CHOOSING A PERCENTILE OF ACUTE DIETARY EXPOSURE AS A THRESHOLD OF REGULATORY CONCERN

Office of Pesticide Programs
U.S. Environmental Protection Agency
Washington, D.C. 20460

March 16, 2000

Fax-On-Demand
Fax Number: (202) 401-0527
Item: 6046

List of Acronyms

aPAD	acute Population Adjusted Dose
aRfD	acute Reference Dose
CEC	Critical Exposure Contribution
CSFII	Continuing Survey of Food Intake by Individuals
DEEM	Dietary Exposure Evaluation Model
DRES	Dietary Risk Evaluation System
FFDCA	Federal Food, Drug, and Cosmetic Act
FIFRA	Federal Insecticide, Fungicide, and Rodenticide Act
IWG	Implementation Working Group
NCHS	National Center for Health Statistics
NOAEL	No Observed Adverse Effect Level
LOD	Limit of Detection
LOQ	Limit of Quantitation
OPP	Office of Pesticide Programs
PDP	Pesticide Data Program
RfD	Reference Dose
SCS	Supplemental Children's Survey

Table of Contents

CHOOSING A PERCENTILE OF ACUTE DIETARY EXPOSURE AS A THRESHOLD OF REGULATORY CONCERN

EXECUTIVE SUMMARY

EPA is responsible for regulating the nature and amount of pesticide residues in food under the Federal Food, Drug and Cosmetic Act (FFDCA). FFDCA sec. 408 authorizes EPA to set a tolerance or an exemption from the requirement of a tolerance if the Agency determines that the residues would be "safe." The Agency performs various types of risk assessments to evaluate the safety of pesticides in food, including analyses to determine the nature and the amounts of pesticides that people might be exposed to over a single day. This paper discusses how EPA generally applies the statutory safety standard to acute dietary risk assessments as to pesticide residues in foods.

The Environmental Protection Agency's Office of Pesticide Programs (OPP) previously announced that, on an interim basis, it intended to use the 99.9th percentile of the distribution of estimated acute dietary food exposures for calculating a threshold of concern when probabilistic assessment techniques are used to model the distribution. OPP stated that it would compare this percentile of estimated exposure to the Population Adjusted Dose (PAD), a value that reflects an amount of a pesticide to which a person may safely be exposed in one day. The Agency published a notice in the *Federal Register* on April 7, 1999 citing the availability of an interim policy and requested public comment so that the views of all interested parties would be considered (US EPA, 1999a).

Based in part on the comments received, this science policy paper was revised and is now being issued in its revised format. This revised document explains OPP's policy and details some of the various concerns that have been raised, additional associated public health-related issues, as well as OPP's plans for further evaluation and implementation. This policy has broad applicability to many pesticides.

OPP's current approach with respect to assessing and regulating the food uses of pesticides, when using a probabilistic method of estimating acute dietary exposure, is as follows:

> If the 99.9th percentile of acute exposure from food, as estimated by probabilistic (e.g., Monte Carlo) analysis, is ***equal to or less than*** the acute Population Adjusted Dose (aPAD) for the pesticide, then OPP would generally consider its threshold of concern in applying that the safety standard of FFDCA sec. 408(B)(2)(A) not to be exceeded with respect to acute risk from food. However, if the analysis indicates that estimated exposure at the 99.9th percentile ***exceeds*** the PAD, OPP would generally conduct a sensitivity analysis to determine to what

extent the estimated exposures at the high-end percentiles may be affected by unusually high food consumption or residue values. To the extent that one or a few values from the input data sets seem to "drive" the exposure estimates at the high end of exposure, OPP would consider whether these values are representative and should be used as the primary basis for regulatory decision making. In either scenario, EPA would consider submissions by interested parties that question the appropriateness of the use of the 99.9th percentile in calculating the threshold of concern for the particular risk assessment in question or question its use generally.

It is important to note here that the above position refers to the 99.9th percentile of exposure and not consumption. The 99.9th percentile of exposure represents the joining of each individual's consumption data set with randomly selected residue values from the residue data set. The consumption values associated with the 99.9th percentile of exposure do not necessarily represent the 99.9th percentile of consumption since it is both the selected consumption value and residue concentration which is responsible for determining exposure.

At this time, OPP's current policy is used only with daily exposures to a single chemical through the food pathway only. Estimates of exposure through drinking water and residential uses are not sufficiently developed to warrant inclusion in a probabilistic assessment. Establishing the threshold of concern for the food pathway using the 99.9th percentile of exposure is considered to be a "first step" toward regulation of exposures on an aggregate, and then cumulative, basis.

OPP recognizes that different types of risk assessments will generally be needed for aggregate and cumulative evaluations and that these assessments might also be associated with different regulatory thresholds. Although OPP is moving toward regulating on the basis of probabilistic aggregate and cumulative exposure assessments, a decision has not yet been made regarding how the appropriate threshold of concern should be calculated for these types of assessments. When exposures through drinking water and residential uses are sufficiently refined to be incorporated into probabilistic evaluations, they will be aggregated and assessed, and may use a different population percentile.

Section I of this paper provides an overview of OPP's present practice for acute dietary risk assessment for residues in food. It describes the statutory, regulatory, and policy framework for this policy, as well as prior reviews and comments. In addition, this section provides background information on dietary risk assessment in general and explains how the previous system (DRES--Dietary Risk Evaluation System) and the current system (DEEM--Dietary Exposure Evaluation Model) work, as well as what input data sources are used and how.

Section II addresses some of the specific issues and concerns raised about using exposures at the estimated 99.9th percentile in calculating the threshold of concern. One issue is whether the nature of the databases available (i.e., robustness, adequacy, etc.) should preclude the use of

the estimated 99.9th percentile for regulatory purposes since some consider the uncertainties associated with this population percentile to be too great. Examples of data used are USDA's food consumption survey data, registrant crop field trials, USDA Pesticide Data Program (PDP) data, FDA monitoring data, market basket surveys, etc. Other issues include the treatment of data "outliers," representativeness and adequacy of the databases, and the impact of Agency default values on exposure estimates. Concerns, therefore, exist about whether the estimates of the 99.9th percentile of exposure are sufficiently representative of actual exposure to be meaningful. This paper summarizes these concerns and how OPP has addressed them.

Section III addresses the issue of protectiveness of the estimated 99.9th percentile of exposure with respect to the general public health. One view is that using the estimated 99.9th percentile of exposure is insufficiently conservative because very large numbers of people could be exposed every day to pesticide intakes which are estimated to exceed the Agency's "level of concern." This section also explores the contrary view – that the policy is over-protective because of the conservative assumptions used in the estimation methods and the retention of potentially unrepresentative values in the data base. The section discusses as well the view that, whether it over- or under-estimates actual exposure, the estimated 99.9th percentile of exposure is simply too uncertain to be used in risk management decisions.

Section III also explains that OPP weighs a number of factors in considering which percentile to use: the size of the exposed population and the proportion that might receive daily doses above the benchmark of safety, the acute Population Adjusted Dose (aPAD); the level of confidence OPP has in its exposure estimates; and the extent to which such estimates may overstate potential exposure because they incorporate conservative assumptions or rely on atypical and unrealistic data. Further, to the extent understood, OPP considers by how much individual exposures would be estimated to exceed the aPAD.

Section III also briefly addresses the issues associated with exploratory analysis conducted by OPP with the DEEM software and the 99.9th percentile issue. Further details and specifics of this analysis are provided in the associated response to public comments.

Section IV provides a list of the documents referenced in this paper.

The Appendix, entitled "Primer on Interpretation of Exposure Distribution Curves," is a "plain English" guide to Monte Carlo analysis and interpretation of its results.

I. OPP's Present Practice and Policy for Acute Dietary Risk Assessment

A. Introduction

Under the Federal Food, Drug, and Cosmetic Act (FFDCA), the Environmental Protection Agency may authorize a tolerance or exemption from the requirement of a tolerance, to allow a pesticide residue in food, only if the Agency determines that such residues would be "safe" (FFDCA sec. 408(b)(2)(A)(I)). The term "safe" is defined as a "reasonable certainty that no harm will result from aggregate exposure to the pesticide chemical residue, including dietary exposures and all other exposures for which there is reliable information" (FFDCA sec. 408(b)(2)(A)(ii)).

To determine whether food is safe to eat, OPP must assess the potential risks from pesticide residues in food. The size of the potential risks depends on the toxicity of the pesticide (how much harm, if any, is caused by specific amounts of the pesticide) and the magnitude of the exposure to the pesticide. Exposure to a pesticide in the food supply depends, in turn, on two factors: the amount of the pesticide present in food and how much food a person eats. It is impossible to know precisely how much food every individual in the country consumes, either over a lifetime or even on a single day. Similarly, it is impossible to know how much residue each specific item of food contains. Thus, the Agency must use available and reliable, representative data to develop estimates of such exposure.

In evaluating the potential risks from pesticides in the diet, OPP assesses both chronic (long term) exposure and acute (short term) exposure. For chronic exposure, OPP estimates the average amount of pesticide residue a person might consume over extended periods, potentially ranging from several months to a lifetime. For acute exposure, OPP is instead interested in the amount that might be ingested on a single day. To evaluate acute dietary exposure, OPP now uses a probabilistic exposure modeling technique, an example of which is "Monte Carlo analysis." For the purpose of discussion, this paper will use the term "Monte Carlo" keeping in mind that other probabilistic techniques may be used as well. This probabilistic assessment technique estimates the different levels of exposure people experience as the result of differences in the types and amount of foods they eat, as well as variations in the level of pesticide residue that may be present, among other factors.

Over the last several years, OPP has been working to expand its capability of evaluating acute dietary exposure and risk using probabilistic techniques of assessment. In early 1998, OPP released an interim policy and a series of guiding principles regarding the use of probabilistic risk assessment techniques (U.S. EPA, 1998a). In part, this policy was based on an earlier Agency policy regarding the use of probabilistic techniques in risk assessment. Specifically, in a 1997 memorandum from Deputy Administrator Fred Hansen, EPA stated that probabilistic analysis techniques, "given adequate supporting data and credible assumptions, can be viable statistical tools for analyzing variability and uncertainty in risk assessments" (U.S. EPA, 1997a). The Agency also enumerated a set of conditions to be considered in judging the acceptability of a probabilistic analysis for review and evaluation; these conditions relate to transparency, reproducibility, and the use of sound scientific methods (U.S.EPA, 1997a). This Agency policy document noted that Monte Carlo analysis is the only probabilistic technique that has been accepted so far, but EPA would be open to considering other probabilistic techniques.

Among other things, the 1998 OPP draft guidance document on probabilistic health assessments indicated that, when probabilistic exposure assessments were available for acute dietary risk, OPP would refer to the 99.9th percentile of estimated exposure in making its risk management decisions. In general, OPP would compare this level of exposure to a safety benchmark, i.e., the acute Population Adjusted Dose (aPAD), in determining whether a particular regulatory action would be consistent with the statutory safety standard established by FQPA (U.S. EPA, 1998a).

B. Previous Review of OPP's Interim Policy

In March 1998, OPP brought its interim policy to the FIFRA Scientific Advisory Panel (SAP). The SAP generally agreed with the probabilistic approach proposed by OPP. They considered, among other things, the use of a 99.9th population percentile of exposure and expressed divergent views on whether using the 99.9th percentile is an (adequately) conservative approach. They noted that, in their view, if the 99.9th percentile is utilized, a percentage of the population could still be exposed daily to estimated levels that exceed the regulatory threshold of concern. They further noted that, even though the percentage was small (0.1%), the number of people represented by that percentage was very large because the exposed group is potentially the entire population of the country. The following additional remarks were made by the Panel:

- To judge whether any given percentile criterion is conservative for acute effects or not, it would be necessary to consider the margin of safety which is already incorporated into the toxicological portion of the risk evaluation.

- To identify the level of risk, variability not only in exposure levels but also in human thresholds for the toxic effects under consideration would be needed. That is, a probabilistic "toxicity" component of a risk assessment should be incorporated into the analysis as well.

- By recognizing and separately modeling subpopulations, it may be possible to choose a lower, less statistically tenuous percentile in calculating a threshold of concern for one or more of these subpopulations. This lower percentile may also be warranted, they indicate, if the risk assessment contains a number of "conservative" assumptions that might result in overestimates of risk even at the 99.9th percentile.

The Agency's approach to acute dietary risk assessment has been discussed extensively by the Tolerance Reassessment Advisory Committee. In addition, in January 1998 the FQPA Implementation Working Group (IWG), an informal coalition of agricultural commodity groups and food processing and agricultural chemical trade associations, submitted comments to OPP. In particular, the IWG asserted that the residue and food consumption data sets used as input for probabilistic exposure techniques contained data points which were "outliers" and which made the resulting estimates of exposure distribution appear unrealistically high. IWG also argued that

other conservative assumptions (assumptions that would likely overstate potential exposure) used in developing the exposure estimates made the use of the 99.9th percentile an inappropriate point for use in calculating a threshold of concern.

Following the SAP review and after considering public comment, OPP revised portions of its interim policy document on probabilistic exposure assessment in OPP. This revised document incorporated many of the changes recommended by the SAP in its March 1998 meeting discussed above. On November 5, 1998, EPA announced in the *Federal Register* (63 FR 59780) the availability of the revised document as a draft science policy paper entitled "Guidance for Submission of Probabilistic Human Health Exposure Assessments to the Office of Pesticide Programs" (U.S. EPA, 1998a). As its title indicates, the science policy paper contained guidance on the submission of exposure assessments; it also stated that OPP would separately present an explanation of its policy decision to refer to the 99.9th percentile of estimated acute dietary exposure in making its risk management decisions. The current document addresses this latter issue.

The Agency and the U. S. Department of Agriculture (USDA) have also discussed the OPP policy with respect to determining a level of concern for regulatory decisions about the risks of acute exposure to pesticides in the diet. USDA has commented that the use of data bases which contain too few data points to project high-end percentiles of consumption of a particular food or levels of residues in a specific commodity with statistical confidence raises questions about the estimates of high-end exposure developed using probabilistic assessment techniques. EPA has considered USDA's comments and revised this paper to explain and address its concerns.

C. OPP's Current Approach to Dietary Risk Assessment

1. Chronic vs. Acute Exposure and Risk Assessment

OPP typically performs a dietary exposure assessment for two different exposure time frames -- short term or "acute" exposures and long-term or "chronic" exposures; each assessment is calculated differently. In chronic exposure assessment, the risk assessor is attempting to estimate a person's average dietary exposure over the long-term (e.g., several months to a lifetime). Consequently, the use of both average (or mean) residue value for each food commodity and average (or mean) consumption of food commodities is generally regarded as appropriate. Estimates of exposure through drinking water are subsequently combined with these estimates of exposure through food to calculate combined dietary exposure through food and water. In acute dietary exposure assessment, however, the risk assessor is trying to estimate the

range of exposures that individuals could encounter on a single day and determine the exposure to which "high-end" persons could be subjected (where "high-end" is defined as a plausible estimate of exposure for those individuals at the upper end of the exposure distribution). OPP is using Monte Carlo techniques (and its current 99.9th percentile approach) for these acute food exposure assessments only. OPP is not using Monte Carlo techniques at this time for chronic exposures due to the limitations of the existing food consumption data. EPA and USDA, however, are exploring statistical techniques that may allow such analyses in the future. The Monte Carlo Guidance document provides additional information regarding the tiering process used in acute assessments, for both probabilistic and non-probabilistic assessments (U.S. EPA, 1998a).

Units of Measure

mg/kg bw/day. Milligrams of pesticide per kilogram of body weight per day.

µg. Microgram.

g. Grams

Measures of Toxicity

aRfD. *Acute Reference Dose*. An amount of toxicant (in mg/kg bw/day) to which a person can be safely exposed for one day

NOAEL. *No Observed Adverse Effect Level*. Largest amount of toxicant (again in mg/kg bw/day) in a controlled toxicological study which produces no adverse effect in a test animal

Uncertainty Factor. A series of safety factors by which the NOAEL is reduced to obtain the aRfD. Usually, these consist of an interspecies factor (10x) and an intra-species factor (10x).

2. The Risk Equation

Dietary risk can be expressed as a function of toxicity and exposure.

RISK = f (toxicity, exposure)

That is, to determine risk – which can be either acute (one-day) or chronic (long-term) – one "combines" a value representative of the toxicity for the pesticide with the amount of pesticide to which an individual is exposed. The notation above is not meant to imply that the two quantities represented by "toxicity" and "exposure" are necessarily multiplied together, just that toxicity and exposure are two quantities which together determine risk.[1]

The ***toxicity*** part of the risk function is typically expressed as an acute reference dose (aRfD, in units of mg/kg body weight per day). An aRfD is an amount of toxicant (in mg/kg bw/day) to which a person can be safely exposed for one day. In general, an aRfD is set at a level at least 100 times smaller than the no-observed-adverse-effect level (NOAEL, in units of mg/kg bw/day), if the NOAEL used is from a controlled toxicological study in laboratory animals. The NOAEL is defined as the largest amount of toxicant (in units of mg/kg bw/day) which produces no observed adverse effects. The factor of 100 is a generally applied adjustment, (sometimes called a "safety factor" or, more frequently, an "uncertainty factor") to account for the potential that humans could be more sensitive to the toxic effects of a compound than laboratory test animals (10 X) and that some humans could be more

[1] When toxicity is measured in terms of cancer-causing potential (e.g., in terms of a slope factor as a Q*) then toxicity and exposure are multiplied together. When toxicity is expressed in terms of an RfD, then the reciprocal of the RfD (which is representative of toxicity) is multiplied by the exposure to obtain the estimated risk.

sensitive than others (10 X).

The dietary food ***exposure*** part of the function is derived from two distinct pieces of information: the amount of pesticide residue that is present in and on food (i.e., the residue level) and the types and amounts of food in a person's diet (i.e., food consumption). The residue information comes mainly from the crop field trials submitted by pesticide manufacturers and USDA or from monitoring data collected by the USDA and FDA (see Section I.C.3.(b)). Consumption information comes primarily from USDA surveys of what people eat (see Section I.C.3.(a)).

The basic equations for acute dietary food risk assessment are:

$$Exposure\ (mg/kg\ bw/day) = Consumption\ (kg\ food/kg\ bw/day)\ x\ Residue\ (mg\ pesticide/kg\ food)$$

$$aRfD\ (mg/kg\ bw/day) = \frac{NOAEL\ (mg/kg\ bw/day)}{interspp.\ factor\ (10x)\ x\ intraspp.\ (10x)\ factor}$$

$$Population\ Adjusted\ Dose\ (aPAD)\ (mg/kg\ bw/day) = \frac{aRfD\ (mg/kg\ bw/day)}{any\ additional\ FQPA\text{-}only\ Factor}$$

$$\%\ aPAD = \frac{Exposure\ (mg/kg\ bw/day)}{aPAD\ (mg/kg\ bw/day)}\ x\ 100$$

Once the aRfD is derived, OPP identifies the population for which the assessment is being done. If this population includes the fetus, infants and/or children, a determination concerning retention, reduction, increase, or removal of the FQPA 10X Safety Factor must be made considering potential pre- and post-natal toxicity and the completeness of the exposure and toxicity databases. If a decision is made to retain an additional safety factor for reasons solely based on the FQPA (e.g., increased susceptibility in infants and children), the aRfD would be modified and the resulting allowable exposure is termed the acute Population Adjusted Dose (aPAD).

The value of the "%aPAD" reflects the relative size of the aPAD and the estimated exposure. If the estimated exposure is ***less*** than the aPAD, the value will be below 100%. Conversely, if the exposure is estimated to ***exceed*** the aPAD, the value will be greater than 100%. Traditionally, if the "%aPAD" is less than 100%, the estimated exposure is considered "safe".

3. Data Bases Used in Probabilistic Dietary Exposure Estimates

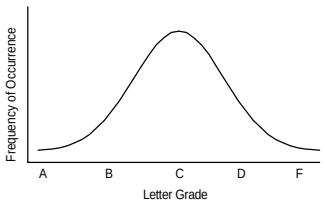
Currently, OPP is developing acute, probabilistic dietary exposure assessments using Monte Carlo techniques that require data on (1) the *distribution* of daily consumption of specific commodities (wheat, corn, apples, etc.) by specific individuals (in g commodity/kg bw/day), and (2) the *distribution* of concentrations of a specific pesticide in those food commodities (in μg pesticide/g commodity). The latter information is generally obtained from crop field trials, USDA's Pesticide Data Program (PDP) data or FDA monitoring data, market basket surveys conducted by the registrants, and other sources while the former is collected by USDA in its Continuing Survey of Food Intake by Individuals (CSFII). These two input data sources, the USDA CSFII and the residue data sources, are discussed below.

(a) *Food Consumption: USDA Continuing Survey of Food Intake by Individuals*

The food survey data used in the OPP's probabilistic exposure and risk assessments are collected by the U.S. Department of Agriculture and are currently from the 1989-91 Continuing Survey of Food Intake by Individuals (CSFII)[2]. The 1989-91 CSFII, conducted as three separate 1-year surveys in 1989, 1990 and 1991, was designed to measure what Americans eat and drink. The USDA has been conducting such food surveys since the 1930's by means of personal interviews in which interviewers ask individuals to recall everything they ate and drank over the previous 24 hours. The uses of Food Survey Research Group survey data are varied and include the assessment of dietary intakes, dietary trends and food consumption economics; the development of policies for food assistance, food labeling and food safety programs; and the implementation of dietary guidance and nutrition education programs. Information from the surveys also is widely used across the U.S. to develop nutrition and education programs, to assess dietary changes associated with participation in food programs, to develop food fortification and enrichment policies, to monitor the safety of the food supply, and to assess demand for agricultural products and marketing facilities. In

What Does the Term *Distribution* Mean?

Think back to the classic bell curve we learned about at some point in our school days. When grades were being determined, some of us had scores that were either on the low -end or high-end of the range while most of us had scores in the middle. If the frequency of occurrence were plotted, the resulting distribution of grades would resemble the bell curve: Essentially what the bell curve tells



us is that most things are near the middle – there are far fewer occurrences at the extremes.

The distribution of the pesticides residues does not typically follow this bell-curve shape. Instead, the curve is generally right-skewed with a long tail to the right. In the above diagram, it would be as if the vast majority of grades were A's and A-' s with very few C's, D's, and F's.

For an additional explanation of how this exposure distribution curve is interpreted in risk-based decision-making, see the Appendix to this document

[2] Data from the recently completed 1994-96 CSFII have now been released by USDA and are expected to be incorporated into the OPP's risk assessments beginning in Spring, 2000.

accordance with federal data reporting guidelines, USDA identifies and cautions users of its databases about the lack of adequate numbers of data points for certain statistical projections. For example, some of the commodities for which EPA sets tolerances are eaten so infrequently that USDA cautions against using the survey data to estimate high-end percentiles of consumption of such commodities, e.g., the 95th percentile or greater.

CSFII (1989-91) data are derived from information provided by 15,128 individuals who participated in the survey. One-day food and nutrient intake data for individuals of all ages were collected between April 1989 and March 1992. Individuals who took part in the survey were asked to provide three consecutive days of dietary data. The first day's data were collected in a personal in-home interview using a 1-day dietary recall. The second and third days' data were collected using a self-administered 2-day dietary record. Intake amounts were reported and energy and nutrient intakes were calculated using the USDA Nutrient Data Base for Individual Intake Surveys. Subject to the cautions about statistical treatment of data, the data collected for such large numbers of survey participants, who have been scientifically selected so that results could be projected from the sample to the U.S. population, constitute a reliable and representative national sample.

(b) *Residue Data Sources: Field Trials, Monitoring, and Market Basket Surveys*

In addition to the food consumption data provided by USDA's CSFII, information on the distribution of residue levels in foods is necessary in order to calculate exposure and risk in a probabilistic manner. Data on the distribution of residues on foods for use in OPP's probabilistic exposure and risk assessments can be obtained from a variety of sources including: (1) crop field trials; (2) FDA enforcement monitoring; (3) USDA PDP monitoring; (4) specialized market basket surveys (usually conducted by a pesticide registrant); and (5) studies on the effects of commercial processing, peeling, washing, cooking or other activities that may affect residue levels. Crop field trials are experimental trials, usually performed by a pesticide company or USDA, in which the maximum usage scenario (with respect to application rate, number of applications, pre-harvest interval, etc.) is simulated. These OPP-required experimental trials are conducted according to Agency guidelines, primarily to determine maximum residues that may be present in fruit, vegetable, grain and other food and feed crops at the earliest point where these food commodities could enter commerce. These data are used to establish legally enforceable pesticide tolerance limits.

In contrast to the pesticide residue data collected during the experimental field trials, FDA and USDA pesticide monitoring data (as well as registrant-sponsored, market basket survey data) represent residue data in crops collected from commercial trading channels (wholesalers, warehouses, distribution centers, retailers, etc.). These data better represent pesticide residues to which consumers are actually exposed because they measure residues in food in commercial channels that is closer to the consumer than food sampled following experimental field trials conducted under maximum application scenarios.

OPP prefers to use data from FDA or USDA PDP monitoring data or market basket surveys, when available, in calculating pesticide exposure estimates. However, these data are not always available or appropriate for use; when this is the case, OPP uses pesticide residue data collected from the experimental field trials. As the field trial data represent residues resulting from a maximum application scenario to which only very few crops are actually subjected, OPP may refine these data to take into account other factors such as residue degradation as a result of transport or storage, or variabilities in farming practices such as use of longer than label pre-harvest intervals and lower than label application rates (U.S. EPA, 1999b). In addition, OPP's exposure estimates can be modified or adjusted, as appropriate, to take into account decreasing or increasing concentrations in processed commodities as a result of commercial processing practices or (generally) decreased residues as a result of cooking or in-home preparation such as washing, peeling, coring etc. (U.S. EPA, 1999c) Finally, information on the percent of the crop which is treated, if available, is also used to adjust the probability of encountering a treated commodity.

4. DRES and DEEM

Until 1998, OPP used a software program called the Dietary Risk Evaluation System (DRES) to conduct its acute dietary risk assessments for pesticide residues in foods. Acute assessments conducted with DRES assumed that 100% of a given crop with registered uses of a pesticide was treated with that pesticide and that all such treated crop items contained pesticide residues at the maximum legal (tolerance) level. The resulting DRES acute risk estimates were considered "high-end" or "bounding" estimates.[3] However, it was not possible to know where the pesticide exposure estimates from the DRES software fit in the overall *distribution* of exposures due to the limits of the tools being used. Thus, risk management decisions were being made not only without a full picture of the distribution of risk among the population, but also without full knowledge of where in the distribution of risk the DRES risk estimate lay.

OPP is now using the Dietary Exposure Evaluation Model (DEEM) computer software program for its dietary exposure and risk assessments for pesticide residues in food. Like the DRES model, DEEM calculates acute and chronic risk using the inputs of: pesticide residues in and on food, food consumption, and toxicity. Also, like DRES, DEEM is able to calculate an estimate of the risk to the general U.S. population in addition to 26 population subgroups, including five subgroups for infants and children:

• U.S. population	• children 7-12	• Females 13+, pregnant/not nursing

[3] A "high-end" estimate is conceptually one that falls between the 90th percentile of the actual exposure distribution but below the exposure to the person in the population who has the highest exposure. It is a plausible estimate of the individual exposure for those persons at the upper end of the exposure distribution. A "bounding estimate," on the other hand, purposely overestimates the exposure or dose in an actual population for the purpose of developing a statement that the risk is "not greater than..." (U.S. EPA, 1992a. Guidance on Risk Characterization for Risk Managers and Risk Assessors).

• U.S. population -spring	• Hispanics	• Females 13+ nursing
• U.S. population--summer	• Non-Hispanic Whites	• Males, 13-19
• U.S. population--autumn	• Non-Hispanic Blacks	• Males 20+
• U.S. population--winter	• Non-Hispanic (other than Black or White)	• Seniors 55+
• all infants	• Females 13-19, not pregnant or nursing	• Northeast
• nursing infants (<1 yr)	• Females 20+, not pregnant or nursing	• Midwest
• non-nursing infants (<1 year)	• Females 13-50	• South
• children 1-6	• Pacific	• West

Unlike DRES, DEEM can generate probabilistic assessments of acute dietary food exposure. DEEM uses a mathematical technique called Monte Carlo analysis to generate estimates of the *distribution* of pesticide dietary exposures. That is, it uses all the individual food consumption and pesticide residue level data points included in a data set to determine the combined (or joint) distribution of exposures (and associated risk). For more information on the interpretation of exposure distribution curves generated by the DEEM software, see Appendix I - "Primer on Interpretation of Exposure Distribution Curves." At this time, OPP uses the DEEM probabilistic (Monte Carlo) model to develop probabilistic exposure estimates only for acute assessments.

The Monte Carlo technique provides a relatively new tool in exposure assessment for more accurately estimating the complete distribution of exposures, and provides probabilistic and statistical assessment of dietary risk using more refined information than was used previously. This analysis uses the actual distribution of pesticide residue levels from either the experimental field trials performed by the registrant or monitoring or market basket surveys, whereas in DRES only a single, high-end residue value was used. Also, it can incorporate information on the percentage of the crop which is treated. That is, it includes the actual distribution of possible consumption and residue values and weighs these possible values by their *probability* of occurrence. Using Monte Carlo, OPP does not assume (as was previously done with DRES) that 100% of the crops with registered uses are treated with the pesticide of interest or that all residues are present in crops at maximum legal (tolerance) levels. Rather than the crude "high-end," single point estimates provided by DRES, Monte Carlo provides more accurate information on the range and probability of possible exposure and their associated risk values.

Monte Carlo techniques are, in and of themselves, neither more conservative nor less conservative than the DRES system they supplement: the "conservatism" is determined, in part, by the risk manager when he or she determines the appropriate percentile of the model's output

distribution (e.g., 99.9th percentile) to be used for regulation in conjunction with the nature of inputs selected and assumptions used. Monte Carlo and probabilistic techniques are simply tools that potentially allow the risk assessor and manager to see a more accurate distribution of risks among the general population and subpopulations.

5. DRES 95th Percentile Dietary Food Exposure vs. Monte Carlo 99.9th Percentile Dietary Food Exposure

The Agency has in the past used the estimated 95th percentile of exposure in calculating a threshold of concern with an acute DRES analysis. Concerns have been raised about what is seen by some as a significant "raising of the bar" by now choosing to refer to the estimated 99.9th percentile of exposure from a Monte Carlo analysis. While it may appear at first that the Agency is taking a more stringent approach, this is actually not so. Estimated exposure at the 99.9th percentile calculated by DEEM probabilistic techniques is significantly lower than exposure calculated by DEEM using DRES-type non-probabilistic assumptions at the 95th percentile for most cases reviewed by OPP to date. There are several reasons for this. An acute DRES analysis assumes that residues are present at tolerance levels in all crops that have registered uses and 100% of the crop is treated at maximum label rates and harvested at minimum pre-harvest interval.

.

In general, Monte Carlo techniques will provide lower and more realistic estimates of exposure than previous DRES techniques when:

- a lower percentage of the crop is treated (e.g., 10% rather than 100%);
- The bulk of residue values from crop field trials are present at low levels and there are only a few high values; and
- a greater number of crops are registered (e.g., 10 crops instead of 2 crops).

For example, a given food item (e.g., cherries) can have several dozen or more individual residue values generated from experimental field trials for a certain pesticide. In an acute DRES analysis, only the highest residue value (or tolerance level) would be used and all registered crops would be assumed to be treated and contain these high residue values. In a Monte Carlo run, the entire set of actual residue data points generated in the crop field trials and the percent of the crop which was treated would be considered. The differences between the estimated exposure numbers generated by these two techniques can be substantial, with the Monte Carlo generated estimated exposures (at the 99.9th percentile) frequently many times lower than DRES-like estimated exposures (at the 95th percentile). Table 1 illustrates some of these extensive differences in exposure estimates for a widely used agricultural pesticide which was recently evaluated by OPP:

TABLE 1. Comparison of DEEM 95th Percentile Exposure and %aRfD Estimates From a Tier 1 Analysis to Monte Carlo 99.9th Percentile Exposure and %aRfD Estimates from a Tier 3 Analysis for One-Widely Used Agricultural Pesticide (expressed on a per capita basis using 1989-91 CSFII Data)

Population Subgroup	Exposure (mg/kg bw/day)		%aRfD[a]	
	DEEM 95th Percentile Estimate (Tier 1)	DEEM Monte Carlo 99.9th Percentile Estimate (Tier 3)	DEEM 95th Percentile Estimate (Tier 1)	DEEM Monte Carlo 99.9th Percentile Estimate (Tier 3)
U.S. Population	0.0192	0.0013	770	50
Infants	0.0375	0.0007	1500	38
Children 1-6	0.0402	0.0017	1610	67
Females 20+/np/nn[b]	0.0126	0.0011	510	45
Males 20+	0.0119	0.0014	480	55

[a] The %aRfD represents the portion of the acute "risk cup" which is occupied. The %aRfD is obtained by dividing the estimated exposure at any given percentile (e.g., 95th or 99.9th percentile) by the aRfD. It should be remembered that the aRfD may be modified to reflect the decision with regard to the FQPA 10x Safety Factor. This modification results in an acute Population Adjusted Dose (aPAD). Comparison of the estimated exposure to the resulting aPAD is then done to determine the acceptability of that exposure.

[b] Females 20+, not pregnant, not nursing

As can be seen, estimated exposures (and corresponding %aRfDs) are significantly lower at the 99.9th percentile DEEM/Monte Carlo analysis than they are at the 95th percentile using a DRES-like analysis. This is almost invariably the case. In fact, at all comparable percentiles, the exposure estimates derived from DEEM/Monte Carlo are lower than the corresponding DRES-like estimates. The advantage of this probabilistic technique is that it can refine the exposure and risk estimates by more fully incorporating all available information and minimizing reliance on values chosen more for their regulatory and administrative convenience than their scientific merit.

In short, DEEM/Monte Carlo analysis tends to provide a lower (but more reliable) estimate of actual exposure in exactly those situations where DRES (or DRES-like analysis) is least realistic. OPP will continue to use the estimated 95th percentile of exposure in calculating a threshold of concern when actual tolerance levels and 100% crop treated assumptions are used during exposure assessment, but recognizes that this approach can significantly overestimate actual exposure levels. In those cases where exposure estimates at the 95th percentile using less refined assumptions are greater than the regulatory threshold of concern, OPP's policy is to use Monte Carlo techniques to assess estimated exposure at the 99.9th percentile using more refined data. In practice, risk assessments done at the estimated 99.9th percentile using more refined data almost invariably result in lower estimated exposures (and corresponding estimated risk) than assessments performed at the estimated 95th percentile of exposure using less refined data.

II. Issues Related to the Methodology and Data Bases Used in Acute Dietary Risk Assessment

Concerns have been raised among the academic, public health, industry, and grower communities with regard to the appropriateness of the estimated 99.9th percentile of exposure as the default decision point for regulation when using probabilistic techniques for acute dietary risk assessment. Specifically, these concerns include: the presence of "outliers" in the pesticide residue and food consumption data; the representativeness of the data sets used in Agency risk assessments; the limited size of the input data bases; the reliance on "uncertain" consumption values which fall at the extreme tails of the distribution when generating exposure estimates; potential variability/uncertainty in the recipes used to convert the reported food items (on an "as consumed" basis) to agricultural commodities used in the DEEM software; and the degree to which OPP's 99.9th percentile estimate incorporates conservative default assumptions. Because of these areas of concern, issues have also been raised about the interpretation of the output developed by the Monte Carlo technique. Some contend that, if the input data are not reliable and representative, neither are the outputs of any technique using such data. Therefore, they contend that the Agency should not use the 99.9th percentile of estimated exposure as a starting point for regulatory decision making and/or should make adjustments in the data sets which are inputs to the exposure assessment. Specific concerns addressed in this section include the following:

- Inclusion of high-end values from USDA food consumption input data sets (see section II.A.);
- OPP use of upper-end residue data (see section II.B.);
- OPP consideration of the size and representativeness of the data bases for consumption estimates and residue profiles (see sections II.C. and D.);
- Potential inaccuracies in recipe translations (see section II.E.); and
- OPP use of conservative default assumptions in treatment of pesticide residue data (see section II.F.).

What's An Outlier?

In a data set, an *outlier* is a number that greatly differs (or is substantially removed) from the bulk of a data set. That is, it is a value that is much larger or much smaller than most of the other numbers in the set. It does not necessarily represent an invalid data point, but may simply represent an unusual or rare, but still very real, occurrence.

A. Treatment of High-End Consumption Values ("Outliers") in USDA CSFII Survey

Concern has been expressed that the USDA's food consumption data have not been properly evaluated to identify potential errors in the

data sets or to assess the potential impacts of outliers on the estimated 99.9th percentile of exposure. As a result, some contend that errors are propagated throughout OPP's Monte Carlo analysis resulting in distributions which inappropriately and artificially inflate estimates of risks at the upper ends of the distribution. Consequently, some believe that tests for outliers should be conducted, and "outliers" should be removed from the data set, so that the high end of estimated risk is not defined by the outliers. They state that OPP's failure to do this means that the results may not be reliable or scientifically-based.

The Agency shares this basic concern and does not want to use data which are not reliable measures of food consumption. By the same token, OPP does not want to ignore data which measure real, but perhaps relatively infrequent, consumption events. Anytime a survey as extensive as USDA's CSFII is conducted, high consumers of a particular food item will be found and reported. It is also true that some portion of the food consumption reports in the initial database may be erroneous. Thus, outliers may be present in the raw survey data. Given that very high food energy intakes do occur in the American population (even though they are not common), considerable judgment is required to determine whether a high-end value should be declared in error and discarded or should be retained.

To ensure that the 1994-96 CSFII data base is free of erroneous or unreliable data points, the USDA extensively validated and cross-checked any questionable survey results prior to their insertion into the 1994-96 CSFII database. For example, USDA survey interviewers were trained to probe for additional information when unusual intakes of various kinds are reported, and to ask questions clarifying large reported amounts, and also if the day's intake was typical or not. If the reported intake was not typical, queries were made about what was atypical, such as the occurrence of a holiday, a social occasion, or the like. On preliminary review of survey data, USDA identified high intakes in various areas (i.e., high consumption of certain foods or high energy intake) and evaluated the reported intake for feasibility, including notations made by the survey interviewer relating to the perceived validity of the reported consumption. USDA carefully reviewed all of the data resulting from its 1994-96 CSFII survey to assure that the reported results are as close to reality as possible. USDA paid particular attention to data points that may be "outliers," and that may have been needed to be removed to characterize food consumption accurately. All reported high intake values retained in the 1994-96 CSFII database have been checked by USDA and resolution or adjudication of values outside specified ranges have been accomplished. Thus, the USDA 1994-96 CSFII data base has been properly evaluated and contains accurate and reliable consumption values that OPP will use in assessment of human dietary exposure to pesticide residues. OPP has similar confidence in the 1989-91 CSFII survey data.

OPP does, however, recognize that unusually high intakes can potentially "drive" calculated exposure and risk estimates and believes that it may be inappropriate to base risk management decisions on unusual consumption values, particularly if these consumption values dominate high-end exposure estimates. Therefore, OPP has decided that risk characterizations will include a "sensitivity analysis" which will take advantage of a recent upgrade to the DEEM

software program which is now capable of generating a "Critical Exposure Contribution" (CEC) analysis when run in the acute Monte Carlo mode. The CEC provides insights into the sources contributing to the exposure estimated for the most highly exposed people in the exposure distribution. This listing contains a detailed exposure analysis for individuals having a total exposure greater than a user-specified "CEC exposure value" (at present, typically around the 99.9th percentile of exposure) in the user distribution profile. The display includes key demographic information (gender, age, body weight), the food(s) consumed, amount consumed, the residue value, the total daily exposure estimate, and the exposure estimate by food. Thus, the CEC provides OPP with comprehensive information on foods (and food-forms) that account for the largest portion of the person's estimated exposure. If OPP finds that the high-end exposures are principally driven by suspect high-end consumption values, OPP's risk mitigation decisions can appropriately consider and weigh these factors.

B. Treatment of High-End Residue Values ("Outliers") in Crop Field Trial or Monitoring Data

As with the food consumption data, some have stated that the residue data included in Monte Carlo assessments have not been properly evaluated to identify potential errors in the data sets or to assess the potential impacts of these outliers on the estimate of the 99.9th percentile risk. OPP acknowledges that it is not uncommon, when field trial residues comprise the data sets used in a probabilistic assessment, that these data include one or more residue values which are significantly higher than the other measured concentrations.[4] Just as with food consumption data, it is important to assure that these data are as accurate as possible. Retaining an erroneous high-end value may result in overestimating exposure, but discarding accurate high-end values may lead to an underestimate of exposure.

Even though OPP may previously have reviewed and relied on a data set, each pesticide residue point in the residue data sets included as input to any Monte Carlo analysis is carefully reviewed and verified by OPP staff scientists because of the recognized potential impacts outliers could have on the high-end exposure estimates. OPP's longstanding approach to outliers has been articulated in the recent draft document "Guidance for Submission of Probabilistic Exposure Assessments to the Office of Pesticide Programs" (U.S. EPA, 1998a) The decision to discard an outlier is based on a scientific or quality assurance basis, and is only made with extreme caution, particularly for environmental data sets which can often contain legitimate extreme values. OPP believes that statistical tests can be used to identify possible outlier data points which require further investigation, but that it is inappropriate to eliminate outliers from analysis on this basis alone unless further review of the suspect data points reveals a significant mistake in protocol which renders a generated residue value irrelevant to label conditions (e.g., wrong tank mix

[4] Frequently, high-end field trial values are the same data initially provided to the Agency by a pesticide company to support OPP's original decision to allow marketing for its product. In fact, these high end residue values likely were used in establishing EPA tolerances. Occasionally, these outliers have represented a sizable fraction of the submitted data sets.

concentration, mistaken application rate, too early a PHI, too many applications, etc.) or there is some other basis to conclude that the data point is not appropriate for use. This is particularly true in cases where the data points in question have been used by the Agency in establishing a tolerance or other regulatory limit.

Occasionally, high-end values may be found among the data from the USDA or FDA monitoring programs. The Agency relies on the extensive QA/QC procedures followed by USDA's PDP program and the FDA program to determine which data points should be retained. Therefore, OPP normally does not discard any of these values, absent other evidence of their invalidity. For example, the monitoring data provided by USDA's Pesticide Data Program are collected under rigorous QA/QC procedures which include method validation (determination of limit of detection and limit of quantitation for each pesticide/crop combination), confirmation of residue identity by alternate detection system, use of blanks, spikes, and internal and external standards, as well as verification of the analyst's performance (check samples, audits, etc.). Similarly, FDA uses official analytical methods that include blanks and fortifications, and requires confirmation of residues of regulatory significance by use of an alternate detection system and verification of results by a different analysis. These measures are intended to ensure integrity of monitoring data from the sample collection to data reporting.

As with high-end consumption values from the USDA food consumption survey, OPP scientists will, as part of the risk characterization, inform the risk manager if high-end residue values are driving the upper ends of the exposure using DEEM's CEC analysis. If specific pesticide residues on certain crops substantially contribute to the majority of the exposures above the specified percentile, the risk manager can incorporate this information into the risk decision and determine an appropriate Agency response.

C. Representativeness

OPP combines both residue data (from crop field trials, monitoring programs, and market basket surveys) and food consumption data (from USDA CSFII surveys), using the DEEM computer software program to generate estimates of the distribution of daily exposures to pesticide residues in food for the general U. S. population and 26 specific subgroups within the U. S. population. The reliability of the estimate of the distribution of exposures depends on the quality of the data used in the model. The data sets used should be sufficiently representative to support reliable estimates.

The relationship between representativeness and the reliability of estimates of the distribution of exposure is easy to understand. Even if all of the values in a data base are accurate (see discussion in II A. and II B.), the use of a data set in a probabilistic assessment will produce unreliable exposure estimates to the extent that the data sample is unrepresentative of the larger population it purports to represent. For example, if no one from low income groups were interviewed about their eating habits, the survey results would miss the very real impact that income has on dietary choices.

The food consumption data used by OPP are collected by the USDA through a survey that is carefully designed to assure that the results would be representative of the U. S. population. The survey design specifically requires that samples be collected from people who differ in ways that could affect the types and amounts of foods they eat. For example, the survey covers people of different ages, genders, ethnicity, regions of the country, and socioeconomic status. People who are selected for interviews are contacted on different days of the week, scattered throughout the year to capture differences due to the time of year or day of the week. A number of other aspects of the survey are also controlled in order to maximize the prospect that the results are representative not only of the entire U. S. population, but also particular subgroups, including those for which OPP generates acute dietary food exposure distributions.

While the USDA food consumption surveys are designed to be generally representative of the U. S. population, it is clear that some factors that can influence dietary choices are not addressed in the survey design. For example, the CSFII surveys do not purport to be representative of people in institutional living arrangements (colleges, nursing homes, etc.) or of different religions or health status. In addition, concern has been expressed about how "representative" the survey results are at the high ends of consumption. This concern, in effect, involves the size of the food consumption data bases. The Agency addresses this concern in Section II.D. of this issue paper and presents a summary of the results of further analysis of the methodology in Section III.

The various data bases on pesticide residues in food raise different set of issues with respect to "representativeness." If market basket or monitoring data are not available, OPP will use residue data sets generated by the registrants and submitted to the Agency for tolerance-setting purposes. The field trial studies are designed to follow the directions on the product labeling and are required to be performed in different areas of the country where the crop on which the pesticide is being used is grown. Multiple field trial sites are required if the crop is a significant component of the diet (e.g. wheat, corn, tomatoes, etc.) and if it is grown in geographically and climatically distinct regions. Of necessity, these are conducted at maximum label rates and minimum label pre-harvest intervals in order to establish maximum legal residue limits (tolerances) on food. Thus, the data set resulting from the required field trials represents the distribution of residues that are likely to be found in a particular raw agricultural commodity following a maximum label application scenario. But, due to the design of these field trial studies, the data are not likely to be representative of the residue values in food, as consumed. As discussed in section I.C.3., adjustments can be made to these data to better represent the amount of pesticides actually used (incorporating, for example, the range of typical application rates and typical PHI's and percent crop treated). Further adjustments and refinements can also be made to better reflect actual exposures; these can include cooking studies, residue degradation studies, washing/home processing studies, etc.

An even more representative picture of the amounts of pesticides in food to which the U.S. population is exposed can be obtained when OPP uses data from market basket surveys or from USDA PDP or FDA monitoring. These data sources are considered to be more

"representative" of actual exposure to consumers than field trials conducted under conditions using maximum rate, minimum PHI, and other use conditions likely to lead to the highest lawful residue. Market basket surveys, for example, are statistically designed and are conducted on a single-serving basis at the point of sale to consumer. These types of studies, thus, best reflect those residues to which consumers are actually exposed. The USDA, too, exercises great care to assure that the food items sampled in its PDP program are representative of the large majority of that type of agricultural commodity consumed in the country. These monitoring data are also designed to be statistically representative of commodities which are typically available throughout the year, except that they represent five pound composite samples (and not single-serving items) collected at distribution points just before release to supermarkets and grocery stores. In addition, PDP commodities are washed, peeled, de-stemmed, or cored, as appropriate, prior to laboratory analysis to represent typical consumer practices. FDA surveillance monitoring data are geared more to tolerance enforcement and not toward OPP's risk assessment needs. Collection occurs as close to the farm gate as possible and the program is not designed to generate statistically representative samples for use in risk assessments. Due to sampling and collection methodologies, residues measured under the FDA surveillance monitoring program likely overestimate pesticide residues to which consumers are exposed. Nevertheless, they are considered more representative of residue levels to which consumers are exposed than the experimental field trial data submitted for tolerance-setting purposes.

D. Size of the Input Data Bases.

In addition to being accurate and representative, the data sets should also be sufficiently large to permit characterization of the overall exposure of the population of interest. As noted earlier, each person in the USDA's 1989-91 CSFII survey currently being used by OPP was asked to contribute information for three different days. While not every respondent provided the full three days of responses, the 1989-91 USDA food consumption survey includes full (three-day) data from 11,912 individuals. The CSFII data base used in the DEEM software collectively represents food records for a total of 35,736 unique person days. Moreover, since each person typically eats many different types of food during a day, there are a great many data points for consumption of specific foods in the data base. Table 2 summarizes the same information for several of the 27 populations that OPP evaluates in its probabilistic exposure assessments.

TABLE 2. Size of 1989-91 and 1994-96 CSFII (with SCS) Data Base

Size of 1989 - 91 CSFII Data Base (all three days)			Size of 1994-96 CSFII Database with Supplemental Children's Survey (both days)	
Population Group	Number of Respondents	Number of Intakes	Number of Respondents	Number of Intakes

U.S. Population	11,912	35,736	20,607	41,214
Infants <1	202	606	1,486	2,972
Children 1-5	1,067	3,201	6,487	12,974
Children 6-11	1,172	3,516	1,913	3,826
Females 12-49	3,459	10,377	3,021	6,042
Males 20+	3,381	10,143	4,751	9,502
Note: OPP is currently using the 1989-91 CSFII dataset, but expects to start using the combined 1994-96 CSFII database and the Supplemental Children's Survey in the second quarter of calender year 2000.				

Despite the overall large scope of the USDA CSFII database, some contend that USDA survey population sample sizes are of insufficient size to provide reliable estimates at the high end of exposure and risk, and that there is a need for specific OPP criteria for a minimum number of samples before an estimate is derived and used in establishing policy. The major focus of this discussion has been the small number of data points at the extremes of the consumption distribution for any given commodity. In other words, a very small number of people may have reported having eaten a food containing a particular commodity. Some state, in particular, that for many infrequently-consumed commodities or for small population subgroups, an adequate number of individuals is not available to calculate a high-end consumption percentile. They say, therefore, that the percentile exposure represented by high-end consumers of infrequently eaten foods is highly uncertain. They further note that USDA has identified minimum population sample size criteria for estimating various percentiles of food consumption, and recommend that USDA flag estimates that do not meet these criteria. They believe that OPP should not use data points that would fall at a percentile which would be flagged by the USDA. Rather, they argue that such high-end (and "uncertain") values should be discarded (or otherwise adjusted) prior to using the data set to perform probabilistic exposure analyses.

OPP recognizes that there are limits with respect to the USDA food consumption data base which would affect the reliability of estimates of high-end consumption of particular commodities. In particular, for many infrequently-consumed commodities and for small population subgroups, an adequate number of individuals may not be available to calculate a reliable high-end consumption percentile[5]. However, the ability to define a high-end consumption

[5] The Agency and USDA, together, recognized this as an issue and initiated a Supplemental Children's Survey (SCS) for the 1994-1996 USDA CSFII. The 1994-96 CSFII contains data for approximately 5700 children up to 18 years of age and the CSFII-SCS will provide intakes on approximately 5000 additional children through 9 years of age, based on OPP sample size needs. The sample design is the same as that used for the 1994-96 CSFII so that the data from the SCS can be merged with data from the 1994-96 CSFII. The size of the merged dataset (1994-96 CSFII and SCS combined) is shown in the right-hand columns of Table 2. In addition, OPP notes that the next food consumption survey will be conducted jointly by USDA and CDC's National Center for Health Statistics (NCHS) and is expected to

percentile for each commodity in an exposure assessment is not necessarily critical to ensuring that the exposure assessment can define high-end exposure. As discussed more fully below, the concern about the size of the input data base is not directed at whether the data sets are adequate to define high-end percentiles of pesticide residue levels in a *particular* food or consumption of a *specific* food form. Rather, the concern is whether the data bases are sufficiently large to characterize accurately the distribution of daily pesticide exposures from all foods which an individual eats in any given day. The distinction between estimating high-end percentiles of exposure and high-end percentiles of consumption (or residue) for a particular commodity is an important one. It may be that the part of the exposure distribution which is derived from (or includes) any single (presumably uncertain) upper-end USDA consumption value, does not necessarily produce an invalid exposure value. Many of the upper-end exposure estimates might not contain upper-end USDA consumption values and thus these uncertain USDA *consumption* values may not be driving the high-end Agency *exposure* estimates at all.

Although this concern might need to be heightened when OPP's probabilistic exposure assessments involve only a single commodity with few residue data points, one or even a few very high food consumption values do not appear likely to be the primary driver(s) of exposure and disproportionately influence the outcome of the DEEM exposure estimates (see Section III of this document for a brief summary of the testing results. A more detailed explanation can be found in the Response to Comments document associated with this policy document). The more commodities which are included in the analysis, the more unlikely it is that the upper-end exposure values are driven by upper-end USDA consumption values.

Finally, as discussed above with respect to the accuracy and representativeness of the input data bases, OPP will perform a sensitivity analysis on all probabilistic assessments of dietary exposure. The CEC module will identify the critical input data points, and the Agency can decide whether to rely on the estimates of high-end exposure in its risk management decisions.

E. Potential Variability/Uncertainty in Recipe Translations

Another area of expressed concern is potential variability/uncertainty in the recipe translations used to convert foods on an "as eaten" basis (e.g., pizza) to foods on an agricultural commodity basis (e.g., wheat flour, tomato paste/puree, milk, beef, etc.). In other words, the USDA CSFII survey requests individuals to report the foods consumed (as consumed) on any given day, but these foods have to be subsequently expressed on an appropriate agricultural commodity basis so that processing factors, appropriate matching, and other considerations can be incorporated into the analysis. This recipe translation information is an intrinsic component of the DEEM software, and concern has been expressed about the potential for these standard recipe translations to inadequately reflect the full range of actual recipes used in practice. For example, if the standard recipe for 100 grams of pizza is assumed to contain 15 grams of tomato paste instead of only10 grams (which may have been present in the particular piece of pizza eaten), then tomato

sample an expanded number of persons.

paste consumption would in this instance be overestimated. The concern is that it is impossible for a single standard recipe to be universally applicable and to account for all "real-world" variations which may exist.

At present, the recipe translation files used by the DEEM software are proprietary and not available for review. USDA and OPP have, however, collaboratively developed recipe translations which are expected to be publically-released early in the year 2000 and to be incorporated shortly after that time into the DEEM software. These translations have been peer reviewed by government and private industry food and nutrition experts.

While OPP acknowledges that it is impossible for one standard recipe to reflect the variability that exists in ingredients selected and quantities used in the many kitchens and food manufacturing facilities across the U.S., OPP does not expect these to be significant sources of error in its exposure estimates. Firstly, in many instances examined to date it is the fresh agricultural commodity such as raw apples and tomatoes which are found to be the primary risk drivers, and commodities in these forms are not subject to translation error (they are not translated). Only rarely would canned baked apples (as in apple pie) or tomato paste/puree (as in pizza sauce), for example, be expected to be risk drivers. Secondly, even considering the variability which exists in the multitude of actual recipes which are used, this variability is unlikely to be particularly significant or to introduce substantial error in our exposure estimate. If, for example, the standard recipe for 100 grams of pizza dough contains 70 grams of flour, it is unlikely that the "true" amount differs from this by more than 20-30% as there are defined limits as to how far a recipe used in practice can deviate from a standard recipe and still produce an identifiable and edible product. Finally, variability in recipes can result in either overestimating or underestimating consumption amounts and exposures and the non-systematic nature of this error (for each of many translations) is not expected to contribute to large errors which occur predominantly in any one direction in estimated exposure.

F. Impact of Agency Default Assumptions on the Choice of a Percentile Exposure Estimate for the Threshold of Concern

Some contend that the use of conservative default assumptions by OPP in its treatment of pesticide residue data results in the estimate of the 99.9th percentile exposure being significantly higher than the actual 99.9th percentile exposure. They point specifically to OPP's use of maximum rate/minimum pre-harvest interval field trial data in the exposure assessment, OPP's treatment of non-detects, and its use of 95th percentile data from monitoring studies. OPP agrees that these three assumptions would lead to an overestimate of dietary exposure, but OPP uses these assumptions only in its early- tiered, screening assessments. OPP's policy is to rely on a screening estimate of exposure only if the estimate indicates that risk would be acceptable. Because such screening estimates overstate exposure, OPP refines its exposure and risk assessments using more realistic data. These refined, higher-tiered estimates do not use the conservative default assumptions likely to overestimate risk. For more information on these tiers,

see "Guidance for the Submission of Probabilistic Human Health Exposure Assessments to the Office of Pesticide Programs " (U.S. EPA, 1998a).

With respect to residue data, OPP prefers to use data (when appropriate and available) from market basket surveys, or PDP or FDA surveillance monitoring data in conducting its pesticide exposure assessments rather than from field trials. However, these market basket or monitoring data are not always available or appropriate for use. When exposure data are obtained from field trials, these data can be modified or adjusted to take into account decreasing or increasing concentrations in processed commodities as a result of commercial processing practices or decreased residues as a result of cooking or in-home preparation such as washing, peeling, coring, etc. OPP is also able to incorporate information about lower than label-specified application rates and longer than label-specified pre-harvest intervals, if available (U.S. EPA, 1999c).

Concerning the treatment of non-detects, OPP has issued two papers containing interim guidance for handling these type of data: "Assigning Values to Nondetected/ Nonquantified Pesticide Residues in Human Health Dietary Exposure Assessments" (U.S. EPA, 1998b); and "A Statistical Method for Incorporating Nondetected Pesticide Residues into Human Health Dietary Exposure Assessments" (U.S. EPA, 1998c). In the first paper, the Agency describes its policy to use ½ the Limit of Detection (LOD) (in place of the full LOD or the limit of quantitation (LOQ)) for treated commodities in cases where the limit of detection has been adequately documented. As explained in the science policy paper, empirical data indicate that it is not unreasonable to assume that treated non-detects contain residues equal to ½ LOD. The Agency will also be performing sensitivity analyses which can determine whether the assumed residue value assigned to the non-detect values is driving the risk estimate. In the second paper, OPP has presented a method for dealing with residue data sets in which many of the observations are below detectable levels. OPP believes that these refinements in the treatment of non-detects will alleviate many of the perceived overly-conservative biases in exposure estimates with regard to the assigning of values to non-detects[6].

With respect to concern about use of the 95th percentile residue value from monitoring data, this policy has recently been revised. Prior to the availability of the probabilistic software, a registrant could, <u>for blended commodities</u> such as corn, soybean, wheat, etc., use <u>either</u> average crop field trial concentrations (which reflect the maximum label use scenario) adjusted for percent crop treated **or** 95th percentile FDA or USDA monitoring data in its acute risk assessment. Either of these would be entered as a *point estimate* for use in an acute risk assessment. The use of 95th percentile monitoring data had been introduced as an alternative to the use of average field trial concentrations since OPP believed that this would be a more realistic (but still

[6] These policy documents have been revised as per previous public comment and combined into one document entitled "Assigning Values to Non-Detected/Non-quantified Pesticide Residues in Human Dietary Exposure Assessments." This is expected to be published in its revised format simultaneously with the publication of the current *Percentile Policy* document.

conservative) estimate of actual exposures which would take into account actual use practices. Because probabilistic software is now available, OPP need not rely solely on point estimates of residue values in its acute dietary risk assessments, and this policy has been revised accordingly. OPP no longer uses a 95th percentile point estimate from monitoring data for blended commodities but instead, uses the entire range of monitoring data and therefore incorporates the entire distribution in its exposure assessment using all of the available monitoring data. Thus, the concern about OPP reliance on upper-end (95th percentile) monitoring data for blended commodities in its risk assessments is no longer justified -- OPP uses the full set of monitoring data, thereby fully incorporating the most refined concentration data available.

III. Issues Related to Public Health Policy

The FQPA directs EPA to set or retain tolerances for a pesticide only if the Agency determines that there is "a reasonable certainty of no harm" from dietary and other non-occupational sources of exposure. OPP will use available and reliable scientific information to characterize the toxicity and exposures of a food use pesticide in deciding whether a particular pesticide meets the FQPA safety standard. Put simply, OPP's goal is to regulate pesticides in such a manner that everyone is reasonably certain to experience no harm as a result of dietary and other non-occupational exposures to pesticides.

To implement this statutory standard, OPP has to make a risk management judgment about what level of pesticide residue in food is consistent with this standard. The reasonable certainty of no harm standard informs OPP judgment on the overall risk management decision as well as on component parts of the decision. OPP has decided to express its risk management judgment for acute dietary risks in quantitative scientific form, as a "threshold of regulatory concern." By a threshold of concern, OPP means that exposures below the threshold generally would not be grounds for pursuing risk mitigation but that exposures above the threshold would, at a minimum, have to be seriously examined to determine whether they exceeded the statutory standard.

A threshold of regulatory concern for acute dietary risks, when based on probabilistic exposure estimates, has two elements: (1) a percentile (or proportion) of the population for use in estimating exposure to a pesticide residue; and (2) a benchmark for judging a safe level of exposure. Due to differences in the various inputs to exposure assessments, OPP will vary the population percentile of exposure used to estimate exposure so as to ensure that in attempting to protect the general population and significant subpopulations (i.e., by not underestimating the exposure to such groups) OPP does not unreasonably overprotect these groups (i.e. by unreasonably overestimating exposure). For exposure estimates using DRES-type high-end assumptions, risk assessors should generally use the 95th percentile of exposure as a reasonable high-end exposure. For probabilistic assessments using more realistic inputs, assessors should generally use in the first instance the 99.9th percentile of exposure. The aPAD will be used as the benchmark of safety. The rest of this section discusses more fully OPP's rationale for choosing the 99.9th percentile and the concerns that have been expressed about that choice.

In adopting this policy, OPP recognizes that the choice of the population percentile for formulating a threshold of concern involves a balancing of a number of factors. OPP considered a variety of factors in formulating the policy and this policy recommends that these factors be considered in decisions regarding the population percentile in individual risk assessment. The first consideration is the size of the exposed population and the proportion which might receive daily doses above the benchmark of safety, the aPAD. A second consideration is the level of confidence OPP has in its exposure estimates, and the extent to which such estimates may overstate (or understate) potential exposure because they incorporate conservative assumptions or rely on atypical and unrealistic data. A third consideration is the degree to which individual exposures would be estimated to exceed the aPAD to the extent this can be understood. A final consideration is the degree of public health protection incorporated into the determination of the aRfD and the aPAD.

Because of the need to balance a variety of factors in selection of a population percentile for calculating a threshold of concern, OPP is issuing its views regarding population percentiles as non-binding policy guidance rather than as a binding rule. Complex risk assessment and risk management issues such as those involved in this policy seldom can be reduced to meaningful rule-style commands. Rather, the scientist and risk manager need to have flexibility in considering a variety of factors and outcomes. This policy is intended to focus the analysis on factors deemed most critical without barring consideration of other factors which may be found to be relevant. As a policy, this guidance does not – in fact, as a legal matter, cannot – draw bright lines or preclude reconsideration of basic principles. EPA would retain the option to depart from the policy. Further, affected parties remain free to challenge the specific application of the policy or the underpinnings of the policy itself.

In initially determining where to establish the threshold of regulatory concern, OPP considered a number of issues and past practices, including the 1992 Guidelines for Exposure Assessment (U.S. EPA 1992b). These Guidelines established a broad framework for Agency exposure assessments by describing the general concepts of exposure assessment and by providing guidance on the planning and conduct of an exposure assessment, including the characterization of uncertainty. Specifically regarding the use of high percentile values, the Agency in these exposure assessment guidelines has stated the following:

> Although the Agency has not specifically set policy on this matter, exposure assessors should observe the following caution when using simulated distributions. The actual percentile cutoff above which a simulation should be considered a *bounding estimate* may be expected to vary depending upon the size of the population. Since bounding estimates are established to develop statements that exposures, doses, and risks are "not greater than...", it is prudent that the percentile cutoff bound expected exposures for the population being evaluated. For example, if there are 100 persons in the population, it may be prudent to consider simulated exposures above the 1 in 500 level *[sic]* or 1 in 1000 level (i.e., above the 99.5th or 99.9th percentile, respectively) to be bounding estimates. Due to uncertainties in simulated distributions, assessors should be cautious about using estimates above the 99.9th percentile for estimates of *high-end* exposures, regardless of the size of the population. The Agency or individual program offices may issue more direct policy for setting the exact cutoff value for use as high-end and bounding estimates in simulations.

Taking the Agency guidance into account, and giving significant weight to the size of the exposed population, OPP uses as a threshold of regulatory concern the 99.9th percentile of estimated daily acute exposure from food only, using probabilistic exposure estimation techniques, should be equal to or less than the acute Population Adjusted Dose (aPAD). Under this policy, when the 99.9th percentile of estimated exposure is equal to or less than the aPAD, the estimated exposure of the vast majority of people would not exceed safe levels. Only those individuals whose exposure is estimated to be in the very high end of the exposure distribution might theoretically receive amounts of pesticide in their food that even approached the level where concern would exist. Based on OPP's experience reviewing Monte Carlo acute dietary exposure estimates, it appears that those with significantly lower exposures (i.e., at lower percentiles of estimated exposure) would be consuming levels of pesticide in their food potentially several orders of magnitude below the aPAD.

Some people, however, have argued that OPP's policy is not sufficiently protective, even at the 99.9th percentile of exposure. Because the group eating food containing pesticide residues is very close, if not equal, to the entire population of the country, currently about 270 million people, they argue that if the 99.9th percentile of exposure is equal to the aPAD, very large numbers of people, including many children, could be exposed at levels which exceed the aPAD each day. Thus, they argue for additional protections to be provided in the risk assessment or risk management process.

While OPP recognizes that, under this policy, a large number of people – including infants and children – would theoretically be exposed at levels potentially exceeding the aPAD, the Agency believes, for several reasons, that allowing this level of estimated exposure would not raise public health concerns. First, OPP believes that actual exposure is unlikely to be greater than that estimated, and in most cases would actually be somewhat lower than the estimates based on data currently available. As discussed in this and other papers, OPP uses a tiered approach and the best available data to develop estimates of exposure to pesticide residues in food. Monte Carlo techniques are used in OPP's highest, most refined tiers – the tiers which are designed to provide the most realistic estimates of exposure. Nonetheless, at the present time, because of data limitations, even the most refined estimates of exposure may still include residue values that are higher than people actually consume for one or more commodities. For some pesticide-commodity pairs, for example, OPP may not have residue monitoring data, such as USDA's PDP data. In such cases, therefore, the estimated exposure at the 99.9th percentile (or any other percentile) may overstate potential exposure, and some portion of the most highly exposed 0.1% of the population would, in actuality, be exposed at levels less than the aPAD. Nevertheless, OPP does not believe that this significantly overstates exposure and, in any case, the sensitivity analysis of Critical Exposure Contribution module will identify the extent to which high residue values for specific commodities account for the upper-end of exposure. In those instances where data limitations result in an exposure estimate which is highly conservative, it may be necessary to generate better data for the specific commodity of concern. Thus, just as when OPP uses the 95th percentile with non-probabilistic exposure assessments OPP is not suggesting that OPP is leaving 5 percent of the population unprotected, OPP is not by choosing the 99.9th percentile for

probabilistic exposure assessments concluding that only 99.9 percent of the population deserves protection. Rather, it is OPP's view that, with probabilistic assessments, the use of the 99.9th percentile generally produces a reasonable high-end exposure such that if that exposure does not exceed the safe level, OPP can conclude there is a reasonable certainty of no harm to the general population and all significant population groups.

Second, exposure at a level above the aPAD would pose public health concerns only to the extent that such exposures might result in harm. Certainly it would be difficult to justify allowing individuals to get doses of a pesticide at such levels, if OPP expected all such exposures to result in harm. From information about the general shape of the distribution curve of dietary exposures, OPP expects the vast majority of individuals estimated to be exposed to residues over the aRfD or aPAD would be exposed only to levels slightly greater than the aRfD or aPAD. OPP believes that its risk estimation methods incorporate sufficiently conservative (health protective) approaches, so that the overall approach provides sufficient protection for the small percentage of people (those above the 99.9th percentile) who may be exposed at levels slightly above the aPAD. OPP sets the aRfD well below (usually 100 to 1000 times lower than) the appropriately-chosen "No Observed Adverse Effect Level" (NOAEL) in the most relevant laboratory animal toxicity study; the aPAD may be the same as the aRfD or even lower depending upon the FQPA Safety Factor decision. Due to conservative modeling assumptions, it is possible that no actual person is exposed to levels predicted by the model. This is why OPP has included sensitivity analysis, commodity contribution analysis, and other techniques to critically examine the results of the analysis prior to making regulatory decisions.

Third, given the size of the exposed population, the occurrence of an "exceedance" would (at the 99.9th percentile level) be very infrequent for the typical individual. For example, at the 99.9th percentile, the time between exceedances, on average, would be once every 2 to 3 years. Depending on an individual's diet, an exceedence may occur more or less frequently. Collectively, this information gives the Agency confidence that its approach to protecting people from risks associated with single-day exposure to pesticides in their diet is adequately protective.

Some have asserted that OPP's policy is overly protective. One concern is that OPP's exposure methodology significantly overestimates actual exposure to the extent that the underlying data bases include outliers or unrepresentative (and unrealistically high) field trial residue data. A closely allied concern is that exposure estimates overstate exposure when the methodology uses unrealistic, conservative assumptions. These concerns are addressed in section II of this paper and, as discussed there, efforts are made to use only reliable, realistic data. Because of the careful quality control measures taken by USDA and FDA in the generation of food consumption and residue monitoring data, OPP typically accepts those data sets compiled by the respective agencies as being reliable and realistic. OPP conducts its own review of residue data from field trials and adjusts these data to better reflect actual residues on food. As discussed in section II.F., OPP believes that it does not use overly conservative assumptions. In sum, OPP does not believe that the data bases used, and the ways in which they are used to develop probabilistic exposure estimates, will produce significant overestimates of exposure at the 99.9th

percentile.

Another concern is that the data bases available for use in probabilistic exposure estimates yield estimates at the 99.9th percentile that are unacceptably uncertain, and, because of the uncertainty, OPP should use a lower percentile (e.g. the 99.5th, 99th, or 95th percentile) of exposure in its expression of the threshold of concern. While OPP agrees that the estimates of the 99.9th percentile of exposure have some uncertainty due to the use of high-end consumption and/or residue data, it does not know whether the probabilistic assessments understate or overstate actual exposure at the 99.9th percentile. Nor does OPP expect that one high-end residue or consumption value will "drive" high-end exposure estimates. In order to evaluate the possible impact of high-end values, OPP will perform sensitivity analyses at the threshold of concern to determine what data account for the largest part of the estimated exposure.

Recognizing that there was considerable concern about both the scientific and regulatory judgments underpinning the policy, OPP has performed further analysis of the methodology to provide both the Agency's staff and the public with a better understanding of the most critical elements of the methodology. Many of the scientific concerns resulted from questions about how this relatively new approach to assessing acute dietary exposure would be performed and what aspects of the methodology had the greatest impact on the outcome. In fact, reliance on probabilistic exposure modeling techniques in regulatory decision making has very few precedents. These analyses were conducted to evaluate a variety of statistical attributes of the distributions produced using the Monte Carlo technique. The activities were designed to provide a better understanding of the most critical elements of the methodology, several of which addressed the issue of the high consumption individual and his effect on the tails of the distribution. Briefly, OPP concluded as a result of these analyses:

1. For even a reasonably large dataset as would be expected for a major agrichemical with extensive nation-wide use, approximately 1,000 iterations of the DEEM software are adequate to produce reasonably stable exposure estimates at the 99.9th percentile (generally varying less than 1-3%);

2. Given an adequate number of iterations, exposure estimates from the DEEM software are reasonably reproducible (i.e., any randomly selected run is unlikely to be more than 2% from the "true value" where the "true value" represents the DEEM exposure estimate which would result if an infinite number of iterations were performed.); and

3. Extreme consumption events are not pervasive in the USDA's CSFII survey and are unlikely to have a significant effect in controlling the exposure estimate at the 99.9th percentile (i.e., acting as primary "risk drivers"). OPP will, in any case, use the capabilities of the DEEM software to identify any extreme eating occasions which might occur and fully characterize any exposure estimates which appear to be driven by high or unusual reported consumption.

If OPP adopted a policy that relied on a percentile of exposure lower than the estimated 99.9th percentile, it would need to justify its decision in public health terms as being consistent with the FQPA statutory standard. As indicated above in the discussion of whether the estimated 99.9th percentile of exposure is adequately protective, the choice of any percentile less than 100% assumes that, to the extent that estimates understate or correspond to actual, real world exposure, some portion of the exposed population could receive an amount of pesticide in excess of the aPAD. As lower percentiles are considered, the estimated size of the population potentially exposed to levels greater than the aPAD increases. Furthermore, if a lower percentile of regulatory concern were selected, a greater proportion of the population would be exposed more frequently to a one-day intake of pesticide residue that exceeds aPAD by a greater margin. For example, at the estimated 99th percentile of exposure, on average, individuals would experience an exceedance roughly once over several months. Moreover, the size of the exposed population potentially exceeding the aPAD at the 99^{th} or 95^{th} percentiles would be 10 and 50 times larger, respectively, than the number at the 99.9th percentile.

In OPP's view, the above analysis raises concerns about routinely using a lower percentile than the 99.9th. If, because of uncertainties associated with using the 99.9th percentile, OPP decided to use a lower percentile of exposure as its threshold of concern, it would still have some uncertainty in its assessment of acute dietary risk from pesticide residues in food. At the lower percentiles, the predicted incidence of these exceedances is quite high, and there is a substantial possibility that some significant number of people would be receiving doses that are considerably higher than the aPAD. But OPP would not have much certainty about either the number of people above the aPAD or, more importantly, how close to the aRfD or NOAEL their exposures might come. Therefore, if OPP chose a lower percentile as its threshold of concern, it would also need to consider whether other steps (e.g. use of an additional safety/uncertainty factor) would be needed to assure that the FQPA safety standard was satisfied.

Nevertheless, the selection of the estimated 99.9th percentile of exposure for use in calculating the "threshold of regulatory concern" should not be regarded as an immutable "bright line" from which deviation is not possible, regardless of the nature of ancillary data and information. OPP retains some discretion to choose a different percentile for regulatory concern if the conditions or situation warrant. A risk assessment should fully characterize the nature of the analysis for consideration by the risk manager and his or her selection of an appropriate regulatory threshold. A number of criteria should be considered including, for example, the exposure estimate's perceived degree of conservatism considering in particular the identity of the risk "drivers," the reliability and characteristics of the input data, the size of the affected populations, the results of a sensitivity analysis, etc. In this manner, the risk manager can evaluate how supportable the 99.9^{th} percentile exposure estimate is and evaluate whether or not it is appropriate to deviate (up or down) from the 99.9^{th} percentile. Specifically, a full and adequate characterization of the risk estimates might include a review of the following (in approximate order of relative importance):

- whether a "high-end" consumption value actually acts as a "driver" in the risk assessment. *(in many cases, high-end consumption values may not be actual "drivers" (*i.e., *significant contributors) in the risk assessment and thus may not be the primary reason behind high estimated exposures at the tails of the distribution)*

- how extreme the upper tails of the consumption curve are. *(for example, is the 95th percentile consumption value greater than four times the mean consumption?; is the 99th percentile value greater than six times the mean consumption?)*

- how far the presumed high-end consumption value is from where it would be expected to be given the pattern (or distribution) of reported consumption values in the lower percentiles. *(*e.g., *if a distribution can be reasonably established for the reported consumption values in the lower percentiles (*e.g., *70th through 95th percentiles), how extreme would the suspected outlier be in an appropriate Q-Q or other statistical plot)*

- the size of the affected subpopulation and how likely exposure estimates for the subpopulation would be subject to undue effects of outliers. *(a high-end value would be expected to have more influence on the upper-end exposure estimates in a small subpopulation than it would in a large subpopulation)*

- from a dietary standpoint, how likely the high-end value is to be a valid reported consumption event. *(for example, although they may be equally extreme from a probabilistic standpoint, consumption of three gingko fruits in a day might be considered more reasonable than consumption of 10 apples)*

- the nature of the inputs both in the overall assessment and (particularly) for the drivers. *(this would include, for example, whether input residue data included field trials vs. PDP data vs. market basket survey data; the use of default vs. actual processing factors; the extent to which single-serving values are measured vs. established by "decompositing"[7], the nature and degree of percent crop treated data, etc.)*

- comparison of exposure and consumption estimates using the 1989-91 data

[7] "Decompositing" is a mathematical deconvolution procedure used by OPP to produce estimates of pesticide residue levels in single items of produce based on the distribution of residues measured in composite samples where the residues measured in the composite samples represent average residues in a group of generally ten or more items.

and 1994-96 data. *(if both the 1989-91 and 1994-96 CSFII data sets produce similar estimates of exposure and contain similar extremes of consumption, it is more likely that high-end reported consumption is indeed an actual value or at least no affecting the exposure assessment in any significant way)*

In sum, OPP believes that the risk assessor should adequately characterize the nature of the assessment (including any biases and uncertainties) and perform a sensitivity analysis, where appropriate, such that the reasonableness of the upper-end percentile estimates (including the 99.9th) can be properly gauged. Any risk assessment performed by OPP should characterize the effect of any high-end points (on the consumption) on the regulatory percentiles of possible regulatory interest. Likewise, it is important for the risk manager, in turn, to consider the entire set of data and information available in deciding if the 99.9th percentile is an appropriate demarcation point for use in regulation. In particular, any risk management decisions should consider the effect of any high-end data values (consumption or residue) or other relevant factors and, when appropriate, be flexible with respect to the regulatory threshold selected. Nevertheless, based on the several dozen risk assessments and sensitivity analyses we have performed to date using probabilistic techniques, we do not expect this review to warrant a departure from the 99.9th percentile in the vast majority of cases.

IV. Policy Not Rules

The policy document discussed in this notice is intended to provide guidance to EPA personnel and decision-makers, and to the public. As a guidance document and not a rule, the policy in this guidance is not binding on either EPA or any outside parties. Although this guidance provides a starting point for EPA risk assessments, EPA will depart from its policy where the facts or circumstances warrant. In such cases, EPA will explain why a different course was taken. Similarly, outside parties remain free to assert that a policy is not appropriate for a specific pesticide or that the circumstances surrounding a specific risk assessment demonstrate that a policy should be abandoned.

V. Bibliography

U.S. EPA 1992a. Memorandum from H. Habicht to Assistant Administrators. "Guidance on Risk Characterization for Risk Managers and Risk Assessors." February 26, 1992.

U.S. EPA 1992b. Federal Register Notice. "Final Guidelines for Exposure Assessment." 57 FR 22888. May 29, 1992.

U.S. EPA 1997a. "Guiding Principles for Monte Carlo Analysis," Appendix A: Policy for Use of Probabilistic Analysis in Risk Assessment at the U.S. Environmental Protection Agency. Office of Research and Development. March 1997. EPA/630/R-97/001.

U.S. EPA 1997b. Memorandum from W. Wood to D. Patton. "Proposed Policy for Use of Monte Carlo Analysis in Agency Risk Assessments." January 29, 1997.

U.S. EPA 1998a. Federal Register Notice. "Guidance for Submission of Probabilistic Human Health Exposure Assessments to the Office of Pesticide Programs." Draft. 63 FR 59780 November 4, 1998.

U.S. EPA 1998b. Federal Register Notice. "Assigning Values to Nondetected/Nonquantified Pesticide Residues in Human Health Dietary Exposure Assessments." Draft. 63 FR 67063-67066. November 4, 1998.

U.S. EPA 1998c. Federal Register Notice. "A Statistical Method for Incorporating Nondetected Pesticide Residues into Human Health Dietary Exposure Assessments." Draft. 63 FR 67063-67066. December 4, 1998.

U.S. EPA 1999a. Federal Register Notice. "Choosing a Percentile of Acute Dietary Exposure as a Threshold of Regulatory Concern." Draft. 57 FR 16962. April 7, 1999.

U.S. EPA 1999b. Federal Register Notice. "Guidance of the Conduct of Bridging Studies for Use in Acute Dietary Probabilistic Risk Assessment" (dated 7/29/99) and "Guidance for the Conduct of Residue Decline Studies for Use in Probabilistic Acute Dietary Probabilistic Risk Assessment" (dated 7/29/99). Draft. 64 FR 42372. August 4, 1999.

U.S. EPA. 1999c. Federal Register Notice. "Data for Refining Anticipated Residue Estimates Used in Dietary Risk Assessment for Organophosphate Pesticides" (dated 3/26/99). Draft. 64 FR 16967. April 7, 1999.

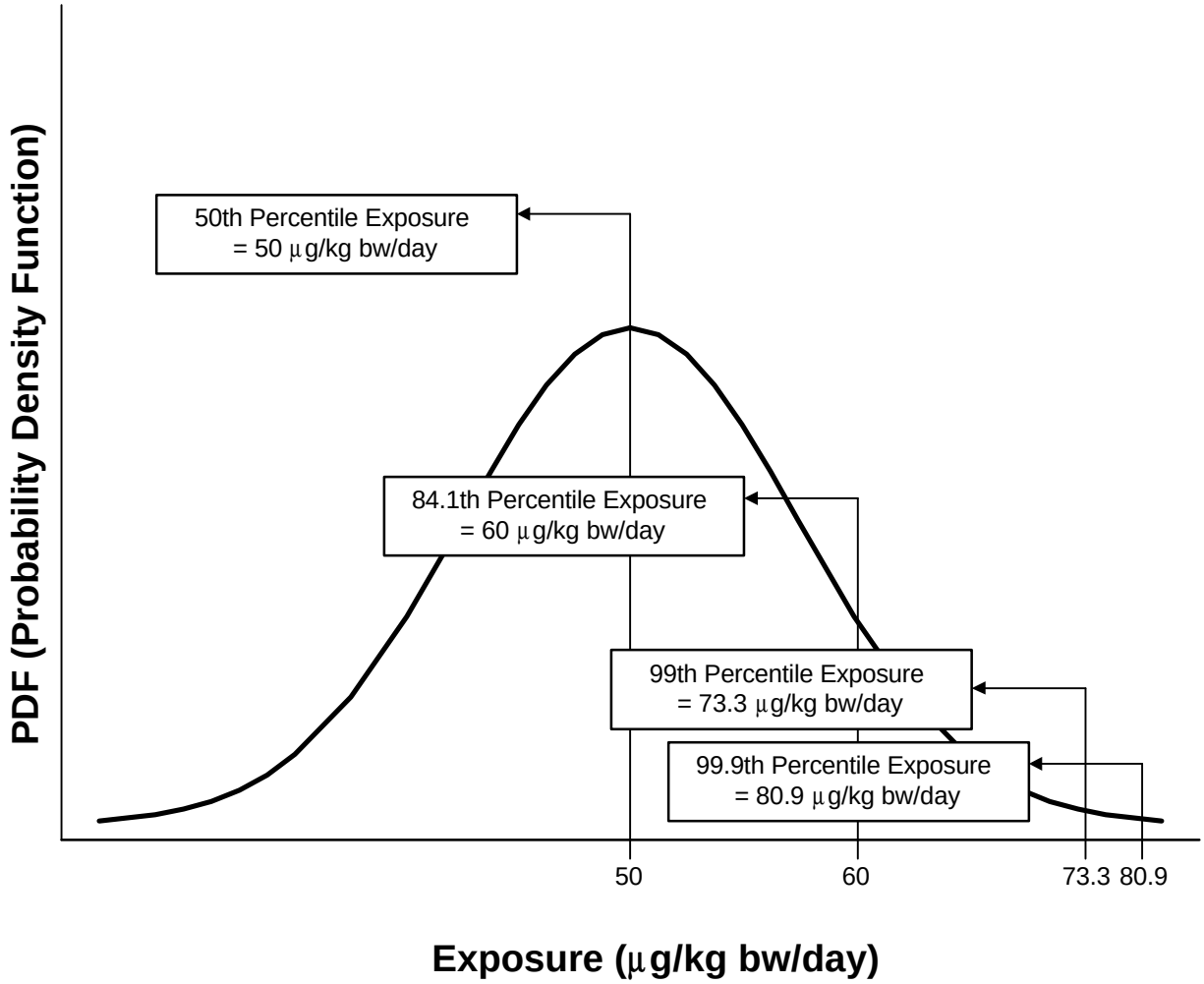
APPENDIX

Primer on Interpretation of Exposure Distribution Curves

Traditionally, OPP has selected a regulatory threshold of concern (e.g., an individual cancer risk of no greater than a range of 1-in-a-million or a %aRfD of no more than 100%) which could not be exceeded. The risk threshold was derived by calculating a high-end (or bounding) point estimate of exposure using certain high-end exposure assumptions and combining it with the toxicological endpoint to determine whether a hypothetical "high-end" individual exceeded the regulatory threshold of concern. If so, exposures were deemed to be unacceptable and mitigation actions were generally sought. However, it was not known whether OPP's "high-end" exposure estimate represented the 95th, 99th, 99.9th, or 99.999th percentile individual or if the high-end exposure estimates were well beyond the exposures received by even the maximally exposed individual (i.e., if high-end exposure estimates were above the 100th percentile).

With the advent of Monte Carlo analysis, OPP is no longer limited to assessing exposure and risks to the population using methodologies which produce only a single high-end point estimate. Monte Carlo analyses permit the risk assessor to not only produce more accurate estimates of exposure but to produce estimates of exposure across the entire population which incorporate the *probabilities* of being subjected to these exposures. This distribution of exposures can be represented graphically as a probability density function similar to the classic bell curve[8]. An example of one of these curves is illustrated on the following page:

[8] As noted in a side bar on p. 12 of the main document, pesticide exposures (or concentrations) do not typically follow the classic "bell shaped" curve, but instead have a large "hump" to the left and a long tail to the right. This is closer to what is termed a *log-normal* distribution of exposures. The use of a bell-curve in this primer permits ready recalculation of the sample tabled values from first principles of classic statistics. If this example were to have used a more realistic log-normal distribution of exposures, then calculation of the appropriate values would have been considerably more complex and would have likely led to considerably more confusion among even those familiar with basic statistics. For the purposes of this primer, OPP believes it makes sense to rely on basic (normal distribution) statistics so as to be accessible to the widest audience. Although use of a log-normal distribution and all its attendant mathematical complexities would have been more realistic, OPP believes that this would have severely hampered understanding and would have contributed minimally to comprehension of the necessary principles at the basic level.



The total <u>area under the curve</u> represents the entire population of interest. As one moves progressively from the extreme left side of the distribution (from an exposure of zero) to the right, an ever higher proportion of the population falls under the curve. As can be seen in the hypothetical example, 50% of the population is exposed to levels of 50 μg/kg bw/day or less, 84.1% of the population is exposed to levels of 60 μg/kg/day or less, 99% of the population is exposed to levels of 73.3 μg/kg bw/day or less, and 99.9% of the population is exposed to levels of 80.9 μg/kg bw/day or less.

In the above bell curve example, the exposures across the population range from a low of about 10 μg/kg bw/day to a high of about 90 μg/kg bw/day, with a mean (or average) exposure of 50 μg/kg bw/day. Also, exposures at the 50th, 84.1th, 99th, and 99.9th percentiles are 50-, 60-, 73.3-, and 80.9- μg/kg bw/day, respectively. Each of these exposure values (in μg/kg bw/day) can be converted to a %aPAD (to be compared to the threshold of concern) which is calculated by dividing each exposure value by aPAD. These

%aPADs are calculated and shown in the table below:

Hypothetical Calculation of %aPAD at Various Percentile Thresholds of Concern			
Percentile Threshold of Regulatory Concern	**Estimated Exposure at Specified Percentile Threshold of Concern[a] (μg/kg bw/day)**	**aPAD[b] (μg/kg bw/day)**	**%aPAD[c]**
99.9	80.9	75	108%
99	73.3		98%
84.1	60		80%
50	50		67%

[a] Obtained as output from DEEM software program

[b] This is calculated by dividing the NOAEL observed in animal studies (in μg/kg bw/day) by the appropriate uncertainty factors, and a decision with regard to the FPQA Safety Factor. In this hypothetical case, the NOAEL is 7500 μg/kg bw/day , the uncertainty factor is 100, and the FQPA Safety Factor has been removed.

[c]The %aPAD is calculated by dividing the estimated exposure at any given percentile by the aPAD. In general, a %aPAD of 100% or less is not considered to be of concern.

From the above table it is apparent that the %aPAD corresponding to the 99.9th percentile of exposure (at 108%) exceeds the threshold of concern. It is also apparent that "acceptable" exposures would occur at the 99th percentile. Thus, exposures would be deemed excessive if the 99.9th percentile is considered to be the threshold of concern and either risk mitigation or refinement of exposure estimates would be considered.