastic Human Exposure and Dose Simulation Model for Multimedia, Multipathway Chemicals (SHEDS-Multimedia): Residential Module. SHEDS-Residential version 4. Technical Manual; http://www.epa.gov/heasd/research/sheds/user_information.html*. Office of Research and Development, USEPA, RTP, NC;2012.

31. US Census Bureau. 2000 Census of Population and Housing. https://www.census.gov/. 2000.

32. USEPA. Child-Specific Exposure Factors Handbook (CSEFH). In: Development NCfEAOoRa, ed. Washington, DC2008.

33. Wilson NK, Chuang JC, Lyu C, Menton R, Morgan MK. Aggregate exposures of nine preschool children to persistent organic pollutants at day care and at home *Journal of Exposure Analysis and Environmental Epidemiology.* 2003;13:187-202.

34. Miller LA, Stapleton FB. Urinary volume in children with urolithiasis. *J. Urol.* 1989;141(4):918-920.

35. Szabo L, Fegyverneki S. Maximum and average urine flow rates in normal children - the Miskolc nomograms. . *Br. J. Urol.* 1995;76(1):16-20.

36. Morgan MK, Sheldon LS, Croghan CW, et al. A Pilot Study of Children's Total Exposure to Persistent Pesticides and Other Persistent Organic Pollutants (CTEPP) - Volume 1. In: US EPA: National Exposure Research Laboratory HEaASD, ed. Research Triangle Park, NC2006.

37. Morgan MK, Sheldon LS, Thomas KW, et al. Adult and children's exposure to 2,4-D from multiple sources and pathways. *Journal of Exposure Science and Environmental Epidemiology.* 2008:1-9.

38. Burmaster DE, Anderson PD. Principles of good practice for the use of Monte Carlo techniques in human health and ecological risk assessment. *Risk Anal.* 1994;14(4):477-481.

39. Kholi JD, Khanna RN, Gupta BN. Absorption and Excretion of 2,4-Dichlorophenoxyacetic Acid in Man. *Xenobiotica.* 1974;4(2):97-100.

40. Morgan MK, Sheldon LS, Croghan CW, et al. A Pilot Study of Children's Total Exposure to Persistent Pesticides and Other Persistent Organic Pollutants (CTEPP) - Volume II. In: USEPA: National Exposure Research Laboratory HEaASD, ed. Research Triangle Park, NC2006.

41. Consumer Reports - Scale buying guide. 2013; <http://www.consumerreports.org/cro/scales/buying-guide.htm>.

42. Thomas K, Sheldon L, Pellizzari E, Handy R, Roberts J, Berry M. Testing duplicate diet sample collection methods for measuring personal dietary exposures to chemical contaminants. *Journal of Exposure Analysis and Environmental Epidemiology.* 1997;7(1):17-36.

43. CDC. Fourth National Report on Human Exposure to Environmental Chemicals. (2009).

44. CDC. Fourth Report on Human Exposure to Environmental Chemicals, Updated Tables, (February, 2015). http://www.cdc.gov/exposurereport/.

45. Morgan MK, Sheldon LS, Croghan CW, et al. Exposures of preschool children to chlorpyrifos and its degradation product 3,5,6-trichloro-2-pyridinol in their everyday environments. *Journal of Exposure Analysis and Environmental Epidemiology.* 2004:1-13

46. Kissel J. The mismeasure of dermal absorption. *Journal of Exposure Science and Environmental Epidemiology.* 2011(21):302-309.