Stochastic Human Exposure and Dose Simulation Model
For Multimedia, Multipathway Chemicals

An Update on the Development of the SHEDS-Dietary Model

July 25th, 2007

Disclaimer: This report is currently undergoing EPA review and should not be considered final.

Prepared by: Steven M. Nako[1], Jianping Xue[2]

[1] U.S. Environmental Protection Agency, Office of Pesticide Programs
[2] U.S. Environmental Protection Agency, Office of Research and Development, National Exposure Research Laboratory

Acknowledgments

We would like to thank our colleagues for their many contributions, in particular: David Miller, David Hrdy, Dana Vogel, Philip Villanueva and Hamaad Syed of the U.S. Environmental Protection Agency, Office of Pesticide Programs, and Valerie Zartarian of the U.S. Environmental Protection Agency, Office of Research and Development, National Exposure Research Laboratory.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACRONYMS AND ABBREVIATIONS

ARS – Agricultural Research Service
CARES™ – Cumulative and Aggregate Risk Evaluation System
CDC - Center for Disease Control
CM - Cooking Method
CS – Cooking Status
CSFII - Continuing Survey of Food Intakes by Individuals
DEEM™ - Dietary Exposure Evaluation Model
DWCS – Drinking Water Consumption Survey (sponsored by Bayer CropScience)
EATS – Eating at America's Table
EDF – Empirical Distribution Functions
EFH – Exposure Factors Handbook
EPA – United States Environmental Protection Agency
ERDEM - Exposure Related Dose Estimating Model platform
FCID - Food Consumption Intake Database
FF – Food Form
FFQ – Food Frequency Question
FIFRA – Federal Insecticide, Fungicide, and Rodenticide Act
FQPA - Food Quality and Protection Act
MOE – Margin of Exposure
NAS - National Academy of Sciences
NCI – National Cancer Institute
NERL – National Exposure Research Laboratory
NHANES - National Health and Nutrition Examination Survey
NHERL – National Human Exposure and Research Laboratory
NMC CRA – N-Methyl Carbamate Cumulative Risk Assessment
NOAEL – No Observable Adverse Effect Level
OPP – Office of Pesticide Programs
ORD – Office of Research and Development
OW – Office of Water
PBPK – Physiologically based Pharmacokinetic
PCT – Percent of Crop Treated
PDP – Pesticide Data Program
PF – Processing Factors
RAC – Raw Agricultural Commodity
SAIC – Science Applications International Corporation
SAS – Statistical Analysis System
SHEDS - Stochastic Human Exposure and Dose Simulation
USDA – United States Department of Agriculture

1. Background and Introduction

EPA's Office of Research and Development, National Exposure Research Laboratory (ORD/NERL) is developing a probabilistic human exposure model, the Stochastic Human Exposure and Dose Simulation (SHEDS) model for multimedia, multipathway pollutants. SHEDS-Multimedia version 3 (Zartarian et al., 2007; Glen, 2007; Stallings et al., 2007) that will be reviewed by the FIFRA SAP on August 14-15, 2007 is an aggregate (single chemical) version that includes only the residential module. SHEDS-Multimedia version 4 will have the capacity to conduct both aggregate (single chemical) as well as cumulative (multi-chemical) assessments. Version 4 will include both the residential and dietary modules, i.e., estimate both residential and dietary exposures for simulated individuals. This paper highlights research and development on the SHEDS-Dietary module.

Over the past eight years, the Agency has requested the FIFRA SAP to review several probabilistic dietary exposure models. These have included DEEM™, Calendex™, CARES™, and Lifeline. However, this is the first time that the SHEDS-Dietary module is being brought to the FIFRA SAP for external review and consultation. SHEDS-Wood (a version of SHEDS for assessing exposures to residues of wood preservatives) was also presented to the FIFRA SAP (August 28, 2002 and December 3-5, 2003); but that application focused on children's dermal and non-dietary ingestion exposures to CCA on treated decks and play sets. A dietary module developed around that time (part of SHEDS-Multimedia version 2) has been used in various ORD research activities, but has not used by the Office of Pesticide Programs to assess dietary risks to pesticides.

Agency staff from the Office of Pesticide Program (OPP) and ORD/NERL are collaborating to develop the dietary module for the SHEDS-Multimedia model. These research efforts are geared toward producing a SHEDS Multimedia Version 4.0 model that may be used to conduct an aggregate risk assessment for a pesticide (single chemical), or a cumulative risk assessment for a group of pesticides that have a common mechanism of toxicity. The Agency anticipates several advantages from a SHEDS-dietary model, including having the ability to: (1) model dietary exposures by eating occasion, (2) conduct sensitivity and uncertainty analyses, (3) evaluate the NHANES food consumption data quickly and efficiently, and (4) evaluate model performance at the per capita level against NHANES biomonitoring data following linkage with physiologically-based pharmacokinetic (PBPK) models (e.g., SHEDS-ERDEM).

There are both short- and long-term interests regarding the SHEDS-dietary module. In the short-term, the module is being used by OPP to provide supplemental analyses on Eating Occasions to support the risk characterization of the ongoing N-Methyl Carbamates Cumulative Risk Assessment (NMC CRA). That particular risk assessment is currently being finalized, and details of how this model was used will be provided in the risk assessment documents for the NMC CRA. The purpose of this update is to obtain advice on issues relating to the development of the dietary module of a SHEDS Multimedia Version 4.0 that the Agency may use to conduct an aggregate risk assessment (single chemical), or a cumulative risk assessment for a group of pesticides (e.g., pyrethroids) or chemicals.[1]

Section 2 reviews the intended use of this dietary model. The particular approach used to model food intake depends upon the question or problem that EPA is attempting to address. The dietary exposure models currently used by OPP will be discussed and how the Agency uses these models to regulate pesticides. A brief discussion of the DEEM-FCID™ model is provided since OPP has relied most often upon this model to regulate pesticides. This paper notes some features in the SHEDS-Dietary model that are believed will enhance OPP's ability to conduct dietary exposure assessments.

Section 3 presents the food and drinking water consumption data available to model dietary exposures. A particular food consumption diary is presented to clarify precisely what information is available, and the role that food recipes and processing factors play in estimating dietary exposures to pesticides. Also presented are some new drinking water consumption survey data provided to the Agency by a pesticide registrant. This survey was designed to collect data on the time and amounts of drinking water consumed by individuals – information that is not available in other consumption data.

Section 4 presents the *cross-sectional SHEDS-dietary* model. This *cross-sectional* model was developed to evaluate alternative methods for modeling dietary exposure by eating occasion. This model has also been used to explore the utility of new data provided to the Agency on the timing and amounts of drinking water consumption. Some features developed in this *cross-sectional SHEDS-dietary* model may be incorporated into the SHEDS-Dietary model.

Section 5 presents the SHEDS-Dietary model that will be included in SHEDS-Multimedia version 4. This Section includes some discussion on issues regarding the modeling of longitudinal food consumption. Section 6 highlight some options for conducting sensitivity and uncertainty analyses. Finally, Section 7 summarizes a list of research activities. In providing this update, we hope to receive the FIFRA SAP's thoughts, comments and suggestions on the development of the SHEDS-dietary model.

[1] The Agency has not made a determination on a common mechanism for pyrethroid pesticides.

2. Background and Purpose of SHEDS-Dietary Model

The SHEDS-Dietary model is being developed for use by the U.S. Environmental Protection Agency's (EPA) Office of Pesticide Programs (OPP). We begin with some background information on dietary exposure models and how the Agency uses these models to regulate pesticides.

The EPA is responsible for registering all uses of pesticides.[2] The Agency must ensure that a pesticide, when used according to label directions, can be used without posing unreasonable risks to the environment. The Agency sets tolerances (maximum permissible pesticide residue levels) for the amount of the pesticide that can legally remain in or on foods when a pesticide may be used on food or feed crops. Under the Food Quality and Protection Act of 1996 (FQPA), "the term 'safe', with respect to a tolerance for a pesticide chemical residue, means that the Administrator has determined that there is a reasonable certainty that no harm will result from aggregate exposure to the pesticide chemical residue, including all anticipated dietary exposures and all other exposures for which there is reliable information." FQPA specifies 'all anticipated dietary exposures' as potential for concurrent exposures from 'other tolerances in effect for the pesticide', and 'all other exposures' as potential for concurrent exposures from 'non-occupational uses', such as lawn care and other residential uses.

Since the passage of FQPA, OPP scientists have generally used the Dietary Exposure Evaluation Model (DEEM™) model to conduct three types of dietary exposure assessments: (i) acute, (ii) chronic, and (iii) lifetime. Our discussion will focus on modeling dietary exposures over an acute or short term duration. For acute dietary risk assessments, DEEM™ is used to obtain a reasonable high-end estimate of aggregate daily, dietary exposure to a pesticide. Agency risk assessors use DEEM™'s deterministic option to obtain an initial estimate. This entails using a single value for residue inputs (tolerances), and assuming that one hundred percent of all registered food and feed crops were treated (PCT=100%). DEEM™ deterministically tabulates total daily exposure for each food diary. The food diaries are based on the United States Department of Agriculture's (USDA), Continuing Survey of Food Intakes by Individuals (CSFII) survey; these data are discussed in Section 3. Given the conservativeness of this assessment, the per capita 95% percentile is generally used as the basis for determining a reasonable high aggregate dietary exposure to pesticides (US EPA 1999).

Pesticide residue monitoring, such as the USDA's Pesticide Data Program (PDP), indicate that tolerances are conservative estimates for food residues. If adequate residue monitoring data are available, Agency risk assessors may conduct refined acute risk assessments using the Monte-Carlo option in DEEM™. Pesticide use data are also used by the Agency estimate the percent of the crop likely to have been treated and to account for exposures from those samples that were treated but left residues at concentrations below the level of detection. The monitoring data, processing factors, the percent crop treated and level of detection, are used to construct empirical distribution functions

---

[2] See OPP's website: http://www.epa.gov/pesticides/regulating/laws.htm.

(EDF) for each registered food and feed commodity. The Monte-Carlo option randomly selects a residue from corresponding EDFs to calculate total daily exposure for a simulated person-day; further details are provided in Section 4. For such refined acute dietary risk assessments, the per capita 99.9th percentile is generally used as the basis for determining aggregate dietary exposure (EPA, 1999). This higher percentile is selected since the simulated aggregate daily exposures more accurately reflect the general population's dietary exposures to pesticides.

In addition to DEEM, some longitudinal models (Calendex-FCID, CARES, and Lifeline) have been used to conduct three cumulative risk assessments for three separate groups of pesticides. The exposure duration of interest for two of these risk assessments (the organophosphates and N-methyl carbamates) were relatively short-term (acetyl cholinesterase inhibition, less than 1 to 28 days), while the average daily exposure for the third assessment (the s-triazines) was relatively longer in duration (neuroendocrine developmental and reproductive effects, 28 – 90 days). The longitudinal models allow EPA to account for seasonal patterns in pesticide use when aggregating exposures from food, drinking water and residential uses (lawn and garden, pets, indoor).

SHEDS-Dietary is capable of modeling aggregate single day or longitudinal dietary exposure to pesticides. For a simulated individual, SHEDS-Dietary constructs a longitudinal profile of food consumption over a 365 day period in a manner similar to that used by the other models. However, the SHEDS-Dietary model will require modifications in order to utilize the food consumption data being collected by the Center for Disease Control's (CDC), National Health and Nutrition Examination Survey (NHANES). The NHANES data does not include information on season (date or time of year data collected); which is used by the SHEDS-Dietary to model longitudinal consumption patterns. While NHANES collects that information, the CDC does not release that data to protect the respondents' privacy. This additional level of security is provided since NHANES collects a wide array of sensitive personal information that the CSFII survey did not, including personal health, drug use, sexual behavior, and biomonitoring (blood, urine and swabs) data to estimate absorbed levels of pesticides and other contaminants.

In revisiting the problem of modeling longitudinal consumption, we returned to the matter of estimating exposures within a single day. Toward this end, a 'cross-sectional' version of the SHEDS-Dietary model was developed to estimate dietary exposures, by eating occasion. After assessing its capability to model exposures within a day, the problem of modeling exposures over consecutive days will be considered – an issue that may also be relevant for 'acute' (one-day) risk assessments. As an earlier FIFRA SAP (1999) noted:

> "*The traditional approach of dividing exposures (and toxicity tests) into "acute," "sub-chronic," and "chronic" time frames has several drawbacks. Shorter-term exposures add to the long-term cumulative burden, and so "acute" exposures can be relevant to chronic endpoints. After episodes of higher-than-usual exposure, the body will build up a burden, either of the agent itself or of the damage caused by an agent. Subsequent exposures must be judged not only in terms of the newly encountered agent (and the time pattern of this encounter) but also bearing in*

> *mind the lingering effects of previous exposures. In a simple illustration, an acute exposure that might be tolerated without ill effect in a previous unexposed individual could end up causing an effect in another individual with a sub-threshold burden of agent from earlier exposures. Thus, dividing exposures into duration categories and comparing each only with toxicity endpoints seen for similar durations potentially misses the key element that examining aggregate exposures is meant to address.*" (FIFRA SAP, 1999, p.30)

The exposure outputs from SHEDS-Multimedia version 4 (dietary and residential) may be used to calculate Margins of Exposures (MOEs), or if physiologically-based pharmacokinetic (PBPK) models are available, as absorbed dose inputs to such models to construct appropriate 'risk' measures (e.g., Peak inhibition, Area Under Curve) for the risk assessment.[3]

[3] The FIFRA SAP (1999) noted that exposure metrics, such as Margin of Exposure (MOE = Exposure/NOAEL), are measures of a fraction of a 'critical dose'; "when calling the MOE..a "risk Metric", the implication is that risk is linearly related to dose, and one adds up the components. … This is not really how one would wish to think about noncancer risks…" (FIFRA SAP 1999, p.43).

3. Consumption Inputs for Modeling Dietary Exposure

This section reviews the consumption data used to assess dietary exposures to pesticides. Information is also presented from a drinking water consumption survey sponsored by a registrant to support the reregistration of a pesticide. This brief overview focuses on the ability to estimate dietary exposures by eating occasion.

3.1. CSFII Food Consumption Diaries

The food consumption diaries in the U.S. Department of Agriculture's (USDA) Continuing Survey of Food Intakes by Individuals (CSFII) database are the primary source of information used to assess dietary exposure to pesticides. The food diaries contain information collected through a multiple pass, 24-hour dietary recall instrument that was administered by trained interviewers in the respondents' homes. Individuals were asked to provide food intake on 2 nonconsecutive days (3 to 10 days apart) as well as socioeconomic and health-related information.[4]

A total of 16,166 individuals provided food diaries during the initial survey period, 1994-1996. A supplemental survey was conducted in 1998 to address FQPA requirements that the USDA provide food intake data for use by the EPA to estimate exposure to pesticide residues. That effort provided food diaries for 5,496 children, 0 to 9 years old. Overall, the 1994-96, 1998 CSFII survey contains a total of 42,269 food diaries - two 24-hour diaries for 20,607 respondents; and one 24-hour diary for 1,055 respondents. The population included non-institutionalized individuals in the 48 contiguous states. These respondents resided in 12,364 distinct households; with the number of respondents per household ranging from 1 to 11 persons – the sampling design did not necessarily recruit all residents within a household. The CSFII survey developed sampling weights for each respondent. The respondents represent approximately 261 million individuals; this projection was based on the 1990 Census.

Table 1 presents the food diary data for 'Day 2' intake of a 1 year old male, weighing 13.6 kg (CSFII identification: Household ID =28517, Person number =2, Day 2). This one year old reportedly consumed 6 fluid ounces (183 grams) of milk at 7:00 am, 2 servings (92 grams) of 'Egg, whole, fried w/Lard' and 2 cups (288 grams) of 'white potato, home fries w/Lard' at 10:15 am, one serving (52 grams) of 'Chicken drumstick' and 2 more cups (288 grams) of 'home fries' at 6:00 pm, and 6 fluid ounces (183 grams) of milk at 8:00 pm.[5]

---

[4] For details, see USDA, Section 3.2 CSFII Data Collection. The NHANES food consumption data are also 24 hour recalls. The first day (Day 1) diary is collected through in-person interviews in the Mobile Examination Centers (MEC), while the second day 24 hour recall diary is collected by telephone, approximately 10 days after the in-person interview.

[5] This particular child's Day 1 intake did not include 'home fries', but did include the following: 183 grams of whole milk at 3:00 am, 80 grams of eggs at 9:00 am, 280 grams of spaghetti and 180 grams of Fruit drink at 5:00 pm, and 183 grams of whole milk at 9:00 pm.

Table 1: An Example CSFII Food Diary (CSFII ID=28517-2-2: 1 yr, M, 13.6 kg)

SEQN	Time of Day	Food Description	Amount (unit code)	Consumption (gm)	Food Source /1
1	7:00 AM	Milk, cow's, fluid, whole	6 fl.oz (10205)	183	Store
2	10:15 AM	Egg, whole, fried W/ LARD	2 XX (60919)	92	Store
3		White potato, home fries W/ LARD	2 C (10205)	388	Store
4	6:00 PM	Chicken, drumstick, with or without bone, roasted, skin eaten	1 XX (61343)	52	Store
5		White potato, home fries W/ LARD	2 C (10205)	388	Store
6	8:00 PM	Milk, cow's, fluid, whole	6 fl.oz (10205)	183	Store

/1 The Food Source variable is based on the question, 'Where was the food item obtained?' (1=store, etc.).

3.2. FCID Recipes

The U.S. Environmental Protection Agency's Food Consumption Intake (FCID) database contains recipes or each food item reported in the 1994-1996, 1998 CSFII diaries. Table 2 presents the FCID recipes for the four food items listed in the diary above. Whole milk is decomposed into fat (3.3%), non-fat solids (8.7%) and milk-water (88%). This recipe composition of fat, non-fat solids and milk-water varies by milk type (e.g., 1% milk has 1.06% fat), and across diary products (e.g., 'Cheese, processed, American and Swiss blends' has 31% fat, 24% non-fat solids, and 38% milk-water). The cooking method and food forms are also listed for each commodity. The ingredients for a particular food item may have different food forms, e.g., a cheeseburger may have uncooked lettuce (FF=110), and cooked beef (FF=213).

Table 2: FCID Recipes for Selected Foods

Ingredient	RAC Code	Commodity (RAC)	Cooking Status	Food Form	Cooking Method	Pct
Milk, cow's, fluid, whole (11111000)						
1	27002220	Milk, fat	1	1	0	3.3
2	27012230	Milk, nonfat solids	1	1	0	8.7
3	27022240	Milk, water	1	1	0	88
Egg, whole, fried with Lard (31105000-100774)						
1	25002930	Pork, fat	2	1	3	4.6
2	70001450	Egg, whole	2	1	3	109
White potato, home fries with Lard (71403000-200001)						
1	1032990	Potato, tuber, w/peel	2	1	3	82.6
2	3002370	Onion, dry bulb	2	1	3	10.6
3	19022740	Pepper, black and white	2	1	3	0.034
4	25002930	Pork, fat	2	1	3	6.8
Chicken, drumstick, with or without bone, roasted, skin eaten (24142210)						
1	40000930	Chicken, meat	2	1	1	78.8
2	40000960	Chicken, fat	2	1	1	11.2
3	40000970	Chicken, skin	2	1	1	10.1

Table 3 presents the coding for the FCID food forms. By convention, 'food form' refers to three fields: 'Cooking Status', 'Food Form' and 'Cooking Method', concatenated in that order. Therefore, food form=110 refers to a fresh, uncooked commodity, such as an 'apple', 'banana', or the fresh 'lettuce' on a cheeseburger.

Table 3: FCID Food Form Coding Scheme

Cooking Status (CS)		Food Form (FF)		Cooking Method (CM)	
Code	**Description**	**Code**	**Description**	**Code**	**Description**
1	Uncooked	1	Fresh or N/S	0	Not specified
2	Cooked	2	Frozen	1	Baked
		3	Dried	2	Boiled
		4	Canned	3	Fried
		5	Cured, pickled, smoked, salted	4	Fried or baked
		6	Not Applicable	5	Boiled or baked
Following FCID convention, 'Food Form' refers to the concatenation of the three fields: CS-FF-CM; or: Apples (raw) has FF='110'=Uncooked/Fresh/Not specified.					

The FCID recipes are important for regulatory decisions since food tolerances are set on raw agricultural commodities (RAC). While it may be nice to know that 'home fries' may contain some average amount of residues, the Agency needs to assess the contributions from potatoes, onions, pepper and pork-fat to assess exposures from the corresponding agricultural uses (food, feed, livestock, etc.). The FCID **food forms** are also important since anticipated residues may be refined based on processing studies (e.g., ketchup and tomato soup may have different residues than fresh tomatoes due to food processing and cooking).

Table 4 presents the **total** daily consumption, by commodity (RAC-FF), for the CSFII diary presented in Table 1. These values reflect the total amount consumed over all eating occasions, from all foods, based on the FCID recipes. These diaries do not contain detailed information on the time of day and corresponding amounts consumed. The dietary exposure models currently utilize these 'FCID' diaries since it eases calculation of the 'total daily' exposures.

Table 4: Example Total Daily (RAC-FF) Diary (CSFII ID=28517-2-2)

Commodity	FCID Code Crop Group-RAC-FF	Total Consumption (gm/kg bwt/day)
Milk, fat	27-002220-110	0.8 *
Milk, nonfat solids	27-012230-110	2.4
Milk, water	27-022240-110	23.6
Egg, whole	70-001450-213	7.4
Pork, fat	25-002930-213	4.2
Potato, tuber, w/peel	1-032990-213	47
Onion, dry bulb	3-002370-213	6
Pepper, black and white	19-022740-213	0.0
Chicken, meat	40-000930-211	3.0
Chicken, fat	40-000960-211	0.4
Chicken, skin	40-000970-211	0.4

* Milk, fat: 0.8 gm/kg bwt/day = {(183 gm + 183 gm)x 0.033(Pct)}/13.6 kg

Table 5 presents the consumption data for the same diary as in Tables 1 and 4, but with consumption listed by eating occasion. These data may be used for an eating occasion analysis since they retained the CSFII information on the time and amounts consumed throughout the 24-hour period.

Table 5: Example Eating Occasion Diary (CSFII ID=28517-2-2)

Time of Day	Food Description	Consumption (gm/EO)	FCID Commodity (RAC-FF)	FCID Pct of Total	Consumption (gm/EO)	Amount Consumed (gm/EO/kg)
7:00 AM	Milk, cow's, fluid, whole	6 fl.oz (183)	Milk, fat	3.3	6.1	0.4
			Milk, nonfat solids	8.7	15.9	1.2
			Milk, water	88.0	161.0	11.8
10:15 AM	Egg, whole, fried W/ LARD	2 XX (92)	Pork, fat	4.6	4.3	0.3
			Egg, whole	109	100.3	7.4
	White potato, home fries W/ LARD	2 C (388)	Potato, tuber, w/peel	82.6	320.4	23.5
			Onion, dry bulb	10.6	41.1	3.0
			Pepper, black and white	0.034	0.1	0.0
			Pork, fat	6.8	26.3	1.9
6:00 PM	Chicken, drumstick, with or without bone, roasted, skin eaten	1 XX (52)	Chicken, meat	78.8	41.0	3.0
			Chicken, fat	11.2	5.8	0.4
			Chicken, skin	10.1	5.2	0.4
	White potato, home fries W/ LARD	2 C (388)	Potato, tuber, w/peel	82.6	320.4	23.5
			Onion, dry bulb	10.6	41.1	3.0
			Pepper, black and white	0.034	0.1	0.0
			Pork, fat	6.8	26.3	1.9
8:00 PM	Milk, cow's, fluid, whole	6 fl.oz (183)	Milk, fat	3.3	6.1	0.4
			Milk, nonfat solids	8.7	15.9	1.2
			Milk, water	88.0	161.0	11.8

3.3. Drinking Water Consumption

Water may be consumed either **directly** or **indirectly** through foods (e.g., infant formula, 'kool aid', coffee, tea, etc.). Both direct drinking water and indirect (cooking) drinking water can be further decomposed by primary source (tap, bottled, other, and miscellaneous).[6] Table 6 presents the average drinking water intake, by age group, type and source. Table 6 indicates that adults (20+ yrs old) tend to consume a greater (absolute) amount of water (~1500 mL/day) than children, but infants have the highest drinking water intake per kilogram bodyweight (65.4 mL/kg bwt/day).

Table 6: Mean Per Capita Intake of Drinking Water, by Age Group and Source

	All Infants <1 yr	Children 1-2 yrs old	Children 3-5 yrs old	Children 6-12 yrs old	Youth 13-19 yrs old	Adults 20-49 yrs old	Adults 50+ yrs old
# Diaries (N)	**2,972**	**4,192**	**8,782**	**4,178**	**2,444**	**9,354**	**9,292**
	Mean Total Drinking Water Consumption (mL/Day)						
	488	**397**	**514**	**647**	**948**	**1,436**	**1,504**
	Mean Total Drinking Water Consumption (mL/kg bwt/Day)						
	65.4	**29.6**	**27.2**	**18.9**	**15.0**	**19.3**	**20.2**
Source	**Mean Direct DW Consumption (mL/kg bwt/Day)**						
Total Direct DW	**8.6**	**17.6**	**18.7**	**13.4**	**10.9**	**11.0**	**10.8**
Tap water	4.8	12.1	13.0	9.8	7.7	7.5	7.7
Bottled water	2.9	3.7	3.9	2.0	1.8	2.4	1.6
Other	0.8	1.6	1.5	1.3	1.0	0.8	1.3
Miscellaneous	0.1	0.2	0.3	0.2	0.3	0.2	0.2
Source	**Mean Indirect DW Consumption (mL/kg bwt/Day)**						
Total Indirect DW	**56.8**	**12.0**	**8.5**	**5.5**	**4.1**	**8.3**	**9.4**
Tap water	39.1	9.9	7.3	4.4	3.5	6.8	7.7
Bottled water	13.0	1.2	0.5	0.3	0.2	0.5	0.5
Other	4.6	0.8	0.7	0.7	0.4	0.9	1.2
Miscellaneous	0.16	0.14	0.04	0.02	0.03	0.04	0.10

There is a considerable amount of variation in drinking water patterns even within a particular age group. Figures 1-3 presents three different scatter plots of drinking water consumption (mL/kg bwt/day), for infants, children ages 1-2 yrs old, and adults 20-49 yrs old, respectively. Each plot presents three values for each food diary: (1) **total** consumption (black dot), (2) total **direct** water consumption (pink dot), and (3) total **indirect** water consumption (yellow dot). Figure 1 illustrates that many infants obtain most of their total intake from indirect drinking water (yellow). An inspection of the data reveals that most of the indirect water consumption is through infant formula. This is

[6] The EPA Office of Water – Science Applications International Corporation (OW-SAIC) report details the method for assigning direct and indirect drinking water, by source. Indirect 'water' does not include consumption of 'water' in manufactured Beverages (e.g., soda pop, beer, etc.), or water content of foods (e.g., watermelon). Similarly, milk-water also is not indirect drinking water. There was no indirect drinking water intake in the food diary presented above (CSFII ID=28517-2-2). The OW-SAIC report is available in the US EPA (2000) FCID CD ROM, most recent version, 3-8-2004.

consistent with the mean intakes presented in Table 6, specifically, 56.8 mL/kg bwt/day of indirect drinking water, and 8.6 mL/kg bwt/day of direct drinking water. Figure 2 indicates that direct drinking water (pink) is relatively more significant to total water intake for many children 1-2 yrs old. And Figure 3 indicates that direct drinking water (pink) is the primary source of total daily intake for most adults 20-49 yrs old (yellow dots appear to dominate due to large N = 9354x3). In all cases, there are some exceptions to this general observation. A review of the CSFII data will indicate that some 1-2 yr olds consume large amounts of indirect water ('kool-aid kids'), as do a number of adults ('coffee club').

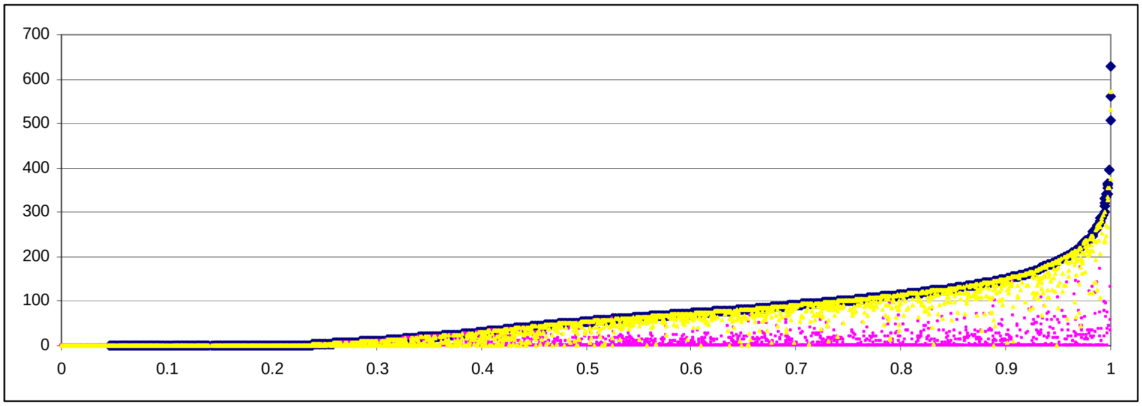


Figure 1: Drinking Water Intake, By Source (Total, Direct, Indirect): All Infants (N=2970)

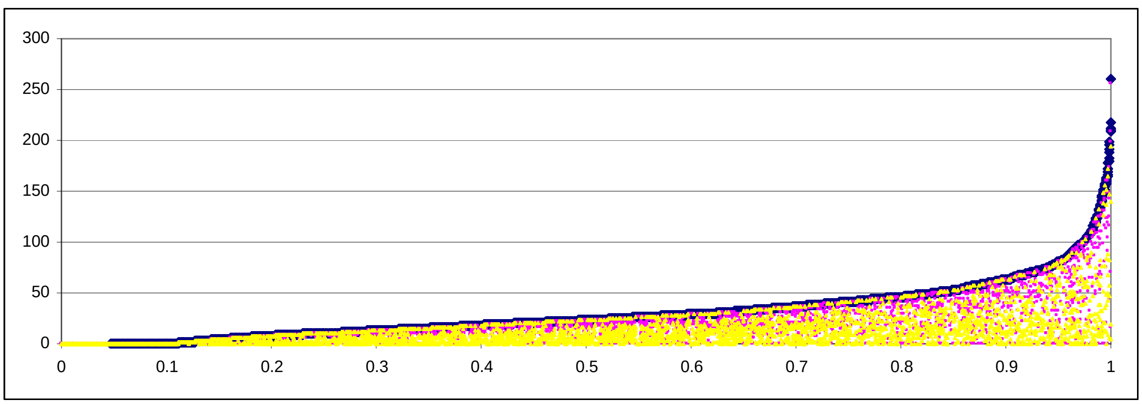


Figure 2: Drinking Water Intake, By Source (Total, Direct, Indirect) Children 1-2 yrs old (N=4192)

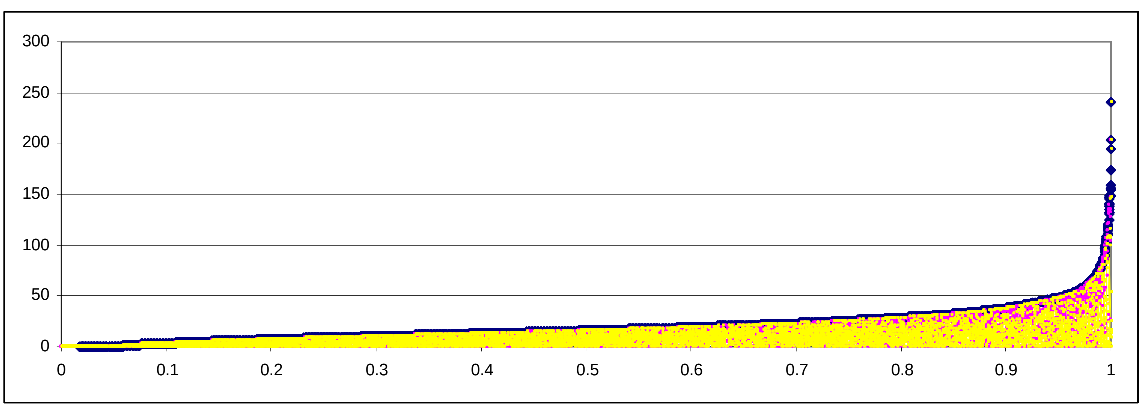


Figure 3: Drinking Water Intake, By Source (Total, Direct, Indirect): Adults 20-49 yrs old (N=9354)

3.4. Bayer Drinking Water Consumption Survey

Since the CSFII food diaries contain information on the time of day and corresponding amounts of indirect drinking water (e.g., infant formula, coffee/tea, soups, etc.), these data support the ability to conduct eating occasion analyses of **indirect** drinking water consumption. For direct drinking water exposures, the CSFII simply asked the question: "*How many fluid ounces of plain drinking water did you consume yesterday?*" The respondents provided an estimate of their direct drinking water intake, but were not asked to detail when and how much water was consumed throughout the day.

Bayer CropScience sponsored a study on direct drinking water consumption entitled "Drinking Water Consumption Survey" (DWCS). The DWCS collected information on how often, when and how much direct water is consumed at specific times during the day. The DWCS was conducted in two waves, in August 2000 (wave 1=- summer), and March 2001 (wave 2 = winter). The report provides the following description on the study design (Barraj, L. et.al., pp.9-10):

> *"The National Product Database group (NPD) was chosen to conduct this survey because of its experience in tracking the consumption habits of the US population since 1980 through its National Eating Trends (NET®) service (NET®, 2004)."*
> "*Two nationally representative samples (one for each wave) were extracted from a core sample of 250,000 households from NPD's Home Testing Institute (HTI) consumer panel. The sample for wave 1 included 3,000 households randomly selected from the core sample of 250,000 households, while in an effort to increase the number of children in the survey, the sample for wave 2 included 650 households randomly selected from households with children less than 6 years of age in addition to 3,000 households randomly selected from the core sample."*
> *"One thousand nine hundred ninety-two participants in 994 households (33% response rate) completed the first wave of the survey, and 2,950 participants in 1,320 households (36% response rate) completed the second wave of the survey.* "

Participants recorded their water consumption (time of day and amount consumed) over a one-week (7 day) period. The following information was collected in the DWCS diaries:

- Date and day of the week
- Age and gender of the household member
- Source of the home's drinking water (municipal, well)
- Time period of water consumption episode (18 hourly intervals starting at 6 am, and one 6 hr interval corresponding to the midnight-6 am period)
- Number of ounces of water consumed per time period (in 2-ounce bins)
- Where the consumption episode occurred (home/work or school/other)
- Whether the water was consumed with a meal
- The type of water consumed (tap/bottled)

A number of diaries were not used due to incomplete or missing information.[7] The resulting database contained data from 4,198 individuals from 2,154 households, providing a total of 27,282 person-day diaries (approximately 82% of the total of all participants returned diaries for all 7 days).

Figures 4 and 5, taken from the DWCS report, indicate that many respondents consume direct drinking water on multiple occasions, and throughout the day. This provides some support for using a simple modeling assumption (e.g., equal amounts allocated across 5 or 6 occasions). However, those population-based distributions do not reflect intrapersonal patterns in drinking water intakes. The report suggested that these data may be used to model drinking water exposures, by eating occasion:

> *"It may be possible, using the information collected by the DWCS to "allocate" the total daily water consumption amount reported in the CSFII into various drinking occasions. Specifically, if each subject in the CSFII survey was randomly matched to subjects in the DWCS, based on survey season, region, age, gender, and total amount of drinking water consumed per day, then the total amount reported by that CSFII participant can be allocated to the same number of drinking occasions as those reported by the matching DWCS participant. Similarly, the proportion of the total daily water consumption allocated to each of these drinking occasions can be assumed to be similar to that reported by the matching DWCS participant. This approach would then allow a less than 24-hour assessment of both food and drinking water (aggregate assessment) for a pesticide."* [8]

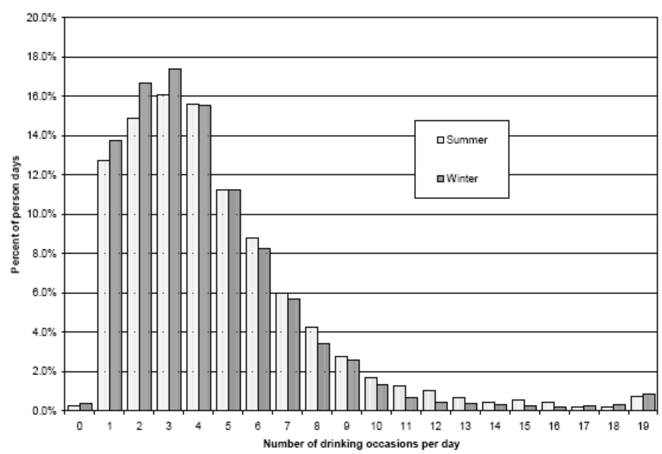


Figure 4: Total Number of Occasions of Direct Drinking Water Consumption

---

[7] Barraj, L.M. *et.al.* (2004), "Some diaries were filled out incorrectly, with duplicate or missing person IDs within a household, or with multiple entries per time interval. … If it was not possible to correctly identify the age of the participants from the demographic data fields, the diaries were discarded. Also, if it was not possible to identify the gender of participants' ages 13 years or more, their diaries were discarded……Thus, 240 respondents from wave 1 and 475 from wave 2 were dropped from the database because they had bad diaries…", pp, 10-11.

[8] Barraj, L.M. et.al. (2004), Exponent®, Inc.; National Product Database (NPD) Group., p.17.

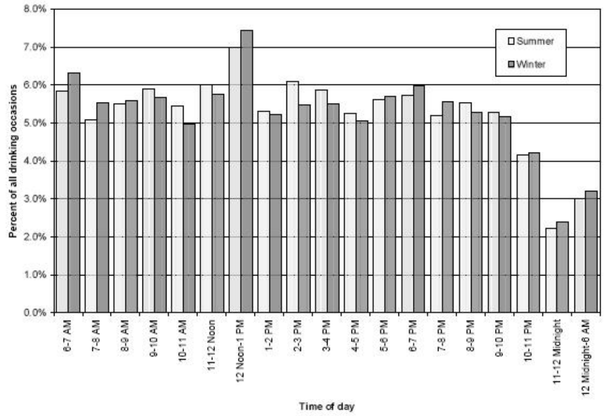


Figure 5: Distribution of Direct Drinking Water Consumption, By Time of Day

Table 7 provides the total number of drinking water diaries in the DWCS by gender, age and season. Infants were not included in this survey. The DWCS contains a reasonably large number of drinking water diaries for most of these ‘bins’. The two adult bins (20-49 yrs, 50+ yrs) contain larger number of diaries since they encompass a greater range of years.

Table 7: Total Number of DWCS Diaries, By Age Group, Gender and Season

Age Group	Gender	Season		Subtotal	Subtotal
		Winter	Summer	Age-Season	Age Group
1 yr	M	98	128	226	391
	F	136	29	165	
2 yrs	M	167	97	264	462
	F	125	73	198	
3 yrs	M	132	81	213	453
	F	151	89	240	
4 yrs	M	128	63	191	438
	F	149	98	247	
5 yrs	M	141	109	250	380
	F	67	63	130	
6-12 yrs	M	663	404	1,067	2,148
	F	624	457	1,081	
13-19 yrs	M	491	322	813	1,758
	F	577	368	945	
20-49 yrs	M	2,871	1,999	4,870	11,450
	F	4,036	2,544	6,580	
50+ yrs	M	1,975	1,688	3,663	9,712
	F	3,332	2,717	6,049	
Total	M	6,666	4,891	11,557	27,192
	F	9,197	6,438	15,635	

Barraj, L.M. *et.al.* (2004) noted that the estimated direct drinking water intakes reported by the DWCS respondents were slightly higher than the 1994-1998 CSFII respondents. For example, the overall mean intake of DWCS respondents were 37.8 oz/day (40.6=summer, 35.7=winter), while the CSFII respondents reported 29.6 oz/day

(32.4=summer, 27.8=winter).[9] The DWCS raw data files did not contain any sampling weights, nor were any formal statistical tests presented. However, the report noted that this difference may be due to the fact that "the DWCS provided participants with a time grid to report their water consumption, thus potentially helping them remember all their water consumption occasions, in contrast to the CSFII general 24hour total consumption recall question."[10]

3.5. Note on Modeling Drinking Water Concentrations

Before turning to the issue of modeling dietary exposures, two practices in the Agency's drinking water risk assessments will be discussed. First, all (**total**) drinking water consumed (both direct and indirect consumption, from all sources: tap, bottled, other and/or miscellaneous) are assumed to contain the same concentration level – i.e., only one concentration value is selected in the Monte-Carlo simulation. While some models have the option of allowing different sources to have different concentrations, those options have not been used by Agency risk assessors. Second, Agency risk assessors have not refined indirect drinking water concentrations (e.g., infant formula, coffee, tea, etc.) with cooking processing factors. Given these two practices, the discussion on modeling drinking water exposure is simplified since water is conceptually similar to any other food consumed: one residue value is randomly selected and multiplied by total intake to obtain drinking water exposures.

[9] Barraj, L.M. et.al. (2004), Table 7, p. 26. Figure 5 provides some estimates, by age groups. (p.31)
[10] Barraj, L.M. et.al. (2004), p.16.

4. Cross-Sectional SHEDS-Dietary

This section presents the basic dietary algorithm used to estimate dietary exposure over an acute – one day – duration. In addition to the FCID recipes, the Agency also developed a database of total daily food diaries, such as the one presented in Table 4, for each CSFII food diary. These 'FCID' diaries were provided to the exposure modelers for use in modeling daily dietary exposures.[11] The following discussion presents the dietary algorithm, and the extension of this algorithm to include eating occasions.

4.1. Basic Dietary Algorithms

In this Section we present two algorithms for modeling dietary exposure. The 'total daily' algorithm, based on the FCID (RAC-FF) diaries, is used by DEEM-FCID and other dietary exposure models. The second algorithm is a simple extension to account for the time of day. This extension is required to assess dietary exposures by eating occasion.

For each diary, a Monte-Carlo simulation is performed selecting a residue value for each commodity (RAC-FF). Each commodity exposure is calculated by multiplying total daily consumption with corresponding residues; and aggregate daily exposure is calculated by summing exposures across all commodities, as depicted in Equation (1).

Equation (1) - Total Daily Approach:

$$\underset{\text{(mg ai/kg bwt/day)}}{\text{Dietary Exposure}} = \sum_{i \in \text{RAC-FF}} \underset{\text{(gm food/kg bwt)}}{\text{Consumption}_i} \times \underset{\text{(1 kg food/1000 gm food)}}{\text{Unit Conversion}} \times \underset{\text{(mg ai/kg food)}}{\text{Residues}_i}$$

where 'Consumption' is normalized based on the CSFII respondents' bodyweights (grams food/kg bwt/day).[12] For simplicity, 'Residue' concentrations also reflect Processing Factors. The exposure models assume that commodity residues are independently distributed.

Detailed food diaries, such as the one presented in Table 5, are needed to assess dietary exposures by eating occasion. Equation (2) is generalized to include the time of day (t), for both consumption and residues:

Equation (2) – Extension to Eating Occasions:

$$\underset{\text{(mg ai/kg bwt/EO(t))}}{\text{Dietary Exposure(t)}} = \sum_{i \in \text{RAC-FF}} \underset{\text{(gm food/kg bwt)}}{\text{Consumption}_i(t)} \times \underset{\text{(1 kg food/1000 gm food)}}{\text{Unit Conversion}} \times \underset{\text{(mg ai/kg food)}}{\text{Residues}_i(t)}$$

In principle, both anticipated (food) residues as well as drinking water concentrations may vary by eating occasion and/or across foods consumed within an eating occasion.

---

[11] Table FCIDFFCM5 in the FCID CD ROM (FCID_LL.MDB, Latest Version Released, March 8th, 2004)

[12] With the exception of the Lifeline model, the dietary exposure models utilized the food consumption data in terms of amount of food consumed per kilogram bodyweight of the corresponding CSFII respondent. Further discussion is provided in Attachment 1.

DEEM-FCID™

The following is a description of how the DEEM-FCID™ model works and how the Agency uses the model to regulate pesticide. For each food diary, DEEM-FCID™ applies a Monte-Carlo simulation to calculate total daily exposure, as depicted by Equation (1).[13] DEEM-FCID™ conducts a fixed number of 'iterations' to each food diary. The number of iterations is a user-specified parameter. DEEM-FCID™ keeps track of the total daily exposure for each simulated person-day, and applies the corresponding CSFII survey weights to project the simulated person-days to calculate per capita based percentiles. If the user specifies only one iteration per diary, then the per capita percentiles reflect interpersonal variability – a particular diary may tend to have high exposures due to high consumption of various foods, but it may also have little or no exposures depending on the residues selected in that simulation. If multiple iterations are specified, DEEM-FCID™ treats each modeled person-day as a separate (independent) simulation, and the per capita estimates of all simulated person-days reflect both intrapersonal variability and interpersonal variability. Agency risk assessors typically run DEEM-FCID™ with 1,000 iterations per diary. This number of iterations produces stable results, i.e., minimal simulation or random seed uncertainty, at the per capita 99.9th percentile for most pesticide risk assessments.[14] This is discussed further in Section 6 (uncertainty analyses). For now, it is noted that the purpose of these Monte-Carlo simulations is to obtain a reasonable high-end estimate of aggregate total daily exposure.

4.2. *Cross-sectional SHEDS-Dietary*

Cross-sectional SHEDS-Dietary applies a fixed number of iterations to each CSFII food diary, and utilizes the CSFII sampling weights to project exposures at the per capita level in a manner similar to DEEM-FCID™. The major difference between these two models is that *Cross-sectional SHEDS-Dietary* retains exposures, by Eating Occasion, as described in Equation (2). Otherwise, the similarity in model design minimizes any modeling uncertainties when comparing total daily exposures against the DEEM-FCID™ model.

Cross-sectional SHEDS-Dietary assumes that only one residue value is selected for each food-RAC-FF (i.e., if a RAC-FF occurs on two separate eating occasions (e.g., 'home fries at 10:15 am and again at 6:00 pm), then each ingredient in that food has the same residue for both occasions). The model draws a separate residue value if that commodity (RAC-FF) is found in two separate foods, e.g., *Cross-sectional SHEDS-Dietary* randomly

[13] The Monte Carlo procedure draws a residue for each RAC-FF. While a particular commodity (Potato, tuber w/peel) may be used in multiple foods, the cooking method may differ, and thus, it will have a different food form. The food form for potatoes used in 'White potato, home fries w/Lard' is 'cooked-fresh-fried' (ff=213, see legend in Table 3). This particular diary may have contained other foods with 'Potato, tuber w/peel' - some of which may have the same food forms, e.g., 71411000- 100701='White potato skins, with adhering flesh, fried, with cheese and bacon', while others have different food forms, e.g., 71603010='Potato salad', 71101110='Baked potato'. If the cooking method is the same (e.g., ''Pork fat' or 'Lard' used to fry eggs and home fries), then the same residue is applied to all those consumption amounts ('home fries', 'White potato skins', etc.). But if the food forms are different (e.g., 'Potato salad' is boiled, ff=212; 'Baked potato', ff=211), then a different residue is independently drawn and applied for those food forms in the total daily simulation.

[14] Risk assessors may increase this to 5,000 or more iterations if the results are sensitive at this level.

selects two residue values for pork lard, one that is used to cook fried eggs, and another value for pork lard used to prepare home fries.

If new residue values were to be selected for each new eating occasion, then this modeling assumption would lead to greater differences with DEEM-FCID™. Otherwise, the major difference between DEEM-FCID™ and *Cross-sectional SHEDS-Dietary* is that the latter model keeps track of consumption and exposures throughout the simulated day. The selection of two residue values for pork lard produces minimal differences for the particular diary discussed above.

Appendix 1 provides a comparison between *Cross-sectional SHEDS-Dietary* and DEEM-FCID™ for chemical 'ABC'. The per capita exposure estimates for this chemical at the upper percentiles (95^{th}, 99^{th}, 99.9^{th}) are relatively similar across the two models. The differences primarily reflect simulation uncertainty since the models both rely upon the CSFII sampling design.

The FIFRA SAP previously noted the importance of respecting differences due to model uncertainty (differences across models).[15] The issue of model uncertainty is beyond the scope of this paper. *Cross-sectional SHEDS-Dietary* was developed to evaluate the incremental effects of specific modeling assumptions. This tool can also help explore the utility of various types of sensitivity analyses (Section 6).

[15] FIFRA SAP (2004), p. 24.

4.3. Options for Sampling Food Residues

Cross-sectional SHEDS-Dietary has the option of conducting Monte-Carlo simulations using residues for food items, as well as commodities (RAC-FF).[16] If the user has residues for a particular food (e.g., milk), then *Cross-sectional SHEDS-Dietary* can randomly draw a residue from that corresponding distribution. Otherwise, *Cross-sectional SHEDS-Dietary* will randomly select a residue for each of the RAC-FF ingredient (e.g., fat, non-fat solids, milk-water) following the same procedure applied by the other exposure models. The current procedure has the potential to overestimate the percent of the population exposed, while underestimating their levels of exposure. To illustrate this effect, assume that about one percent of all milk (whole, skim or reduced fat) contains residues of a pesticide or contaminant. If the same residue distribution is applied to each of the three ingredients (fat, non-fat solids, and milk-water), and residue values are randomly selected from this same distribution for each of the three ingredients, then approximately three percent of the milk consumers will have a residue on one or more of these ingredients – and only a subset of these consumers will have residues (likely different values) in all three ingredients. This food residue option is equivalent to assuming perfect correlation across the ingredients – an individual that drank a glass of milk either did or did not select a residue, on all three ingredients or none of these three ingredients. We anticipate that this option may be useful for a subset of food items, such as milk (versus other dairy products), and meats ('beef steak' versus other meat products).

If residue data for specific commodities are used, then other information may be used to refine these anticipated residues. For example, if we have monitoring data on fresh potatoes, we may use the pesticide use data to refine the estimate of total samples analyzed that were treated versus not treated. The user can determine how many samples that did not have detectable levels residues should have residues with 'half the level of detection' versus zero values (not treated). Processing factors may also be used to refine the anticipated residues, e.g., a cooking factor may be applied to refine the residues for home fries (fried) or mash potatoes (boiled).

Finally, we are exploring other 'distributional fitting options' for commodities. In principle, residues may be correlated across commodities and over time. For example, pesticide use on apples and pears may be positively correlated due to similar pest pressures, or someone consuming apples over multiple eating occasions (or sequential days in a longitudinal model) may have some autocorrelation in residues from occasion to occasion and day to day.

This extension may also apply across chemicals for a cumulative risk assessment. In two cumulative risk assessments, the Agency has relied upon the USDA Pesticide Data Program (PDP) to empirically capture co-occurrence across chemicals within a particular

---

[16] This option for assigning residues at the food level is also available in the Lifeline model, although the Agency has not utilized that option to date.

sample. Since PDP analyzes for many pesticides and metabolites of concern, empirical use of these data enabled the Agency to account for both the probability and magnitude of two or more chemicals being found on a particular commodity-sample. The drawback with this approach is that it is difficult to utilize residue data from two different sources for the same commodity – e.g., a chemical (or metabolite) may require a separate analytical method and so one may data from a special study (*e.g.*, market basket) in addition to the primary data source (*e.g.*, PDP). We can better utilize such data from different sources by estimating correlations in the pesticide use data, or in older residue data (if available), and allowing the user to specify such correlations across commodities, over time, and across chemicals.

4.4. Modeling Food Residues by Eating Occasion

An earlier version of DEEM™ included an Eating Occasions option that performed a Monte Carlo simulation by selecting a new food residue for each eating occasion. The Agency asked the FIFRA SAP, "Under what circumstances should the EPA consider using the (DEEM™) Eating Occasion approach?" The FIFRA SAP (1999) noted:

> *In terms of the actual exposures, the "Daily Total" approach seems most appropriate in situations where an individual would have multiple servings from a single unit of food (e.g., several slices of a single watermelon) over the course of day. ….. If one is looking for acute toxicity in a fast-clearing pesticide, then only the "Eating Occasion approach is appropriate. However, DEEM™ does not consider "binge" and other special eating habits, and data on rarely-eaten foods will come from relatively few individuals, and these factors may limit the validity of "Eating Occasion" estimates.*:
>
> "*Dietary exposure analysis is an extremely complex process. It utilizes many pieces of data from different sources, each carrying its own limitations and deficiencies for the purpose. Therefore, a careful documentation of the database limitations and the uncertainties associated with the estimated exposure is essential for a proper interpretation of the exposure estimates.*" [17]

The qualifying comments reflect a complexity in accounting for differences in eating habits across the population. To illustrate this point, the food consumption diary highlighted earlier (CSFII ID=28517-2-2) indicated that the 1 yr old consumed the same food ('home fries') on two different eating occasions. It is possible that the child had 'leftovers' in the evening meal. If that were the case (or if his mother prepared more home fries from the same bag of potatoes), then it would be appropriate to assume that the same level of pesticide residue was present in the potatoes on both eating occasions. On the other hand, if the child had consumed 'home fries' from two different fast food restaurants, then it may be more appropriate to randomly draw separate residue values for each eating occasion. Such conditional modeling decisions can better made after a closer inspection of the food consumption data.

[17] FIFRA SAP (2000) Report No. 2000-01B, May 25, 2000, Pages *33-35*.

One research activity that is planned is to conduct a systematic review of the food consumption diaries (Section 7). Such analyses may lead to better decision rules for selecting residues for different eating occasions. A decision rule could require the modeler to determine if the commodity comes from the same food item and if so, require that the same residue value be used. If not, the model would compare the foods' sources, time of eating occasions, and if the foods were eaten at home. If those factors differ by some condition, then the model would select different residues values. The potential for different plausible decision rules suggests that some type of uncertainty analyses may be appropriate to address this data uncertainty.

4.5. Modeling Drinking Water Exposures

The Agency generally uses environmental fate models (e.g., PRZM-EXAMS) to generate predicted drinking water concentrations for pesticides. The duration of the modeling period, often 30 or more years, is based on the availability of meteorological data for a particular location. For example, rainfall (and other meteorological) data from January 1st, 1960 through December 31st, 1990 may be available for a particular site in the Midwest. The environmental fate models predict drinking water concentrations based on that rainfall data as well as data from other inputs (e.g., half-lives, soil types, and pesticide use - application method, rates and timing, etc.). This modeling effort may produce over 11,000 (=31 years x 365 days/yr) predicted drinking water concentrations.

DEEM-FCID™ model randomly draws one value from the empirical distribution of predicted drinking water concentrations for each simulated person-day. The current versions of both *Cross-Sectional SHEDS-Dietary* and SHEDS-Dietary apply a similar procedure for utilizing the drinking water concentration data. Appendix A (Figure A.1) plots DEEM-FCID™ per capita estimates for 17 separate drinking water scenarios for chemical 'C' at the 99.9th percentile for nine subpopulations against *Cross-Sectional SHEDS-Dietary* estimates – the two models produce similar estimates.

Cross-Sectional SHEDS-Dietary utilizes the CSFII data to assess the timing and amounts of **indirect** drinking water intake within a simulated person-day. The model contains two options for allocating **direct** drinking water consumption throughout the day: (1) fixed approach, and (2) empirical use the Bayer DWCS data. In the fixed approach, SHEDS allocates the CSFII respondents' total direct drinking water consumption (mL/day) as equal amounts over 6 fixed occasions (6:00 am, 9:00 am, 12:00 noon, 3:00 pm, 6 pm, and 9:00 pm). The second option uses the Bayer DWCS data to allocate the amounts and times that **direct** drinking water was consumed throughout the simulated person-day.

Although the DWCS study did not have the same level of rigor as the CSFII, we think that these data are useful to assess the timing of direct drinking water intake for several reasons, including: (1) the marketing firm, the NPD group, has extensive experience at monitoring eating trends in the US and Canada, (2) the design of the data collection instrument may have led to better 24-hour recall, (3) the survey received reasonable response rates (>30%), and (4) a relatively high percent of respondents completed 7 day diaries (82%), and (5) the 7-day study period reduces the need to model intrapersonal variability over this duration.[18] The DWCS option involves the following steps:

1. Generate cohort ('bins') by gender, age, season (Table 7)
2. For each DWCS diary, calculate the percent of Total Direct DW, by Occasion
3. For each CSFII diary, randomly select a Bayer DW diary from appropriate 'bin'
4. Use the percentage of DW from DWCS data to allocate the total amount of direct DW (CSFII) across drinking occasions throughout the simulated person-day

This second option does not apply to the infant subpopulation since the DWCS survey did not include infants; however, this concern is partly alleviated for this subpopulation due to the relative importance of indirect drinking water (formula) versus direct drinking water.

Some preliminary analyses indicate that the DWCS option provides similar, but slightly higher peak exposures at the per capita 99.9th percentile than the fixed option with 6 equal allocations. The similarity between these two options appear consistent with the DWCS data that many people consume direct water on multiple occasions, but often less than six, throughout the day – fewer occasions mean higher intake per occasion.

We found it appealing to empirically utilize the DWCS diaries to allocate direct drinking water consumption throughout the day. However, we do not know precisely what these CSFII respondents actually did, and therefore some type of uncertainty analyses may be helpful to characterize these results. The FIFRA SAP (2005) suggested allocating water intake over five events – with three occasions during meals (25% per occasion) and two occasions in between meals (12.5% each).[19] This option will require a bit more programming since different people have different eating patterns (less than or greater than three meals per day). Further exploration of the DWCS data may suggest that other factors, in addition to age, gender and season (e.g., consumed with meal, total number of occasions, total amount consumed, etc.) may be used to 'bin' diaries.

---

[18] This algorithm can be modified for longitudinal models, 'binning' respondents (persons), rather than diaries (person-days) to retain the intrapersonal information contained in these 7-day drinking water diaries.

[19] "In further development of this approach, EPA should make use of any reliable source of relevant empirical data on daily patterns of drinking water consumption; ideally adapted to the likely consumption behavior in specific regions or smaller areas of the country." FIFRA SAP 2005, p.59.

5. SHEDS-Dietary Model

5.1. SHEDS-Dietary

SHEDS-Dietary can be used to model longitudinal dietary exposure. SHEDS-Dietary creates a modeled individual (reference population) by randomly drawing a person from a demographic table based on the 2000 U.S. Census.[20] The CSFII food diaries are grouped by age and gender, and 'diary pools' are created based on Season and Day of Week (weekday or weekend) for each age-gender cohort.[21] For each modeled individual, SHEDS-Dietary constructs a longitudinal profile of food consumption by randomly selecting 8 food diaries - one weekend and one weekday, for each of the four seasons - from the appropriate diary pools. Appendix 2 contrasts this with the approaches used by the other models.[22]

5.2. Discussion on Modeling Longitudinal Food Consumption

The SHEDS-Dietary model will require change in order to use the recent National Health And Nutrition Examination Survey (NHANES) food consumption data. NHANES does not provide data on location (region) or calendar dates (season), the latter data is used by SHEDS to model longitudinal consumption. We plan to pursue to two approaches for modifying SHEDS-Dietary model. The first approach follows the new method for developing longitudinal activity profiles described by Glen et.al. (2007). The method is designed to develop longitudinal patterns from which intrapersonal and interpersonal measures can be derived. The issue is to determine which measure to use (x-score) as a potential for dietary exposure. There are many potential covariates and measures of diversity with respect to food consumption patterns across a subpopulation and over time. The diet, health and nutrition literature contains a rich volume of research, indicating that food consumption patterns may vary by race, ethnicity, lifestyle (activities and energy requirements) and socio-economic factors.

This literature also suggests an alternative method for modeling longitudinal consumption profiles.[23] A promising effort is described by a team of researchers from government (NCI) and academic and some private institutions. In the first of three papers, Dodd *et.al.* (2006) review existing methods used to estimate long-term dietary intake using cross-sectional data. Tooze *et.al.* (2006) present a new method for estimating long-term intake of episodically consumed foods using food frequency questions (FFQs). An example of a FFQ is, 'How often have you (respondent) consumed fish during the past 30 days?' The team presents a two stage model with the first part (logistic regression) predicting the

---

[20] This is discussed in the SHEDS-Multimedia Version 3 Technical Manual.

[21] This methodology is similar to the construction of longitudinal activity profiles from the Consolidated Human Activity Data base (CHAD).

[22] Appendix 2

[23] Tran, N.L. et.al. (2004) also present a method for combining food frequency and survey data to estimate long-term (30 day average daily intake) exposure to mercury from fish consumption. One of our research items is to conduct a literature review.

probability of consuming a particular food and the second part (regression on log transformed consumption amount) predicting the amount of food consumed (>0). In their third paper, Subar *et.al.* (2006) apply this method to the Eating at America's Table Study (EATS) data. Subar *et.al.* (2006) also provide a brief review of the development of the Food Propensity Questionnaire, a set of FFQs that was introduced in the 2003-2006 NHANES. Following a more comprehensive review of this literature, we may develop more than one approach for modeling longitudinal consumption to obtain a range on how different approaches may affect particular exposure measures.

Longitudinal food consumption is modeled to account for seasonal patterns in exposures from food, drinking water and residential uses. For food alone, the longitudinal dimension does not appear to affect per capita estimates of a single day aggregate exposures.[24] The focus on single day exposures may expand as the Agency develops physiologically-based pharmacokinetic models (PBPK) for pyrethroids and other chemicals. The SAP (2005) noted that one-day simulation models may underestimate risks if carry-over effects from consecutive days of high exposures are of concern.[25] Single day models, like DEEM-FCID™ and *Cross-Sectional SHEDS-Dietary* are unable to account for carry-over effects. Consecutive days of relatively high exposures may occur in all three sources (food, drinking water and residential). Strong patterns in drinking water exposures may be anticipated since most people consume water daily and both the surface water and ground water models generally produce drinking water concentrations that exhibit positive autocorrelation. Similarly, exposures from residential uses will reflect seasonal patterns in persistence in surface residues following product use and the daily habits of many people. For food, one can conceive of an individual purchasing a bag of apples and consuming one or a few apples from that bag over consecutive days. The relative contribution to total exposure from these three sources (food, drinking water, and residential) varies by chemical as well as across individuals within a subpopulation. Lu *et. al.* (2006a, 2006b) report that residential uses appear to contribute more toward typical exposures to some synthetic pyrethoids, while dietary exposures tend to be the bigger contributor for some organophosphate pesticides. While the dietary pathway may play a lesser role for pyrethroids, we anticipate that these efforts on modeling longitudinal food consumption may be important for future dietary exposure and risk assessments to other pesticides and other chemicals.

---

[24] See US EPA (2004) for some comparisons. The Agency has not refined food residues, by region nor season. Matching consumption with residues by season (and region) may have a modest effect on the overall results. For seasonal residues, DEEM-FCID and older version of Calendex-FCID would need to be run for only a subset of the population (Winter, Spring, Summer, Fall); and outputs cannot be merged across separate runs to calculate an overall per capita 99.9th percentile.

[25] FIFRA SAP (2005), Minutes, p.10, "In particular, if one applies a 4.1-fold inter-species scaling factor to the 5.4 hr half-time for reversal of brain AChE inhibition in rats, one obtains a predicted half-time of 22 hr in the 70 kg human adult. Such a long half-time would force the risk assessment model to address carryover of inhibition from one day to the next. In considering this issue, the Agency should take into account cases where there is a dose dependency for inhibition reversal half-lives." p.56.

6. Potential Applications for SHEDS-Dietary

The FIFRA SAP has noted the importance of conducting sensitivity and uncertainty analyses on several occasions. An earlier FIFRA SAP (1998) commented on the potential for extending probabilistic methods to the toxicity data to address uncertainty in the toxicity endpoints.[26] However, the focus here is on the exposure side. The following quote is from an earlier FIFRA SAP reviewing the DEEM-FCID™ model:

"*An uncertainty analysis should accompany a dietary exposure analysis. The complexity of dietary exposure estimates underscores the importance of presenting the commodity contribution and uncertainties associated with an analysis. In light of the lack of a built-in uncertainty analysis tool in DEEM™, it is recommended that multiple sets of dietary exposure analyses be routinely conducted to capture the impact of the critical factors that are identified in the steps leading up to the dietary analysis (e.g., the choice of residue data, whether to combine residue data from regions, seasons, or years, differences in eating habits and preferences). A simple hand-calculation test is recommended for testing the reality of the exposure and risk estimates from a dietary exposure software program.* " (FIFRA SAP minutes, 2000, p. 35)

Sensitivity and contribution analyses are a routine part of OPP risk assessments. While DEEM-FCID™ does not have automated procedures, a user can modify inputs and rerun the model to evaluate incremental effects. OPP risk assessors conduct various types of analyses, including: (i) using different sources of residue data (e.g., Market Basket Survey), (ii) refining estimates on the percent of samples treated (Half Level of Detection), and (iii) 'dropping' commodities to see how various mitigation measures affect exposure at a high percentile. These analyses help inform the risk manager how exposures may change when certain model inputs are modified. These modifications to the model inputs are typically performed "one at a time" to permit isolation of the effect. In a typical risk assessment, all the dietary consumption data (*i.e.*, reported CSFII diaries) are used along with the best available pesticide residue data. OPP risk assessors specify a sufficiently large number of Monte-Carlo iterations such that exposure estimates are stable with respect to the random seed.

Table 8, below, lists other types of analyses that can be performed to complement a dietary exposure assessment.[27] The first application is a simple tool, but a potentially valuable one for users that do not have data base querying capabilities. If potatoes are significant contributors to dietary exposure, then the user can view the food diaries for

---

[26] "The FIFRA SAP concluded that it is appropriate for the Agency to move toward probabilistic techniques for toxicity endpoints. Agency policy concerning probabilistic methods does not prohibit or exclude the possibility of applying distributions to toxicity data. …Whether they are explicitly recognized or not, variability and uncertainty in toxicity estimates are key contributors to variability and uncertainty in resulting risk estimates." FIFRA SAP (1998), p.2 of 9.

[27] The Agency's Exposure Factors Handbook (EFH) outlines three types of uncertainty: (1) Scenario Uncertainty (Descriptive errors, aggregation errors, judgment errors, incomplete analyses), (2) Parameter Uncertainty (measurement errors, sampling errors, variability – time/space, surrogate data), and (3) Model Uncertainty (relationship errors, modeling errors). EFH Volume I, General Factors, Chapter 2, pp.2-5, 2-6.

the top 10 potato eaters and see how much potatoes was consumed and what other foods they reportedly consumed that day. Following a recent FIFRA SAP (2004) recommendation to publish descriptive statistics on food consumption patterns, this tool will make it easier to for users to see what percent of all children eat 'potatoes' on any given day, and a typical (50%) and high amount (90%) consumed.

Table 8. Potential Applications of SHEDS-Dietary

Variable/Modeling	Description
User-only Analysis	
Data Viewer: Top 10 Eaters, Foods & Descriptive Statistics	A simple tool that shows what people eat, and who are the top 10 eaters
Eaters-only Report	Deterministic calculation of exposures among people that consume a treated commodity or food
Contribution Analysis	
Shares of Total Exposure, by Commodity, Food or Diaries	Current reports provide shares of total exposures (99.9^{th} – 100^{th} percentiles), by commodity or by food
Shares of Total Exposure, by Chemical – Commodity, Food or Diaries	For cumulative exposure assessments, SHEDS keeps track of residues, by chemical (i.e., not used RPF combined residue)
Shares of Total Person-days, by Commodity, Food or Diaries	(i) 'Exceeders' or shares of total person-days (99.9^{th} – 100^{th} percentiles), by commodity. (ii) focus on diaries: percent of simulations exceeding target
Sensitivity Analysis	
Consumption 'Outliers'	Effect of Diaries with Reported High Amounts Consumed
Percent Samples Treated	Effect of the Estimated Percent of Samples Treated (Half Level Of Detection (Half-LOD) used for monitoring data)
Percent Crop Treated	Effect of Annual Fluctuations in Percent Crop Treated (assuming all other factors constant)
Processing Factors	Effect of Estimated Processing Factors
Uncertainty Analysis	
Uncertainty - Cohorts	Effect of Different Factors for Developing 'Cohorts' or 'Bins' for Food Diaries
Uncertainty - Subsamples	Effect of using a Subsample of the Food diaries and Residue data on per capita estimates (200 person-years)
Uncertainty – Subsamples of residues, by commodity	Residue by commodity
Uncertainty - Models	Comparing Results Across All Models

The second item in the table is a simple Eaters-only report. We define an 'Eater' as someone that consumes a treated commodity. This is different from the DEEM-FCID™ users' report which provides percentiles based on all individuals (diaries) that consumed any food, whether or not any of the commodities were treated. This report will be provided on a commodity basis to "ensure that a RAC with a high level of risk would not 'slip through the system'."[28]

Contribution analysis is an important tool for risk management. If exposure at the upper percentile is high, then risk managers need to know which commodities are contributing to that result. Agency risk assessors use DEEM's Critical Exposure Contribution Analysis report to rank commodities based on the commodities' shares of the total

[28] FIFRA SAP 2004, p.26.

exposures from all simulated diaries in the top percentile (e.g., $99.9^{th} – 100^{th}$). This ranking is used to determine which uses can be mitigated, such that aggregate exposure at the particular percentile does not exceed the target level. Following an earlier FIFRA SAP's recommendation, we will explore other measures for quantifying contributors.[29] For example, we can calculate a frequency exceeded report that informs the risk manager how often a commodity tends to show up in the tail, and the average exposures that eaters have from consuming treated commodities. Other measures may help to get a sense of interpersonal variation in exposures across the subpopulation. For example, if a subpopulation contains a total of 1000 diaries, and the user specified 1000 Monte-Carlo simulations per diary, then we can calculate how many diaries are found – at least once - in the tail, and how frequently (out of the 1000 simulations) those diaries tend to be there.

6.1. Sensitivity Analyses

The most difficult part of conducting sensitivity analyses is in the problem formulation: defining a particular issue of concern, evaluating the available data inputs, developing method(s) to assess how sensitive the results are to that concern, and characterizing the degree to which that analysis addresses that concern. This section briefly describes two applications using *Cross-Sectional SHEDS-Dietary*: (1) sensitivity analyses on food and drinking water consumption 'outliers', and (2) uncertainty analyses focusing on the selection of food consumption diaries. These two examples provide an overview the problem formulation and the potential applications of SHEDS-dietary.

A component of the Agency's risk characterization is to "Evaluate the tails of the food exposure distribution to verify that unusual consumption patterns are not inappropriately impacting on the results of the assessment."[30] Identifying 'unusual' consumption patterns requires inspection of the food diaries. If the amounts consumed appear reasonable, then no further analyses are required. As the FIFRA SAP noted,

> "*The CSFII is designed to be representative of the population as a whole. Hence the "tails" of the distribution are still part of the distribution and, therefore, cannot be said to impact the results of the assessment inappropriately.*"[31]

If consumption values are unusual so as to bring into question the accuracy of the data (e.g., measurement or data entry error), then quantitative 'what-if' analyses may be appropriate. We want to know how sensitive exposure at the upper percentiles is to one or a few such data records. The open source coding of *Cross-Sectional SHEDS-Dietary* enable the Agency to perform such analysis in a quick and cost-effective manner.

---

[29] FIFRA SAP (2004), p.26 "Both CEC and 'frequency-exceeded' analysis are valuable for identifying significant contributors to the dietary exposure. … Until an all-inclusive measure for risk management is identified, all methods of measures should be explored."

[30] EPA SAP, 2005, p.187.

[31] SAP minutes, 2005, p.36.

Sensitivity Analyses on Food Consumption 'Outliers'

Figure 6 presents a Box and Whisker plot of potato consumption among children ages 1 to 2 years old.[32] The amount of potatoes consumed by the CSFII diary highlighted earlier (ID=28517-2-2) is more than twice as high as the second highest eater in this age group. This amount appears to be an outlier when focusing on only 'fried' potato consumption, but not so much the case when considering potato consumption in other food forms (e.g., boiled). As an absolute amount consumed, it does not appear to be implausible, i.e., a 1
yr old, 13 kg boy eating 300 grams of home fries on two occasions. But a considerable amount of resources may be expended to defend that assessment. Using the *Cross-Sectional SHEDS-Dietary* model, one can determine if the per capita estimates are sensitive to this one diary. In particular, if either (i) this 'outlier' was removed from the Monte-Carlo simulations, or (ii) the amount consumed was adjusted to a lower level (e.g., second highest amount), the per capita estimates at the 99.9th percentile would not change considerably. For one particular set of residues we analyzed, we found that the results are fairly robust to this one diary. Consequently, there does not appear to be any reason for conducting further sensitivity analyses for this commodity.

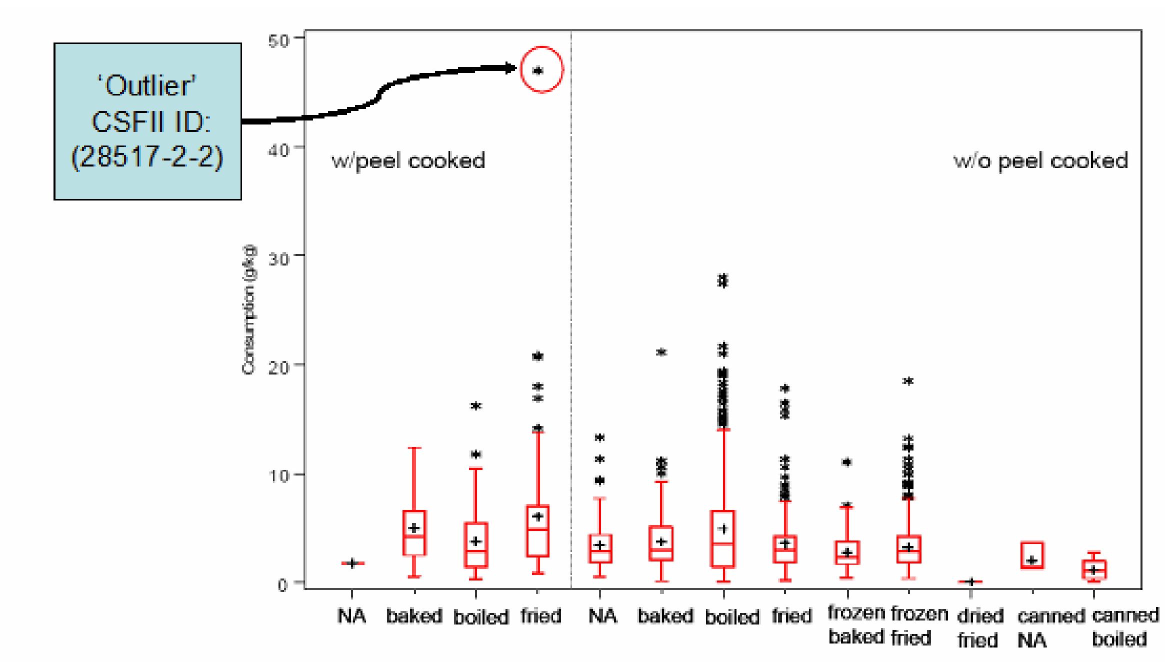


Figure 6: Box-and-Whisker Plot of Potato Consumption by Day for Children 1-2 Yrs. Old

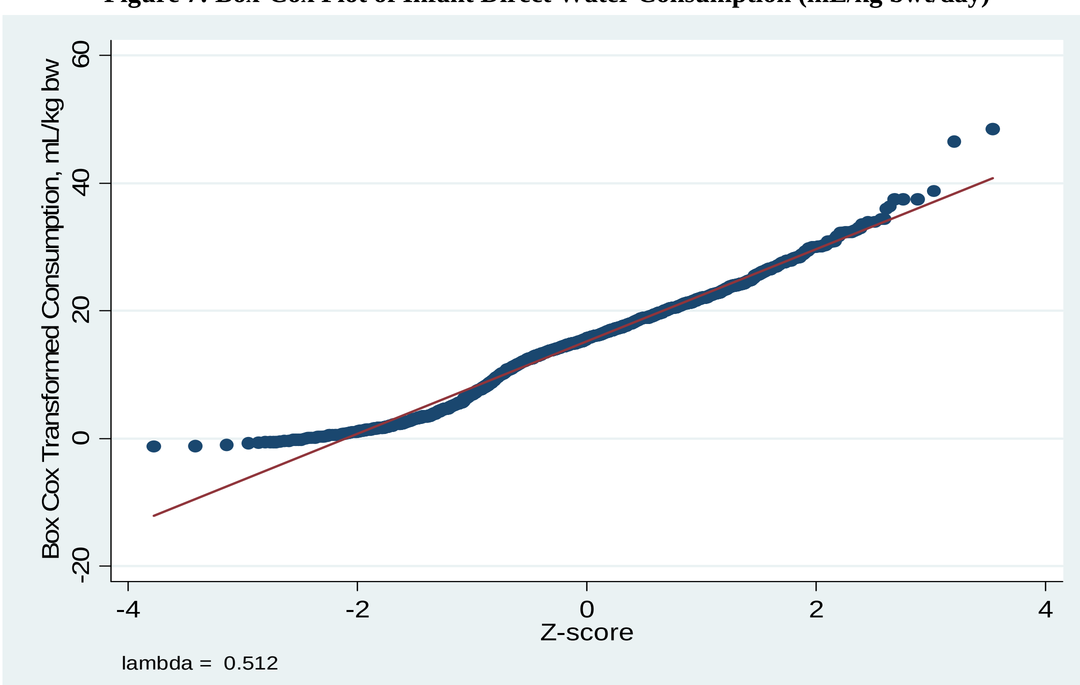
[32] Figure 7 was taken and modified from Lantz et.al. (2006).

Drinking Water Consumption Outliers

As in the case of various food items, there are some high reported drinking water consumption amounts in the CSFII. Figure 7 presents a Box-Cox transformation of drinking water consumption (mL/kg bwt/day) for all infants in the CSFII data base. The two highest amounts are located in the upper right hand corner – deviating above the otherwise linear pattern established by the majority of the remaining reported consumptions**.** These two values are, respectively, 52% and 41% higher on an ml/kg bwt basis than the next (third) highest reported consumption value. An inspection of the food diaries indicate that a set amount of formula was reportedly prepared and consumed by these two infants on multiple occasions throughout the day. The first infant diary (28892-2-1) was for a newborn (0 month old) weighing 3.2 kg that reportedly consumed a total of 1,997 mL that day (1,819 mL indirect, 117 direct), or about 624 mL/kg bwt/day.

An inspection of the CSFII diary indicates that this infant consumed a total of 8 oz of formula (6 ounces consumed directly + 2 oz used to prepare 0.25 cup of dry rice cereal) at 8:00 am, and at 9:30, 11, 1:30, 4:30, 6:00, 10 and 11:30 pm; an additional 4 oz of formula alone was prepared/consumed at 1:00 am. The second infant-dairy (26837-3-2) was a one month old that weighed 3.6 kg, and consuming a total of 2,044 mL that day (1,926 mL indirect, 118 direct), or about 568 mL/kg bwt/day. An inspection of this second diary indicates that that infant consumed 8 oz of formula on nine different occasions throughout the day, at 4:00 am, and at 6, 8, 10, 12, 2, 6, 8 and 10 pm. These two drinking water intake amounts appear to be 'outliers' based on the available references, and a brief review of the pediatric literature (e.g., US EPA, Child-Specific Exposure Factors Handbook, forthcoming).

Figure 7: Box-Cox Plot of Infant Direct Water Consumption (mL/kg bwt/day)



Analyses of the *SHEDS* simulated output for the infant subpopulation indicated that two food diaries constituted about 70% of all high simulated outputs in the top 0.1% of simulated person-days. Again, the question of concern was how sensitive are the estimates at the upper percentiles to the drinking water intakes reported for these two respondents. Agency staff used *cross-sectional SHEDS-dietary* to conduct two 'what-if' scenarios: (1) drop these two diaries from the Monte Carlo simulation, and (2) reduce the reported amounts consumed by 50 percent. As in the sensitivity analyses for the potato eater discussed earlier, and these high infant water intake diaries, the estimated exposures for these high infant water consumers also did not change considerably at the per capita 99.9th percentile in either of these two sensitivity analyses.

The Agency previously noted the robustness of the results to residue outliers:

"*…it is often not the extreme upper tail of a residue distribution which is responsible for driving the 99.9th or 99th percentile exposure levels, but rather a combination of reasonable (but high end) consumption and reasonable (but high end) residue levels of one or two frequently consumed agricultural commodities.*" US EPA (1999), pp. 21-22.

While that quote referred to **residue** 'outliers', the two case studies above suggests that a similar level of robustness appears to hold for consumption 'outliers' as well. While such analyses cannot be performed if the consumption diaries are fixed in the code, the open source code of *cross-sectional SHEDS-dietary* provides Agency modelers with complete access to all of the underlying data and algorithms. This feature enables the Agency to quantitatively address other questions that risk managers may have as PBPK models are used to assess dietary risks to pesticides and other chemicals.

6.2. Uncertainty Analyses

The Agency has not conducted any formal uncertainty analyses for dietary exposure assessments. In a typical risk assessment, all of the (CSFII-FCID) consumption data are used, along with the best available residue data to estimate dietary exposures to pesticides. There is generally little (random seed) uncertainty in the per capita 99.9th when users specify 1000 iterations per diary (results vary by about one percent with different random seeds). This random seed aspect does not capture other aspects of uncertainty regarding the dietary exposure assessment, including: limited food consumption data (CSFII), food recipes (FCID), available residue data (e.g., PDP monitoring, crop translations), and processing factors. This preliminary list of factors expands with longitudinal measures and the use of PBPK models.

SHEDS-Dietary has a simple bootstrapping method for conducting uncertainty analyses - utilizing only a subset of the consumption and residue data inputs (subset of foods). This proposed method was designed to gain some insight about 'How much better would our estimates be if we had **more** data?', by conducting the uncertainty analyses in the other direction 'How far off will our estimates go if we used only a subset of the consumption diaries and foods?'.

The current bootstrap procedure entails the following steps:

1) randomly draw certain percentage (1/20 or 5%) of person-day from CSFII data or/and randomly draw certain percentage (1/4 or 25%) of the commodities having specified residues
2) repeat the step 1 many times, say 200 times
3) get estimated per capita percentiles from each run
4) conduct uncertainty analyses from different runs (e.g. 200 times). 200 50^{th}, 95^{th} and 99^{th} values can be acquired respectively. The ratio of 95^{th} vs. 5^{th} percentile can be used to evaluate the uncertainty
5) obtain important sources contributing to the total uncertainty

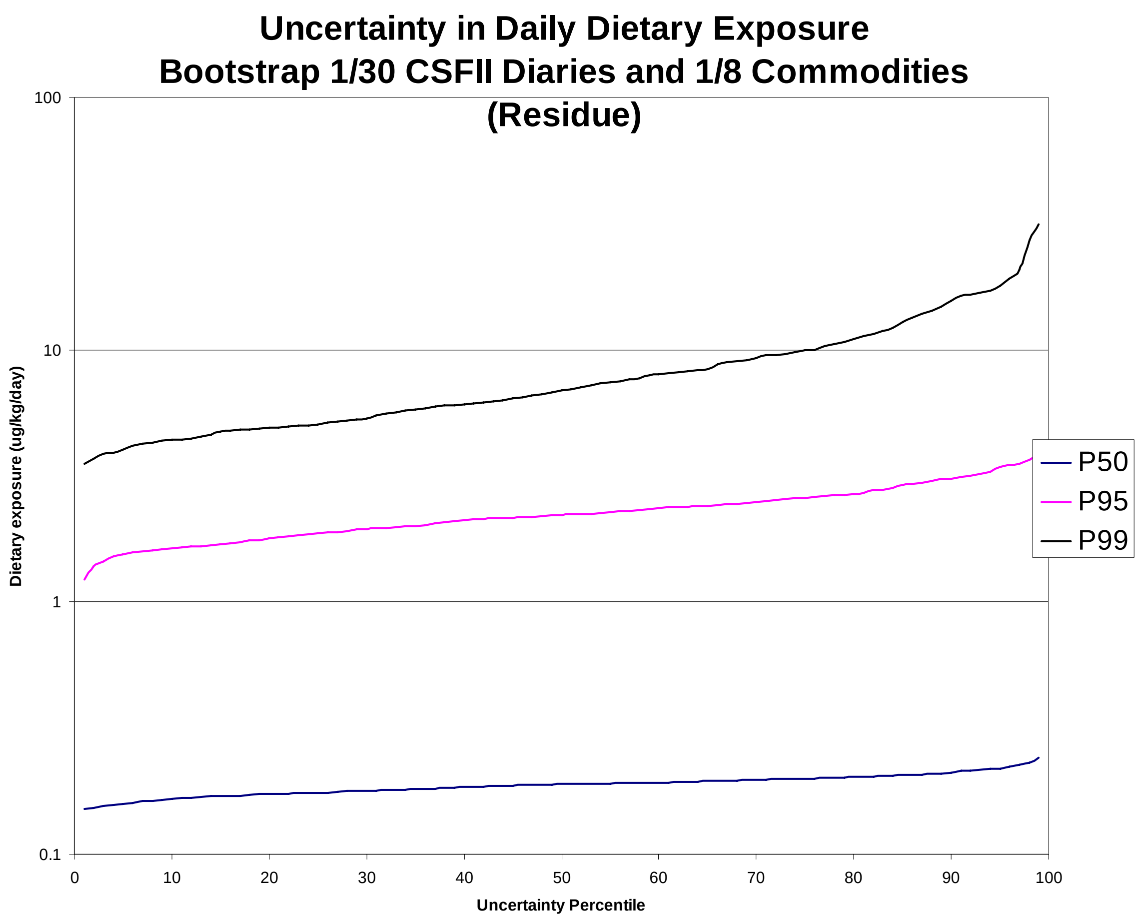


Figure 8: Uncertainty Plot

Figure 8 presents results for the following uncertainty analyses: Bootstrap 1/30 of CSFII diaries 200 times and 1/8 of the commodities (residue distributions). The bootstrap procedure was run 200 times, producing 200 values for each percentile. Figure 8 plots these 200 values for the 50^{th} percentile (P50), 95^{th} percentile (P95) and 99^{th} percentile (P99). For each of these three percentiles, we can calculate the ratio of the 95^{th} percentile to the 5^{th} percentile. Table 9 presents these ratios for this subset and other subsets of food consumption diaries and commodities. For the $1/30^{th}$ CSFII diaries $1/8^{th}$ Commodity (residues) scenario in Figure 8, these ratios are 1.39 (P50 ratio of $95^{th}/5^{th}$), 2.22 (P95 ratio of $95^{th}/5^{th}$) and 4.47 (P99 ratio of $95^{th}/5^{th}$), respectively. The ratios for other bootstrap scenarios in Table 9 indicate that there is greater uncertainty (intervals) at higher percentiles, but that both the consumption data and the commodity (residue) data contribute to overall uncertainty.

These results reflect extreme subsets of the consumption and residue data – as noted, Agency risk assessors use all data (both food diaries and commodities) in dietary risk assessments. The sensitivity of the results to selecting a subset of foods (residues) is somewhat anticipated from past observations that a few commodities generally account for a significant share of the highest modeled exposures. Sensitivity to selecting a subset of diaries (1/8 or less) suggests that having a moderate number of food diaries can be important for estimating exposures at the upper percentiles. While this exercise cannot quantify how much less uncertainty there would be if we had more food consumption data, it provides some information for evaluating the utility of gathering additional food consumption diaries (e.g., Supplemental Children's Survey), or waiting until such data is collected. Similarly, uncertainty analyses on a subset of residues (rather than commodities), might provide some insight into the benefits of gathering more PDP samples for various commodities.

Table 9. Uncertainty Analyses – Subset of Consumption Diaries and Commodities (Residues)

Bootstrap Method	Ratio of Confidence Intervals (95th/5th) for various per capita percentiles		
	50th	95th	99th
Res (1/8) Con (1/30)	1.39	2.22	4.47
Res (1/8) Con (1/20)	1.30	1.99	3.87
Res (1/8) Con (1/10)	1.26	1.69	2.52
Res (1/8) Con (1)	1.20	1.31	1.73
Res (1/4) Con (1/20)	1.24	2.03	3.40
Res (1/4) Con (1/10)	1.20	1.66	2.43
Res (1) Con (1/20)	1.19	1.93	3.28
Res (1) Con (1/8)	1.14	1.51	2.14

How many person-days (or person-years) to simulate?

The aggregate exposure models may conduct between 36.5 million to 146 million person-day simulations in any given simulation.[33] Agency risk assessors typically specify 1,000 iterations per diary during a DEEM-FCID™ simulation, providing for about 41 million person-day simulations (=41,214 person-day diaries x 1000 iterations/diary). Except for extremely unusual circumstances, this number of iterations provides stable results at the per capita 99.9th percentile for all subpopulations (i.e., not much 'simulation' or 'random seed' uncertainty). Similarly, users can specify any number of iterations per diary using *Cross-Sectional SHEDS-Dietary*. The sensitivity analyses presented in this section were based on only 150 iterations which appeared to be sufficient when comparing baseline

[33] Calendex-FCID is typically run with 10 iterations per person-year, providing for a total of 75 million person-days (75,215,550 = 20,607 persons x 365 days/person x 10 iterations/person). Lifeline is often run for 5,000 individuals, providing for a total of 146 million person-days (146,000,000=5,000 persons x 80 yrs/person x 365 days/year). And CARES has a fixed population of 100,000 persons with food match data base of a total 36.5 million person-days (36,500,000 = 100,000 persons x 365 days/person).

results (i.e., no eating occasion) with DEEM-FCID™. We specified fewer iterations (150 versus 1000) since *Cross-Sectional SHEDS-Dietary* retains all of the output from each simulated person-day. While this requires available hard disk space (creating 4 GB in output with 150 iterations), it also enables sensitivity analyses to be conducted more efficiently. For many cases, 150 iterations per diary using cross-Sectional SHEDS-Dietary provide reasonably stable results – as far as the random seed is concerned. This can be verified for other residue inputs with a few simulations. For the SHEDS-Dietary longitudinal model, the Agency will need to conduct further case studies with real residue data before developing guidance on the number of person-years to simulate.

7. Items for Further Investigation

This section provides a list of items for further investigation. We plan to incorporate some of the options presented in this paper in SHEDS-Multimedia Version 4. We currently have two models – *Cross-Sectional SHEDS-Dietary*, and the longitudinal model, SHEDS-Dietary. *Cross-Sectional SHEDS-Dietary* is designed around the CSFII two day respondents, and hence uses the CSFII sampling weights to project exposures at the per capita level. In contrast, SHEDS-Dietary models individuals based on the U.S. Census, and hence the modeled population (percent of race/gender) is representative of the general U.S. population for any age-gender subgroup specified in the simulation. While the longitudinal model may estimate total daily exposures for acute risk assessments, we plan to include *Cross-Sectional SHEDS-Dietary* into SHEDS-Multimedia Version 4 since it provides risk managers with a reference to DEEM-FCID results.

1. SHEDS-Multimedia Version 4 – Dietary Module:

a) Develop Graphical User Interface (GUI) for the Dietary Module
b) Summary Output/Reports
c) Write Dietary Chapters for User Manual, and Tech Manual

2. Top 10 Eaters (Food Consumption Diaries)

Options to allow users to view the food consumption data, and FCID recipes

3. Import\Export Anticipated Residues to Other Models

This will make it feasible to conduct further comparisons with other models.

4. Develop Methods to Model Longitudinal Food Consumption

This will require periodic review as new data are collected and new statistical methods are developed.

5. Develop Option for using DW concentrations (Select Year, Apply Daily Values)

Develop a two step procedure for using predicted drinking water concentrations – randomly selecting a year (1960 – 1990), then applying the daily concentrations to the modeled days, by Calendar day (Jan 1st – Day 1,,, December 31st- Day 365). This approach has the advantage of retaining autocorrelation in predicted drinking water concentrations.

6. Missing Values in the CSFII & Data Imputations

Some alternative approaches for imputing values for missing data will be explored. The two fields of interest are: (i) direct drinking water, and (ii) time of eating occasion. The modeled results appear to be relatively robust with respect to data imputations on these two variables. Approximately 738 diaries, or 1.8% of the total 41,214 CSFII food diaries did not report any information regarding direct drinking water consumption; this is different from a response 'zero ounces consumed'. *Cross-sectional SHEDS* assumed that these individuals did not consume any direct drinking water. For eating occasions, approximately 3,948 records, or 0.6% of the 598,829 food records in the CSFII database had missing values for the time of day question. *Cross-sectional SHEDS* replaced those missing values with 12:00 noon.

7. Inspection of Food Consumption Data for Eating Occasions (residues). Develop decision rule to determine if different residues should be drawn for subsequent eating occasions, and/or foods. The current version of Cross-Sectional SHEDS-Dietary randomly draw 1 residue value for each commodity (RAC-FF) is applied to all foods, on all eating occasions.

8. Preliminary Use of NHANES Food Consumption Data
The Agency has received comments on the general relevancy of the 1994-96, 1998 CSFII data for assessing dietary exposure to pesticides (FIFRA SAP 2005, p.33). *Cross-Sectional SHEDS-Dietary* can be modified to assess the sensitivity of results when using the newer NHANES food consumption data. FCID recipes for new foods need to be developed. The 1994-1996, 1998 CSFII data base included 5,845 food items consumed by respondents. The NHANES (1999-2004) respondents reported consuming many of those same foods, as well as approximately 580 new foods that were not reported during the CSFII survey. The Agency is planning to develop FCID recipes for these new foods. In the interim, the FCID recipes of 'similar' foods may be applied as placeholders. This modification is fairly straightforward, and results between the two data bases can be compared using *Cross-sectional SHEDS-dietary model*. We anticipate more uncertainty using the NHANES data since it has fewer food diaries for children (*e.g.*, CSFII has 2,972 infant diaries versus 1,717 diaries in NHANES; for children age 1-2 years old, CSFII has 4,287 diaries vs. 2,160 diaries in NHANES). Another issue is whether or not to use all food diaries, or only the two day diaries. In contrast to CSFII data, only one day of food intake was collected during the first four years (1999-2002) of the NHANES survey. Therefore, NHANES has a slightly larger total number of one day (only) diaries as it does two day diaries (N=19,344 versus N=16,330). This issue affects *Cross-sectional SHEDS-dietary model* since it will use the NHANES (vs CSFII) sampling weights; whereas SHEDS-Dietary can use all food diaries.

9. Develop Options for Conducting Sensitivity Analyses and Uncertainty Analyses

10. Compare with NHANES Biomonitoring Data

REFERENCES

Barraj, L.M., Daniels, C.L. (2004) "Drinking Water Consumption Survey", September 15, 2004, Sponsored by BayerCropScience, Performing Laboratory, E^x^ponent.®, pp.40.

Dodd, K.W., Guenther, P.M., Freedman, L.S., Subar, A.F., Kipnis, V., Midthune, D., Tooze, J.A., Krebs-Smith, S.M. (2006) "Statistical Methods for Estimating Usual Intake of Nutrients and Foods: A Review of the Theory", Journal of the American Dietetic Association, pp. 1640-1660.

Glen, G., Smith, L., Isaacs, K., McCurdy, T., and Langstaff, J. (2007) "A New Method of Longitudinal Diary Assembly for Human Exposure Modeling." Unpublished monograph.

Glen, G.; Smith, L.; Zartarian, V.; Stallings, C.; Isaacs, K.; Xue, J. (2007) Planned Methodologies for Extending SHEDS-Multimedia Model Version 3 (aggregate) to SHEDS-Multimedia Model Version 4 (cumulative or aggregate), June 12, 2007, p.30.

Lantz, J.L., Young, B.M., Jones, R.L. and Mihlan, G.J. (2006), "Aldicarb Acute Dietary Risk Assessment Including Food and Drinking Water", Volume I, July 14th, 2006, Bayer CropScience.

Lu, C., Toepel, K., Irish, R., Fenske, R.A., Barr, D.B., and Bravo, R. (1996a) "Organic Diets Significantly Lower Children's Dietary Exposure to Organophosphorous Pesticides", Environmental Health Perspectives, February 2006, Volume 114, No. 2, pp. 260-263.

Lu, C., Barr, D.B., Pearson, M., Bartell, S., and Bravo, R. (1996b) "A Longitudinal Approach to Assessing Urban and Suburban Children's Exposure to Pyrethroid Pesticides", Environmental Health Perspectives, September 2006, Volume 114, No. 9, pp. 1419-1423.

National Academy of Sciences (2007) "Models in Environmental Regulatory Decision Making", Committee on Models in the Regulatory Decision Process, National Research Council.

Subar, A.F., Dodd, K.W., Guenther, P.M., Kipnis, V., Midthune, D., McDowell, M., Tooze, J.A., Freeedman, L.S., Krebs-Smith, S.M. (2006) "The Food Propensity Questionnaire: Concept, Development, and Validation for Use as a Covariate in a Model to Estimate Usual Food Intake", Journal of the American Dietetic Association, pp. 1556-1563.

Tooze, J.A., Midthune, D., Dodd, K.W., Freedman, L.S., Krebs-Smith, S.M., Subar, A.F., Guenther, P.M., Carroll, R.J., Kipnis, V. (2006) "A New Statistical Method for Estimating the Usual Intake of Episodically Consumed Foods with Application to Their Distribution", Journal of the American Dietetic Association, pp. 1575-1587.

Tran, N.L., Barraj L., Smith K., Javier A., Burke T.A. (2004) "Combining food frequency and survey data to quantify long-term dietary exposure: a methyl mercury case study.", Risk Analysis, February 2004, 24(1), pp.19-30.

US Department of Agriculture (not dated), Agricultural Marketing Service, "Documentation: Supplemental Children's Survey (1998) to the 1994-96 Continuing Service of Intakes by Individuals", http://www.ars.usda.gov/Services/docs.htm.

FIFRA Science Advisory Panel (1998), SAP Meeting Comments, March 1998. "A Set of Scientific Issues Being Considered by the Agency in Connection with Policy for Review of Monte Carlo Analyses for Dietary and Residential Exposure Scenarios." http://www.epa.gov/scipoly/sap

FIFRA SAP (1999b), SAP Report No. 99-02B, March 25th, 1999, "Session II – A Set of Scientific Issues Being Considered by the Environmental Protection Agency Regarding: Time Sensitive Reversibility of Aldicarb Induced Cholinesterase Inhibition as a Factor in Acute Dietary Risk Assessment", February 23-24, 1999. http://www.epa.gov/scipoly/sap

FIFRA SAP (1999c), SAP Report No. 99-02C, March 25th, 1999, "Session III – A Set of Scientific Issues Being Considered by the Environmental Protection Agency Regarding: Consultation on Development of Aggregate Exposure Assessment Guidance Document for Combining Exposure from Multiple Sources and Routes".

FIFRA SAP (2000), SAP Report No. 2000-01B, May 25, 2000. Session II – A Set of Scientific Issues Being Considered by the Environmental Protection Agency Regarding: Dietary Exposure Evaluation Model (DEEM™) and MaxLIP (Maximum Likelihood Imputation Procedure) Pesticide Residue Decompositing Procedures and Software. http://www.epa.gov/scipoly/sap

FIFRA Science Advisory Panel (2004), SAP Minutes No. 2004-04, "A Model Comparison: Dietary and Aggregate Exposure in Calendex, CARES, and Lifeline", http://www.epa.gov/scipoly/sap

FIFRA Science Advisory Panel (2005), SAP Minutes No. 2005-04, "A Set of Scientific Issues Being Considered by the Environmental Protection Agency Regarding: Prelminary N-Methyl Carbamate Cumulative Risk Assesment", August 23-26, 2005.

US EPA (1999), Exposure Factors Handbook, Volume I. General Factors.

US EPA (2000a), Office of Water, 'Estimated Per Capita Water Ingestion in the United States, Based on Data Collected by the United States Department of Agriculture's 1994-96 Continuing Survey of Intakes by Individuals, EPA 822-R-00-08. Available in the U.S. EPA FCID CD-ROM

US EPA (2000b), Food Consumption Intake by Individuals. Documentation (1/11/200) and raw data files are available on the FCID CD-ROM. See US EPA (2004) website.

US EPA (2000c), "Available Information on Assessing Exposure from Pesticides, A User's Guide", Office of Pesticide Programs, See web link: http://www.epa.gov/fedrgstr/EPA-PEST/2000/July/Day-12/6061.pdf.

US EPA (2000d), "Guidance for Performing Aggregate Exposure and Risk Assessments", FIFRA SAP, February 1st, 1999. http://www.epa.gov/scipoly/sap

US EPA (2000e), "Choosing a Percentile of Acute Dietary Exposure as a Threshold of Regulatory Concern, March 16th, 2000.

US EPA (2003), "Physiologically-Based Pharmacokinetic/ Pharmacodynamic Modeling: Preliminary Evaluation and Case Study for the N-Methyl Carbamate Pesticides", FIFRA SAP, November 10th, 2003.

US EPA (2004) "A Model Comparison: Dietary and Aggregate Exposure in Calendex, CARES, and Lifeline", April 29 - 30, 2004.

US EPA (2005), "A Set of Scientific Issues Being Considered by the Environmental Protection Agency Regarding: Prelminary N-Methyl Carbamate Cumulative Risk Assesment", August 23-26, 2005.

US EPA (2007), Child Specific Exposure Factors Handbook (forthcoming).

Zartarian, V.; Glen, G.; Smith, L.; Xue, J. (2007) SHEDS-Multimedia Model Version 3 – Technical Manual, Draft June 13, 2007, pp. 136.

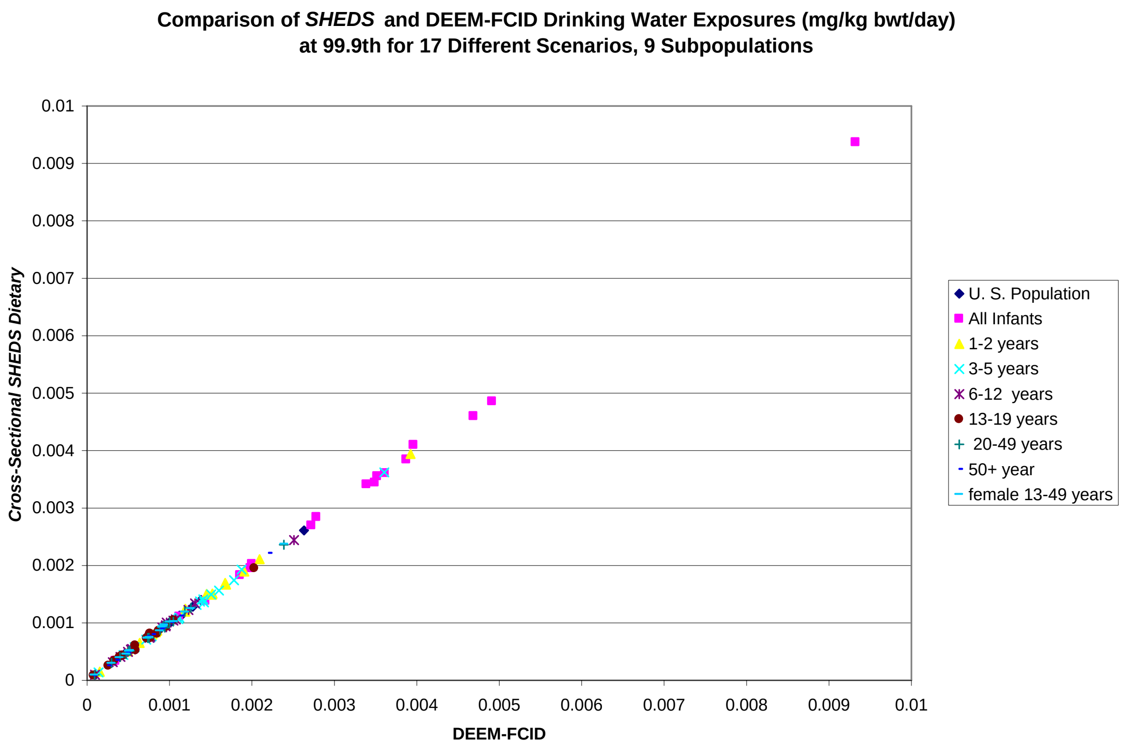
Appendix 1. Comparisons with DEEM-FCID™

Table A.1 presents DEEM-FCID™ and *cross-sectional SHEDS-dietary* estimates of total daily exposure at selected percentiles for chemical ABC, for 9 subpopulation groups. These two models produce similar results across these subpopulations for this particular set of anticipated (food) residues. Children often have higher exposures than adults due to higher intakes of many foods, as a percent of their bodyweight.

Appendix 1 Table 1: Comparison of DEEM-FCID™ and Cross-Sectional SHEDS-Dietary for Chemical ABC

DEEM-FCID™ results (1 simulation w/1000 iterations)			
	95^{th} Pctile	99^{th} Pctile	99.9 Pctile
Subpopulation	**(mg/kg/day)**	**(mg/kg/day)**	**(mg/kg/day)**
U.S. General	0.00209	0.01076	0.04873
All Infants (< 1 yr)	0.00402	0.01661	0.05982
Children 1-2 yrs old	0.00931	0.03261	0.12403
Children 3-5 yrs old	0.00688	0.02717	0.10643
Children 6-12 yrs old	0.00328	0.01515	0.06653
Children 13-19 yrs old	0.00137	0.00762	0.03755
Adults 20-49 yrs old	0.00130	0.00714	0.03410
Adults 50+ yrs	0.00178	0.00879	0.03748
Females 13-49 yrs old	0.00139	0.00792	0.03780
***Cross-Sectional SHEDS-Dietary* results (150 iterations)**			
	95^{th} Pctile	99^{th} Pctile	99.9 Pctile
Subpopulation	**(mg/kg/day)**	**(mg/kg/day)**	**(mg/kg/day)**
U.S. General	0.0021	0.0108	0.0476
All Infants (< 1 yr)	0.0037	0.0158	0.0556
Children 1-2 yrs old	0.0094	0.0326	0.1228
Children 3-5 yrs old	0.0070	0.0272	0.1041
Children 6-12 yrs old	0.0034	0.0154	0.0697
Children 13-19 yrs old	0.0014	0.0078	0.0361
Adults 20-49 yrs old	0.0013	0.0071	0.0323
Adults 50+ yrs	0.0018	0.0086	0.0365
Females 13-49 yrs old	0.0014	0.0079	0.0358
Ratio (DEEM-FCID™/SHEDS)			
	95^{th} Pctile	99^{th} Pctile	99.9 Pctile
Subpopulation	**(mg/kg/day)**	**(mg/kg/day)**	**(mg/kg/day)**
U.S. General	0.99	1.00	1.02
All Infants (< 1 yr)	1.09	1.05	1.07
Children 1-2 yrs old	0.99	1.00	1.01
Children 3-5 yrs old	0.97	0.99	1.02
Children 6-12 yrs old	0.97	0.98	0.95
Children 13-19 yrs old	0.97	0.97	1.03
Adults 20-49 yrs old	0.99	1.00	1.05
Adults 50+ yrs	1.01	1.02	1.02
Females 13-49 yrs old	0.98	0.99	1.05

Figure A.1 plots the DEEM-FCID™ and *cross-sectional SHEDS-dietary* estimates of drinking water exposure at the per capita 99.9th for 17 different drinking water scenarios, for 9 age groups. This plot indicates that these two models produce similar results at this percentile. The infant subpopulation (pink) has the highest drinking water exposures due to relatively higher drinking water intakes per kilogram bodyweight (mL/kg bwt/day).



Appendix 1 Figure 1: Chemical A – 17 DW scenarios, 9 age groups

Appendix 2. General Overview of Probabilistic Risk Assessment Models

Table A.2 Model Comparison Framework

Factor	DEEM-FCID™	Calendex-FCID™	CARES	Lifeline	Cross-Sectional SHEDS-Dietary	SHEDS-Dietary
Reference Population	CSFII Diaries	CSFII Persons	Census (PUMS)	Natality (NCHS)	CSFII Survey Individuals	Census
Model Longitudinal Consumption	**No** Fixed # Iterations per diary	**Yes**	**Yes**	**Yes**	**No** Fixed # Iterations per diary	**Yes**
Food Diary (Binning Approach)	- NA -	Random (2 Day Diary*)	Gower Dissimilarity Index	Random (Age, Season)	- NA -	Random (Age, Gender, DayOfWeek, Season)
Food Consumption (RefPop-Bodyweight)	CSFII normalized	CSFII normalized	CSFII normalized	Lifeline (NHANES)	CSFII normalized	CSFII normalized
Model Weight (per capita)	CSFII	CSFII	CARES (Stratified)	Equal Weights	CSFII	Equal Weights
Food: Anticipated Residues	RAC	RAC	RAC	RAC/Foods (Seasonal)	RAC	RAC/Foods
Drinking Water: Predicted Concentrations	Random	Daily	Daily (multiple sites)	Seasonal CDF	Random	Random

* A recent version of Calendex-FCID (Ver 3.36) contains a Dietary Matching File (DMF) Generator, that allows users to generate consumption profiles by matching CSFII respondents with other respondents based on other demographic characteristics (gender, ethnicity, region, household income, body mass index, breastfed status).

Table A.2 presents a framework for comparing the models. This framework, presented to the Panel in 2004, was developed to help explain why one model might provide higher estimates than another at the per capita 99.9th percentile for any particular set of residues. The report describes the various models in further detail US EPA (2004). The underlying principle is that we can - based on the respective model designs – calculate the expected number of times that a model will use each CSFII food diary, and the model weights that will be applied to each use, from which per capita percentiles are computed. That information is used to construct expected consumption distributions for each commodity (RAC-FF) that can be used to estimate exposures one commodity at a time. If one commodity dominates the tail in a complex assessment, or if residues are specified for only one commodity, then this model should provide extremely accurate predictions for the aggregate probabilistic models. If there are multiple commodities, then a per capita estimate may still be developed by assuming that high exposures for different eaters can be summed across commodities – with the assumption that few individuals receive moderate amounts of exposures from more than one commodity at the tail. These model approximations can provide surprisingly accurate predictions for chemicals in which a relatively high percent of the upper tail obtained most of the total exposures from one commodity.

The upper half of this table lists four design features: (1) Reference Population, (2) Food Diary – Binning Approach, (3) Food Consumption (Reference Population Bodyweight), and (4) Model Weights. The first two items get at the question: 'What is the expected number of times that the model will use each of the food diaries in any given simulation?' This is straightforward for the DEEM-FCID™ and Cross-Sectional SHEDS-Dietary models (user specified parameter); but requires a little work for CARES, Lifeline and the SHEDS-Dietary models. The fourth item gets at the question: 'What weights will the model apply to each simulated person-day to make per capita projections?' This is also straightforward for the DEEM-FCID™ and Cross-Sectional SHEDS-Dietary models (CSFII sampling weights); for CARES, Lifeline and the SHEDS-Dietary models.

The third item pertains to the bodyweight of the modeled individual – and how that information is used with regards to food consumption in terms of grams food/kg bwt – this issue affects the Lifeline model. DEEM, Calendex, CARES and SHEDS-Multimedia 1.0 uses the CSFII respondents' bodyweight when using the CSFII food consumption data (grams food/kg bodyweight/day). In this way, the consumption amounts are 'normalized' based on the CSFII respondents' bodyweights, are used from each food diary. Lifeline uses only the amount of food consumed (grams food/day) from the CSFII diaries, and applies its own anthropomorphic bodyweight model to determine the resulting dietary exposures to these modeled individuals. In this manner, the Lifeline model models a greater range of food consumption (grams food/kg bwt/day) than what is reported in the CSFII diaries.