UNITED STATES ENVIRONMENTAL PROTECTION AGENCY
WASHINGTON D.C., 20460

OFFICE OF
PREVENTION, PESTICIDES AND TOXIC SUBSTANCES

November 8, 2007

MEMORANDUM

SUBJECT: Transmittal of Meeting Minutes of the FIFRA Scientific Advisory Panel Meeting Held August 14 - 15, 2007 on Review of EPA/ORD/NERL's SHEDS-Multimedia Model, aggregate version 3

TO: Debbie Edwards, Director
Office of Pesticide Programs

FROM: Steven M. Knott, Designated Federal Official
FIFRA Scientific Advisory Panel
Office of Science Coordination and Policy

THRU: Elizabeth Resek, Acting Director
Office of Science Coordination and Policy

Attached, please find the meeting minutes of the FIFRA Scientific Advisory Panel open meeting held in Arlington, Virginia on August 14 - 15, 2007. This report addresses a set of scientific issues being considered by the Environmental Protection Agency pertaining to the Review of EPA/ORD/NERL's SHEDS-Multimedia Model, aggregate version 3.

Attachment

cc:

James B. Gulliford
James J. Jones
Anne Lindsay
William Jordan
Margie Fehrenbach
Janet Andersen
Steven Bradbury
William Diamond
Tina Levine
Lois Rossi
Frank Sanders
Betty Shackleford
Richard Keigwin
Jack Housenger
Dana Vogel
Valerie Zartarian
Jianping Xue
Steven Nako
Douglas Parsons
Dale Kemery
Vanessa Vu (SAB)
OPP Docket

FIFRA Scientific Advisory Panel Members

Steven G. Heeringa, Ph.D. (FIFRA SAP Chair)
John R. Bucher, Ph.D., D.A.B.T.
Janice E. Chambers, Ph.D.
Stuart Handwerger, M.D.
Gary Isom, Ph.D.
Kenneth M. Portier, Ph.D.
Daniel Schlenk, Ph.D.

FQPA Science Review Board Members

John Adgate, Ph.D
James J. Chen, Ph.D.
Lutz Edler, Ph.D.
Wendy J. Heiger-Bernays, Ph.D.
Sastry S. Isukapalli, Ph.D.
John Kissel, Ph.D.
Chensheng Lu, Ph.D.
Peter D.M. Macdonald, D. Phil., P. Stat.
Nu-may Ruby Reed, Ph.D., D.A.B.T.

SAP Minutes No. 2007-06

A Set of Scientific Issues Being Considered by the Environmental Protection Agency Regarding:

Review of EPA/ORD/NERL's SHEDS-Multimedia Model Aggregate version 3

August 14 - 15, 2007
FIFRA Scientific Advisory Panel Meeting
held at the
Environmental Protection Agency Conference Center
Arlington, Virginia

Notice

These meeting minutes have been written as part of the activities of the Federal Insecticide, Fungicide, and Rodenticide Act (FIFRA), Scientific Advisory Panel (SAP). The meeting minutes represent the views and recommendations of the FIFRA SAP, not the United States Environmental Protection Agency (Agency). The content of the meeting minutes does not represent information approved or disseminated by the Agency. The meeting minutes have not been reviewed for approval by the Agency and, hence, the contents of these meeting minutes do not necessarily represent the views and policies of the Agency, nor of other agencies in the Executive Branch of the Federal government, nor does mention of trade names or commercial products constitute a recommendation for use.

The FIFRA SAP is a Federal advisory committee operating in accordance with the Federal Advisory Committee Act and established under the provisions of FIFRA as amended by the Food Quality Protection Act (FQPA) of 1996. The FIFRA SAP provides advice, information, and recommendations to the Agency Administrator on pesticides and pesticide-related issues regarding the impact of regulatory actions on health and the environment. The Panel serves as the primary scientific peer review mechanism of the Environmental Protection Agency, Office of Pesticide Programs (OPP), and is structured to provide balanced expert assessment of pesticide and pesticide-related matters facing the Agency. FQPA Science Review Board members serve the FIFRA SAP on an ad hoc basis to assist in reviews conducted by the FIFRA SAP. Further information about FIFRA SAP reports and activities can be obtained from its website at http://www.epa.gov/scipoly/sap/ or the OPP Docket at (703) 305-5805. Interested persons are invited to contact Steven Knott, SAP Designated Federal Official, via e-mail at knott.steven@epa.gov.

In preparing these meeting minutes, the Panel carefully considered all information provided and presented by EPA, as well as information presented by public commenters. This document addresses the information provided and presented by EPA within the structure of the charge.

TABLE OF CONTENTS

SAP Minutes No. 2007-06

A Set of Scientific Issues Being Considered by the Environmental Protection Agency Regarding:

Review of EPA/ORD/NERL's SHEDS-Multimedia Model Aggregate version 3

August 14 - 15, 2007
FIFRA Scientific Advisory Panel Meeting
held at the
Environmental Protection Agency Conference Center
Arlington, Virginia

Steven G. Heeringa, Ph.D.
FIFRA SAP Chair
FIFRA Scientific Advisory Panel
Date: NOV 0 8 2007

Steven M. Knott, M.S.
Designated Federal Official
FIFRA Scientific Advisory Panel
Date: NOV 0 8 2007

Federal Insecticide, Fungicide, and Rodenticide Act
Scientific Advisory Panel Meeting
August 14-15, 2007

Review of EPA/ORD/NERL's SHEDS-Multimedia Model Aggregate version 3

PARTICIPANTS

FIFRA SAP Chair

Steven G. Heeringa, Ph.D., Research Scientist & Director for Statistical Design, University of Michigan, Institute for Social Research, Ann Arbor, MI

Designated Federal Official

Steven M. Knott, M.S., FIFRA Scientific Advisory Panel, Office of Science Coordination and Policy, EPA

FIFRA Scientific Advisory Panel Members

Janice E. Chambers, Ph.D., D.A.B.T., William L. Giles Distinguished Professor and Director, Center for Environmental Health Sciences, College of Veterinary Medicine Mississippi State University, Mississippi State, MS

Stuart Handwerger, M.D., Professor of Pediatrics, University of Cincinnati Children's Hospital Medical Center, Cincinnati, OH

Kenneth M. Portier, Ph.D., Program Director, Statistics, American Cancer Society, Statistics and Evaluation Center, Atlanta, GA

Daniel Schlenk, Ph.D., Professor, Aquatic Toxicology, Department of Environmental Sciences, University of California, Riverside, CA

FQPA Science Review Board Members

John Adgate, Ph.D., Associate Professor, Division of Environmental Health Sciences, School of Public Health, University of Minnesota, Minneapolis, MN

James J. Chen, Ph.D., Senior Mathematical Statistician, Division of Biometry and Risk Assessment, National Center for Toxicological Research, US Food and Drug Administration Jefferson, AR

Lutz Edler, Ph.D., Head of the Department of Biostatistics, German Cancer Research Center, Heidelberg, Germany

Wendy J. Heiger-Bernays, Ph.D., Associate Professor, Department of Environmental Health Boston University School of Public Health, Boston, MA

Sastry S. Isukapalli, Ph.D., Assistant Professor, Department of Environmental and Occupational Medicine, University of Medicine and Dentistry of New Jersey - Robert Wood Johnson Medical School, Piscataway, NJ

John Kissel, Ph.D., Professor, Department of Environmental Health and Occupational Health Sciences, School of Public Health and Community Medicine, University of Washington Seattle, Washington

Chensheng Lu, Ph.D., Assistant Professor, Department of Environmental and Occupational Health, Rollins School of Public Health, Emory University, Atlanta, GA

Peter D.M. Macdonald, D. Phil., P. Stat., Professor of Mathematics & Statistics, McMaster University Hamilton, Ontario, Canada

Nu-may Ruby Reed, Ph.D., D.A.B.T., Staff Toxicologist, California EPA, Sacramento, CA

INTRODUCTION

The FIFRA Scientific Advisory Panel (SAP) has completed its review of **EPA/ORD/NERL's SHEDS-Multimedia Model, aggregate version 3**. Advance notice of the SAP meeting was published in the *Federal Register* on May 16, 2007. The review was conducted in an open panel meeting August 14 – 15, 2007 held in Arlington, Virginia. Dr. Steven Heeringa chaired the meeting. Steven Knott served as the Designated Federal Official.

The purpose of this review was to request input from the SAP on EPA/ORD/NERL's Stochastic Human Exposure and Dose Simulation for Multimedia, Multipathway Pollutants (SHEDS-Multimedia; also referred to in this document as SHEDS), aggregate version 3 model. SHEDS-Multimedia version 3 is a state-of-the-science computer model for simulating human exposures to multimedia, multipathway environmental pollutants including pesticides. It is a physically-based, probabilistic model that predicts, for user-specified population cohorts, exposures incurred via eating contaminated foods or drinking water, inhaling contaminated air, touching contaminated surface residues, and ingesting residues from hand- to-mouth or object- to-mouth activities. To do this, it combines information on chemical usage, human activity data (e.g., from Consolidated Human Activity Database (CHAD) time/activity diary surveys and videography studies), environmental residues and concentrations, and exposure factors to generate time series of exposure for simulated individuals. One-stage or two-stage Monte Carlo simulation is used to produce distributions of exposures for various population cohorts (e.g., age/gender groups) that reflect the variability and/or uncertainty in the input parameters. While the core of SHEDS-Multimedia is the concentration-to-exposure module, there are various options (built-in source-to-concentration module; user-entered time series from other models or field study measurements) for obtaining concentration inputs, and SHEDS-Multimedia exposure outputs can be used as inputs to PBPK models.

Finally, the SHEDS-Multimedia version 3 single chemical model can address many useful aspects of aggregate and cumulative risk assessment, related to population aggregate exposures for different multimedia chemicals and the important contributing pathways and factors. Such information will be useful in identifying populations and exposure scenarios of greatest concern for different classes of chemicals. These populations and exposure scenarios will in turn be used to determine the most relevant chemical (e.g., pyrethroid) combinations for which hazard/exposure factors information will need further development in order to support a PBPK dose modeling approach. EPA plans to extend the current single chemical aggregate version of SHEDS to a cumulative version. The cumulative version of SHEDS will be used to estimate exposure resulting from cumulative exposure to pyrethroid pesticides. At this meeting, the FIFRA SAP was asked to review the following: The dietary module of SHEDS version 3; the residential module of SHEDS version 3; and planned methodologies for extending SHEDS-Multimedia version 3 (aggregate) to SHEDS-Multimedia version 4 (cumulative).

Review of the dietary module included the methodology and model evaluation. Review of the residential module included the SAS code, graphic user interface (GUI), Technical Manual, and User Guide. Review of the planned methodologies to extend the single chemical aggregate version of SHEDS (version 3) to the cumulative version (version 4) included: algorithms for multiple chemicals and co-occurrence; fugacity-based module for residential concentration predictions; new methodologies for enhanced longitudinal activity diary simulation; Sobol methodology for enhanced sensitivity analyses; planned approach for combining residential and

dietary modules; and planned coding and GUI changes for version 4. The Panel members were not asked to review chemical-specific inputs or evaluate outputs at this SAP meeting.

Tina Levine, Ph.D., Director of the Health Effects Division in the Office of Pesticide Programs, provided opening remarks at the meeting. The agenda for this SAP meeting included an introduction of the issues under consideration provided by Dana Vogel, Health Effects Division, Office of Pesticide Programs, EPA. Presentations of technical background materials were provided by Valeria Zartarian, Ph.D. and Jianping Xue, M.D., National Exposure Research Laboratory, Office of Research and Development, EPA and Steven Nako, Ph.D., Health Effects Division, Office of Pesticide Programs, EPA.

PUBLIC COMMENTERS

Oral statements were presented by:

Iain D. Kelly, Ph.D., Jennifer L. Lantz, Ph.D., Gary J. Mihlan, Ph.D., Bruce M. Young, Ph.D., on behalf of Bayer CropScience.

Written statements were provided by:

Iain D. Kelly, Ph.D., Jennifer L. Lantz, Ph.D., Gary J. Mihlan, Ph.D., Bruce M. Young, Ph.D., on behalf of Bayer CropScience.

SUMMARY OF PANEL DISCUSSION AND RECOMMENDATIONS

Overall the Panel was impressed with the current version of the SHEDS model and future plans for the model. It was developed with good quality assurance, incorporating suggestions from earlier Panels. The majority of the Panel had no difficulty loading and running the program. The GUI is helpful and straightforward, but most Panel members felt it could be improved by adding context specific information that is clearly available at all times. One of the major issues with using the program is that it places great responsibility on the user to make sure that the values used for inputs are logical. The Panel suggested a number of proposed features that could alleviate this issue.

The Panel had a number of recommendations for future revisions of the model. The Panel noted a need for clearer instructions on run times, and it is important that units of input variables be displayed because specification of units fosters better user understanding. Future versions of the model need to: 1) provide a case study data set to help orient the user and promote understanding of model features and outputs; 2) improve dermal exposure algorithms; 3) implement the tracking of both pesticide residues and relevant environmental media; 4) develop and use standard populations for the dietary analysis; and 5) improve methods for dealing with correlation and parameterization of variables. Moreover, the Agency needs to clearly state the criteria and goals that will be used to evaluate the outputs of model sensitivity and uncertainty analyses, and provide tools for repeating or reenacting simulations to perform model quality assurance.

The Panel members believe that the model algorithms appear to be technically correct, with the caveat that it is impractical to perform a line-by-line review and verify technical correctness of the model code. The Panel members agree that the code is clear enough that the algorithms can be easily followed. Adequate and well described commenting in the code (about 30% comment lines out of a total of about 8500 lines of SAS code), along with a logical organization, are the positive features of the code. Logical separation of data and code, a sound practice, has been followed in the SHEDS development process. However, the naming of the variables within the SHEDS code can be significantly improved.

The Panel had a number of suggestions on the organization, clarity, completeness, and usefulness of the model manuals. In general, the Panel felt that the audience and scope should be more clearly defined, and that the case study presented in the manual needs to illustrate the strengths and limitations of the model. Both the User Guide and the Technical Manual have important, different functions, and the documents should reflect this difference. The User Guide has the primary task to inform users about installing and running the model and walking them through common tasks, while the Technical Manual presents the details. Given the centrality of the CHAD diaries and human activities for developing model runs, it is important that these data be presented in a clear, illustrative, and concise manner.

The Panel felt that the logic of extending the application of SHEDS from single to multiple chemical assessments is plausible and warranted. However the nature of such proposed work would be complex, time and resource consuming, and the process itself may pose significant challenges. The Panel expressed several concerns about the planned methodology. Although the move from aggregate to cumulative involves relatively few changes to the code, model structure, and interface, there are a number of challenges inherent in doing so. The Panel focused its comments on issues associated with tracking of mass of media and chemical, product co-occurrence, pesticide metabolites, programming issues, scenarios, database representativeness and absorption and bioavailability.

The Panel was asked to comment on a number of technical issues associated with proposed improvements to SHEDS. In general, the Panel felt that often there was insufficient information provided to make a sound decision because the overall effect of proposed changes was unclear. The Panel agrees that chemical redistribution following residential pesticide application is scientifically well established and that the fugacity approach is an appropriate method for addressing this phenomenon. The Panel believes that the fugacity approach is well suited to treatment of multiple chemicals and is therefore compatible with other development goals for SHEDS, but the Panel had less unanimity about whether the fugacity option would provide more useful results than approaches currently used in the model. The Panel expressed a desire to see example calculations and sensitivity analyses that illustrate the performance of the fugacity-based approach. The Panel notes that EPA proposes to evaluate the fugacity module using chemical transport data from a 1-house study. While this is a reasonable place to start, the Panel encourages more extensive testing including application to cases in which both environmental data and biomonitoring data are available.

The Panel believes that the concept behind the new longitudinal diary assembly method proposed by the Agency has scientific merit, and the technical aspects of the proposed method based on the D-statistic and 1-day lag autocorrelation is sound. However, the application of this method for predicting aggregate human exposures to environmental chemicals may be somewhat limited. This is particularly true for activities that may lead to potential exposures that are either difficult or implausible to track as continuous key variables. For example, the D-statistic may be useful for tracking exposures over time when intake rates are constant and highly correlated, but less useful for simulating exposures that may be highly variable and/or dependent on multiple factors, such as the frequency of hand contact on a contaminated surface.

The Panel believes that it is likely that the current activity database (CHAD) may not contain enough diaries to allow random draws using the new longitudinal diary assembly method and several members raised concerns about the representativeness of the US population in the CHAD dataset. Since a significant number of CHAD diaries were excluded for various reasons, further proposed decisions that reduce an already small dataset might be problematic for accurately determining the within-subject variation. Furthermore, as stated in the documentation, two of the most important pieces of information, chemical usage patterns and contact possibilities, are not included in the diary database. These concerns dampened the enthusiasm of the Panel for implementing this new longitudinal diary assembly method in the upcoming revision of SHEDS.

The Panel concurs that including a lag autocorrelation function (currently set at 1-day) would be necessary to eliminate repetitive activities from the longitudinal diaries that are likely unrealistic. However, the Panel suggests that the Agency's efforts should focus on behaviors relevant to potential residential pesticide use, such as frequency of gardening activity. The use of realistic and well documented scenarios in the model is, in the Panel's judgment, as high or higher in priority as further refinement of the diary assembly and autocorrelation methods. For the explicit purpose of estimating the individual and population exposures, it may be more constructive for SHEDS 4 to suggest defaults for key variables and allow for deviation from these defaults as needed by informed users to display the strengths and limitations of the future versions of the model.

The Panel was divided on the utility of Sobol's method, and wondered whether the new sensitivity analysis method is needed. Several Panel members expressed concern that the required number of simulations for sensitivity analysis appeared to be very low. The Agency's view that a single sensitivity run can be considered to represent multiple "individual-days" of simulations is not correct, because a single variability run will ultimately produce a single population metric (e.g., the 95^{th} percentile for population exposures). It appears that the proposed use of Sobol's method will allow the computation of "local sensitivity," without addressing the corresponding uncertainty. Global sensitivity analysis, on the other hand, allows the estimation of sensitivity as well as uncertainty, and thus would be a preferred approach.

The Panel was not comfortable choosing from the list of proposed future model features as a dialogue with the developers is needed to better inform the Agency's priorities. As a starting point, the Panel felt that the developers should do what is simple and doesn't involve a major reorganization of the code. Most Panel members thought that PBPK modules should not be a high priority, but that the output of SHEDS runs should be available for PBPK modules. The Panel felt that the Agency should get more experience with the application of this model before getting into a major reorganization of the code. It is important that the developers anticipate future data sets, especially longitudinal, and try to keep the code scalable. Autocorrelation can have an important effect on the tails of the exposure distributions. Anything that helps the user understand uncertainty and variability is important. Several Panel members emphasized that improved graphics were important for wide model comprehension and acceptance. They also felt it was useful to add utilities to make the model more accessible and useful, e.g., code to review input values and note extreme or illogical combinations, or a script generator to run designed experiments on the model.

Regarding the Agency's inquiry about decision rules for dietary exposure, the Panel noted that the document did not include a clear list of rules for review. A decision diagram or flow chart is needed to allow for substantive critical review. As for setting rules for assigning residue levels, the Panel was in general agreement that simple rules have the advantage over complex rules for having clarity, uniformity of application, and transparency in operation.

The Panel believes that SHEDS should incorporate the timing of eating occasions in dietary exposure assessments, either by assigning the same or different residue to foods consumed on multiple eating occasions. The majority of the Panel believed that it is not sufficient to construct the longitudinal dietary consumption patterns based on the 8-day eating occasions. Recent research suggests that a longitudinal dietary survey (a minimum of 6 consecutive days' dietary consumption in each of the 4 seasons) would be adequate to be representative of a one-year dietary consumption pattern at the individual level and would offer substantial improvements for assessing dietary exposures compared to standard cross-sectional study design with repeated sampling, i.e., the approach currently used in SHEDS. The Panel recommends that the Agency should conduct analysis similar to those already performed with subsets of the CHAD database, to see if the pattern of D-statistic may be useful for longitudinal implementation. The Panel generally agreed that significant value would likely be added to the SHEDS model by collecting longitudinal dietary consumption, as well as individual behavior/activity data, particularly activities at the microenvironment level from a small group of study participants stratified by age and gender. As with the CHAD database, some Panel members raised concern about the representativeness and reliability of the data for the current US population.

With the Bayer DWCS water consumption data, EPA has the opportunity to change and possibly improve the drinking water model in SHEDS. The current SHEDS approach is to assign each individual a total amount of water consumed each day according to the daily water consumption amounts in the Continuing Survey of Food Intake by Individuals (CSFII) diaries and then distribute this amount in six equal amounts to six time points spread evenly through the day. The data from DWCS could be used to either assign new total daily consumption amounts, use the drinking events data to create new allocations patterns throughout the day or both.

In attempting to answer this question regarding the Bayer data, the Panel raised questions of its own, detailed in the report, and noted that the summary of the analysis made available to the Panel indicated that the approach currently implemented in SHEDS appears to be representative. A number of Panel members wondered if the DWCS database was properly weighted to produce population relevant information. In particular, there is the potential for bias in study results because of the low response rate and high proportion of incomplete diaries. An attempt should be made to compute sampling weights to better match study results with the demographics of the US population.

The Panel noted that exposure from drinking water sources is a very small fraction of total exposure in many cases, and some Panel members felt that because of its low contribution to total exposure, further drinking water consumption modeling would not noticeably change model results. Other Panel members felt that in some exposure scenarios the contribution from drinking water could represent a large fraction of total exposure and in those cases different water allocation models could produce quite different results. The importance of water consumption will depend on the type of chemical and the likelihood of that chemical entering the drinking water supply, and further analysis by the Agency is needed to demonstrate the need for significant work on this issue. A more substantive scientific issue noted by the Panel was the assumption in SHEDS that all water consumed in one day has the same concentration, regardless of source, a situation that is unlikely to be true given the diversity of sources of drinking water for most Americans over the course of a day. Previous SAP consultations on this topic remarked at the lack of adequate information on drinking water consumption patterns to properly analyze the fraction of exposure from this source. A good portion of the Panel do not think that the issues

related to exposure from drinking water sources are likely to be less in the future and suggest that drinking water's impact be examined in more detail.

The Panel was asked to comment on the uncertainty approach of daily dietary exposure presented in the SHEDS Dietary Model The approach proposed creates limited CSFII person-day datasets each containing a fraction of the full dataset and a fraction of the commodities-containing-residues database. Exactly what was being proposed and how to interpret the results were not uniformly clear to all Panel members. In general, those Panel members commenting felt that both the median tail spread and the median upper extreme percentile of the response distribution were reasonable indices. Working with these indices and statistics (median and upper percentiles of the 95/5 percentile ratio distributions), members of the Panel had suggestions for other analytical approaches for model uncertainty, noting that the uncertainty estimates as presented involve four variables: the percentage of CSFII food consumption data used (denoted p_1), the percentage commodities used (denoted p_2), the number of samples generated in one run of the SHEDS model and from which the response distribution was constructed (referred to as *b*), and the number of SHEDS runs for the uncertainty analysis (denoted as *r*). The four methods noted by the Panel were: 1) measuring the impact of changes to p_1; 2) post hoc power analysis; 3) inter- and intra-run variability; and 4) alternate indices. The strengths and limitations of these approaches were outlined by the Panel, but without further analysis the Panel cannot recommend one approach over another, but noted that all will likely provide more useful information.

Finally, the Panel was asked to determine the extent to which region- and season- specific dietary consumption amounts and patterns might be important in developing dietary exposure estimates. The NHANES (National Health and Nutrition Examination Survey) database does not include information on season nor region because of concerns over privacy, but food consumption likely varies with season and from region to region. Panel members are not aware of any existing statistical methods that might be directly used to determine the extent to which region- and season- specific dietary consumption amounts and patterns might be important in developing dietary exposure estimates. If different regions or different seasons would result in different exposure estimates, then the question is the utility of NHANES dietary consumption data given the lack of access to information on the season and region. Some Panel members suggested that it may be possible to generate region-specific or season-specific food lists. A simple analysis of variance (ANOVA) approach can be applied to generate region-specific commodity lists, and the Panel described a method to generate season-specific food consumption lists.

There are other issues, such as fasting over night or over a period of 8 hours prior to reporting for medical examination, associated with the NHANES protocols that would further limit the use in SHEDS development. The Panel strongly recommends that the Agency should be involved in the design stage for the future NHANES and other longitudinal studies to address missing data/information issues for SHEDS.

The limitations in access to seasonal and geographical information in NHANES data as well as other limitations of these data discussed by the Panel initiated discussion of alternate sources of data. The issue of the regional and seasonal differences can be addressed through analyzing the population and region- or season-stratified datasets, such as data from, for example, CSFII. Using bootstrap methods, the percentiles of the exposure estimates can be computed for each region, each season, and region and season combination. The 50^{th}, 95^{th}, and 99^{th} percentiles can be compared to evaluate the impact of regional and/or seasonal differences in diaries, consumption, and exposure on the exposure estimates.

PANEL DELIBERATIONS AND RESPONSE TO CHARGE

The specific issues addressed by the Panel are keyed to the background documents, references, and the charge questions provided by EPA.

Issue 1: Documentation, completeness, and clarity of the technical aspects and usability of the SHEDS-Multimedia version 3 (aggregate residential) model

Question 1-1:

a) Were Panel members able to load the software on to their computers? What, if any, difficulties were encountered in loading or running the software?

The majority of the Panel had no difficulty loading and running the program. It is good that administrator privileges were not required and all SHEDS files were placed in a convenient location. The Panel recommends, however, that the prerequisite of SAS version 9.1.3 needs to be clearly stated in the introduction and installation instructions up front. This information is only found in the Users Manual after the introduction to the Demonstration Case Study. Furthermore, it would be helpful to provide users with some idea as to run time expected in the documentation.

The Panel also noted that it is important that units of input variables were displayed on most, but not all prompts. Units of input variables should be clearly stated in all cases because specification of units is important to user understanding.

Some Panel members experienced complications because of the installation problems with SAS and the underlying Java Runtime Environment. One issue identified by one Panel member was that the current SHEDS installation program does not appear to check for SAS installation prior to installing SHEDS. Subsequent installation of SAS (after installing SHEDS) does not help SHEDS run. So, the SHEDS installation program should check for SAS and abort installation if SAS is not installed. Panel members also noted a bug in the Installation Process: the batch file, "multimedia.bat" in the installation folder, contains incorrect information on the installation folder location, and hence does not run. It should be fixed with the correct folder location.

Cross-platform compatibility was one issue that the Panel felt should be addressed in future iterations of the program. Since the SHEDS program is essentially a SAS program with a selected set of parameters to be configured (e.g., just the paths to the main modules), and since

no operating system specific commands are included in the underlying SAS code, it should be straightforward to run the model on other platforms. Therefore, it may benefit users of other operating systems if the installation package is also provided as a zip archive (i.e., the users can then extract the folder and change the SAS invocation based on the requirements of their operating system). This facilitates, with a minimal amount of effort, use of SHEDS on other operating systems. Such a package could be provided through EPA's ftp servers.

It appears that the uninstall program simply removes the underlying SHEDS program and all the simulations the user may have defined and run until that point. This should be clearly noted, and the user should be provided detailed information and prompted about this step.

b) The SHEDS-Multimedia version 3 graphical user interface (GUI) was designed to be user-friendly to exposure modelers and risk assessors. Please comment on the organization and usability of the GUI, any difficulties you encountered, which features and output capabilities were most useful, and whether any additional options would be helpful. Please also comment or offer suggestions for improving the GUI/model interface.

The GUI is helpful and straightforward, but most Panel members felt it could be improved by adding context specific information clearly available at all times. One of the major issues with the program is that it places great responsibility on the user to make sure that the inputs are logical. The user can, for example, assume an exponential distribution for the value of a parameter, with no suggestion that it has to be between 0 and 1. The Panel also felt that there should be a more convenient way to see a summary of the inputs to facilitate checking that all values are logical.

The Panel identified three major issues that should be addressed in future versions of the software. First, the Panel felt that there was need for a substantive case study to guide novice users, preferably using realistic data. Second, Panel members noted that there were many hard-wired approaches in the model, and these need to be clearly identified in the program and the manual and preferably open to adjustment by advanced users from within the program. Lastly, Panel members expressed concerns about the quality of the graphics and plots generated by the program. Although this is a problem associated with SAS itself, providing systems and/or advice to the user on methods to provide reasonable quality graphics and plots is important for promoting SHEDS usage.

A novice user, without appropriate data sets, will have significant difficulty or be overwhelmed by the number of inputs to provide, even for a simple case. This highlights the need for an example case that will also allow users to make minor customizations and not have to develop a case from scratch. One Panel member wanted to focus on modeling female subpopulations, and then changed the inquiry to include both sexes, only to have the program lock up with no ability to return to previous screens. Since there was no option to cancel, the user needed to restart SAS. The user should be allowed to return to previous selections or the screen revised to include a running output of the selections made. For example, it would be useful to be able to return to the "specify exposure scenario details" screen easily from submenus.

Error checking and reporting need to be improved within the program in terms of the number of simulations that can or should be run. At present the system allows the input of an unreasonably large number of individuals to simulate (on the order of 10^{200}: higher numbers are reset to 10),

without checking or reporting disk space or processing time requirements. Similarly the specification of run names needs to be made more systematic. When a run name is not specified, instead of a quick error message, the application goes to the next screen and provides a cryptic error message. Later, a blank case is created. It is probably better to have a default such as "untitled" following the conventions used in many applications.

When checking for errors (using the "Check for Errors" button), the application gave very cryptic messages such as: "deleted", "library not found", etc., and then finally reported "no errors found." The error messages are likely due to outputs of some underlying operations that do not impact the specific case, so they should be appropriately filtered out.

Additional **major features** that would be desirable for SHEDS are

1) the ability to specify alternative locations or folders for storing simulations;

2) the application should allow the user to specify different folders for storing simulations (this will often be dictated by disk space constraints, etc.) and allow a user to organize simulations in a hierarchical manner. A major advantage of this approach is that the previous simulations are saved even when SHEDS is uninstalled;

3) the model documentation should clarify what is meant by some terminology within the program, for example: the use of terminology such as "Application Scenarios" versus "Run Files" can be confusing;

4) the interface should include a "tool tip" for each menu item, which provides descriptive information while being unobtrusive; and finally,

5) the application should include functional help buttons.

Desirable additional **minor features** noted by the Panel include having the "Run Name" shown on the title bar of the pop-up windows at all stages. This is even more important in the context of the following software bug noted by one Panel member: when the user asks to overwrite an existing run with the one he/she is editing, an error message is displayed (saying that the run already exists), and the application loads up the existing old run. The user is not provided an opportunity to force the change, and has to cancel the operation. However, upon canceling the operation, the old run becomes the "current run", and any further changes made will be reflected in that run. Having the run name displayed prominently will alert the user right away.

c) Please comment on the organization, clarity, completeness, and usefulness of the User Guide document and provide any suggestions for improvement.

The User Guide is clearly written, and easy to understand. Panel suggestions and recommendations for improvement are listed below.

- The SAS log window will accumulate numerous unessential messages, so it might be useful to have a separate SHEDS window to show progress of a simulation.

- Graphics quality is poor. It is very important that better graphics be included or allow export of data to another program that can create legible black and white and color graphics.
- The issue of truncation of distributions is important for the user to be able to review. The program should warn the user when inappropriate values are selected.
- The overview of the SHEDS interface, shown in Figure 1 of the User Guide, is inconsistent with the flow of the program.
- The installation documentation should provide information relevant to Windows Vista, especially given that in a time frame of one or two years, a significant portion of the user community may be using Vista.
- The minimum requirements for running SHEDS, as described in the User Guide, appear unrealistically low (e.g., the actual requirements for Windows XP and SAS may actually exceed the requirements described in the manual). The actual requirements for running the model are unlikely to be met by most of the commonly used computers. Specifying more realistic requirements for a "reasonable performance" or just specifying the requirements as "a computer that runs SAS (version x)" are two possible alternatives.
- Sections should be numbered, and additional cross-referencing should be included, including hyperlinks between the User Guide and Technical Manuals and a substantive glossary.
- The captions for figures and tables in the User Guide should be more descriptive.
- The User Guide also should contain some of the technical details of the probability distributions (can be directly copied from the Technical Manual). As written, the manual expects users to understand the type of distribution they are using.
- Other models developed by EPA and revised over the years have addressed many user issues in a readable manner. A review of the IEUBK (Integrated Exposure Uptake Biokinetic Model for Lead) interface and User Guide will address some of the challenges faced by new users.
- "Specify Application Exposure Scenarios" does not include the option to add new exposure scenarios relevant to pesticide application. The model should allow assessment of non-standard application methods.

Question 1-2:

a) Please comment on whether the descriptions of specific model components are scientifically sound and whether the algorithms described in the Technical Manual represent the state of the science for performing exposure assessments. Please also comment or offer suggestions for improving or modifying these algorithms or other aspects of the model construct.

In general the model is scientifically sound and appears to represent the state-of-the-science, with the notable exception of the dermal exposure analysis section, as discussed below. The larger overlying issue is the quality and the transparency of "hard-wired" approaches within the model. The representation of some of these issues, e.g., hand to mouth contact, are in existing EPA documents or peer-reviewed publications, but are largely invisible to the user and not clearly identifiable without careful examination of the underlying SAS code. While it is not necessary to present this data in the Technical Manual, several of the suggestions about the transparency of processes and equations noted by the Panel are important for understanding the functioning and

reasonableness of the model. Equations, units, and key decision points should be clearly stated and, if needed, hyperlinked to important decision variables.

The Panel's greatest concern was about dermal exposure analysis. SHEDS appears to use both the transfer efficiency (TE) and transfer coefficient (TC) approaches in its assessments. The Panel noted that, at present, the documentation appears to favor the use of the TE over TC approach. The Panel urges use of a scientifically defensible approach to dermal exposure estimation, and notes that the current state-of–the-practice of dermal modeling in the regulatory sphere is weak. The SHEDS dermal protocol is regarded as state-of–the-practice, not-state-of–the-science. The Panel noted that the TE approach is misnamed, as it does not represent efficiency, has no inherent internal logic (it merely represents observed similarities), and cannot incorporate different types of surfaces or those with varying degrees of contamination. As an example, one Panel member noted that one oft-cited use of a TE approach (Gurunathan et al., 1998) violates conservation of mass. Therefore, the Panel urges caution in use of this approach, as it needs to be anchored with and compared to other relevant data, such as biomonitoring. A joint EPA/Cal DPR/Health Canada document (presented at the January 2007 FIFRA SAP meeting) compared observed biomarker levels to predictions based on dosimetry using data from the Pesticide Handlers Exposure Database (PHED). Both under- and over-prediction of biomarker results was reported. Chemical-specific analysis reveals that this outcome was not random. Exposures were more likely to be under- or over-predicted depending upon the particular chemical examined. For example, chlorpyrifos exposure was generally underestimated. One likely explanation for at least a portion of the shortfall is the default assumption of dermal absorption of 3% or less for chlorpyrifos, regardless of exposure conditions. This estimate is based on studies conducted at high loadings that are likely not relevant to low level, residential exposure scenarios.

The Panel noted that the paragraph at the bottom of page 43 is incorrect in its assertions about the rate of dermal transfer. Dermal absorption (at least in the sense of binding to stratum corneum as opposed to systemic absorption) is not necessarily slow relative to diary events in the context of residential exposure scenarios. For example, Fenske & Lu (1994) showed apparent rapid uptake of chlorpyrifos at low skin loadings. Failure to account for dependence of absorption efficiency on skin loading is a significant weakness of current practice.

The age grouping in SHEDS is slightly different from the dietary exposure analysis default groups. Operationally, this may not be a major problem as long as it is possible for the user to conduct model runs for a custom population of choice. However, since the apparent differences in age cohort pooling has underlying emphasis that is specific to certain key parameters of interest, it would help to include a brief explanation for the basis of the SHEDS specific age cohort system. The same should be presented for dietary and water components when they are added to the next version of SHEDS. For practicality, the Agency may want to consider consolidating the two different age cohort systems, and informing the user of the rationale.

The current model construct allows only pair-wise correlations although more complex multivariate dependencies may exist. The Panel noted that techniques of partial correlation coefficients could be used to allow for multiple correlations. The system also could be augmented by allowing default correlations such as low, medium, or high correlations.

The Agency is encouraged to consider providing default shapes of continuous probability distributions for user specified input parameters. The Agency scientists have invested substantial effort in reviewing these data and appear to be in a good position for suggesting default distribution types, while still allowing the user to deviate from them. Having defaults has the advantage of more easily comparing different runs by different users, making the deviation from defaults more readily identifiable and more transparent.

The description of embedding a simple PK model on SHEDS version 3 was not sufficient to allow Panel members to make constructive comments. Estimated absorbed dose is an outcome many would like to see as part of a SHEDS run. Panel members were aware that the Agency has separate groups working on various PBPK models. Instead of developing their own simple PBPK model specifically for SHEDS, the Panel recommended that the SHEDS development staff keep abreast of developments by these other PBPK model teams with the idea of integrating SHEDS with these PBPK models at some future point.

Verifying the precision and accuracy of SHEDS predictions is not described in the materials provided to the Panel for either version 3 or 4 of the model; except for the proposed sensitivity and uncertainty analyses. The Panel felt that before implementing SHEDS or other models in risk assessment, it is critical to understand how the model predicted values compare to the reasonable range of exposures that can be measured in the general population, or in important population sub-groups. SHEDS offers a very flexible platform for users to input variables and parameters and other characteristics that will affect the outcomes of the simulations. Evaluating predictions from SHEDS using exposure measurements is one exercise that will inform the Agency about the accuracy and robustness of the model. Extremely wide or extremely narrow distributions of predicted exposures compared to observed values are suggestive of problems with the model that could lead to questions of its ultimate utility in describing a situation.

b) Please comment on the organization, clarity, completeness and usefulness of the Technical Manual and provide any suggestions for improvement.

The Panel had a number of suggestions on the organization, clarity, completeness, and usefulness of the Technical Manual. In general the Panel felt that the audience and scope should be more clearly defined sooner in the manual. The most important issue identified by the Panel was the need for a comprehensive case study that would illustrate the strengths and limitations of the model.

A second issue identified by the Panel was the need for a standard data set that could be used to orient various model clients; users, reviewers, Agency risk managers, and stakeholders are all important audiences, and a comprehensive example is important for explaining the utility of the model as well as serving as a benchmark for evaluating future changes to SHEDS.

As part of these improvements, the manual needs an early statement as to what the user can and cannot change in the model. A fuller description of the fixed inputs and decisions embedded in the model is needed: particular datasets (e.g., CHAD, the contacts data base for dermal exposure) need fuller explication to achieve user understanding and model transparency. Because the manual lacks a description of the strength of the underlying datasets, risk assessors have no way of estimating confidence in the output distributions. As the Agency does with weight of evidence descriptors in toxicology studies and confidence qualifiers in IRIS, this

would allow the assessor to understand whether exposure estimates are based on studies or diaries that capture the characteristics of the exposure of the population or individuals being evaluated. It also provides an understanding of data gaps and can be used to prioritize resources, both human and financial.

Both the User Guide and the Technical Manual have important, different functions, and these functions should be more clearly distinguished in future revisions. The User Guide has the primary task to inform users about installing and running the model and walking them through common uses. Given the centrality of the CHAD diaries and human activities for developing model runs, it would be extremely helpful if the User Guide presented a typical diary of one weekday and one weekend day as part of the example case study.

The Technical Manual has to serve as a detailed reference guide and should inform on technical issues associated with installing and running the program. It should address the limits of the model, ways to avoid common mistakes and errors, the process of obtaining results, modification and extensions of the program, and interpretation of results. As presently constructed it has no central logic in terms of variable names and definitions: the Panel noted that it would be useful to have a central definitions list that consolidates this information in one location. The document should provide a table detailing representative run times and disk space requirements for a set of population and uncertainty run sizes. This can help in planning for a simulation beforehand.

In general, the Panel felt that the description of CHAD data could be expanded and improved. Although several criteria are listed in the manual, the following factors should also be considered in the random selection of representative activity patterns for an individual: temperature, season, and employment status. Some of the criteria for selecting activity diaries (e.g., diaries that have non-zero sleep time) can be strengthened. For example, selecting a diary with "sleep > 0" is reasonable for a short-term simulation. However, repetition of such a diary for an extended period will likely result in unrealistic scenarios.

The Technical Manual should also list the intermediate files generated, provide relevant details so that users can write scripts for assessing model correctness. Information on restarting a simulation from an interrupted portion should also be provided. The Technical Manual should also provide details on weighting population based on the distribution in a simulated population region (e.g., a county or a state), i.e., sufficient documentation should be provided for sampling from "study specific" population characteristics.

A number of specific recommendations about particular pages in the manual are summarized by the following bullets.

- The Quality Assurance section should be moved to the end of the Technical Manual.
- The captions for the figures and tables should be made more descriptive.
- The descriptions of the probability distributions should be moved to an Appendix.
- Numbering chapters and subchapters is strongly recommended to be used instead of the current, unnumbered heading-based organization. Formulae should be presented using mathematical notation, in contrast to the textual expansion of terms (e.g. F = ma versus Force = Mass x Acceleration), especially for complex equations.. If common explanation is thought to be necessary, an informal presentation could be added. The usage of some

terms in the manuals differ slightly from their definition in statistical literature, e.g., interchangeable usage of the terms "variability distribution," "probability distribution," and "probability vector." These inconsistencies should be addressed.

- The manual addresses at several instances the issues of accuracy, e.g., when sampling from a population and generating the population variability and then comparing the center of the distribution of an exposure endpoint with the extremes of the distribution, e.g., percentiles. A formal definition of accuracy would clarify the discussion of these issues. When addressing the characterization of the variability distribution the Panel noted that the population variability as an outcome of SHEDS is not different from the variability of endpoints investigated in an empirical study (e.g., a large epidemiological study). In both cases principles of statistical data analysis have to be applied. The document should refer to such general principles where possible. Similarly one may require that the exposure modeler set up a study design in a sort of study protocol before starting SHEDS for sensitivity and/or uncertainty analyses. This would then resolve some of the fallacies discussed on page 76. It might be reasonable, to combine the two issues identified in separate locations within the manual (on pages 47 and 75).
- The description of the probability vectors needs to be revised so that it covers also a variable with only two discrete outcomes; i.e., a Bernoulli variable with one single parameter p, the event probability, e.g., to parameterize existence of pet or lawn.
- The associated text on page 49 would benefit from more formal probabilistic and statistical terminology. If the authors think that would be too formal, that part could be moved to an appendix. Pages 51-54 on the distributions should go to an Appendix with mathematical formulae for the distributions described in a formal mathematical type, referring to the statistical literature, e.g., that of Johnson, et al. (1994, 1995, and 2004).
- On page 24 there are instances where no dust is contacted. The basis for this decision is unclear: is it for model internal consistency or based on empirical data?
- The manual needs to clarify if and how the model addresses region-based seasonal events, such as spring-time applications, in calculation of the annual exposures.
- There is insufficient description to determine how the calculation or estimation of application rates is determined. The manual should clearly state if this is proprietary information, and if so can they be updated over time with empirical data? Otherwise, the information should be presented in detail.
- Because this is a residential model, and because residents of multi-family dwellings have less control over their pesticide applications, the model should consider a multi family housing designation as a person level variable and address whether it is correlated with other microenvironments or parameters.
- The Technical Manual and the User Guide provide a wealth of information and many details on SHEDS' construction and structure, on the approaches used and on their programming. One Panel member noted that the structural description of SAS procedures as given in SAS Manuals could be a template for the description of the various procedures of SHEDS.

Question 1-3:

a) Please comment on whether the annotated code is sufficiently clear such that the algorithms can be followed and understood.

b) Please also comment on whether these algorithms are technically correct and consistent with the descriptions provided in the Technical Manual.

The Panel members agree that the code is clear enough and that the algorithms can be easily followed. Adequate and well described commenting in the code (about 30% comment lines out of a total of about 8500 lines of SAS code), along with a logical organization are the positive features of the code. Logical separation of data and code, a sound practice, has been followed in the SHEDS development. The code base is of reasonably high quality, especially considering the relatively low number of users that are likely to make changes.

Several suggestions have been made for further improvement in the code and documentation. The need for even higher standards is based on the assumption that detailed external code reviews and adaptations by the general scientific community are desired for SHEDS, especially given that there is an anticipated expansion of the number and types of chemicals, as well as the user base.

In terms of code readability, it is desirable to have adequate cross-referencing between the Users' Guide, Technical Manual, GUI, and the actual code. This would allow a user to check the underlying code for a given operation without reading the code file that contains multiple modules. This can be accomplished by providing hyperlinks in the electronic versions of the documents, and by providing the annotated code separately (e.g., in the form that can be easily used in a "code browser", or in the form of HTML with the comments as "tool tips"). Furthermore, separating the utility modules from the core modules in the current SHEDS code is relatively straightforward and recommended as this will enhance code readability.

The naming of the variables within the SHEDS code can be significantly improved. Currently, the code uses cryptic names such as washprob, absr_gr, absr_gm, bioavm, and bioavr, with subtle changes in variable names for different quantities – this can potentially lead to difficult to notice errors. The code can be improved by providing more informative names, and through the use of a "nested" or hierarchy of variables (e.g., through the use of objects/classes in SAS). The variable naming can also be improved by using a consistent naming scheme. For example, the current version of the SHEDS code uses variables such as washprob, has_lawn_p, timeofday_indoor, and re-entry_indoor, all denoting probabilities.

The code changes across different versions should be adequately documented. In fact, it is desirable to have an archive of major versions of SHEDS available to interested researchers so that they can use simple text differences between different versions of the SHEDS code or individual modules. The technical documentation and code comments do not mention issues regarding the backward compatibility of the current version with earlier versions, which is very important. Additionally, a bug tracking/reporting system is also desirable for users and developers.

The technical documentation or code comments should provide details on the intermediate variables in the SHEDS simulations, and on how a user can enable these to be stored and analyzed. Experienced users can then easily obtain summary statistics, improved visualizations, and other desired outputs.

The Panel members believe that the algorithms appear to be technically correct, with the caveat that it is impractical to perform a line-by-line review and verify technical correctness of the code by the Panel. However, some suggestions for improving algorithms as well as for enhancing the consistency between the Technical Manual and code were made. These include providing reproducibility of simulation results (by reusing a random "seed" if it is provided by the user), addressing or acknowledging issues associated with running stochastic programs on multiple machines and associated potential correlations in the random number generation process.

A minor comment from the Panel members includes documenting the descriptions of probability distributions used in the Technical Manual and User Guide, because the distributions are often parameterized in different ways, and it is important to clearly describe how these distributions have been parameterized in SHEDS.

Issue 2: Technical Aspects of Planned Methodologies to Extend SHEDS-Multimedia version 3 (aggregate) to version 4 (cumulative)

Question 2-1:

Please comment on the technical aspects and usefulness of the planned methodology for extending SHEDS-Multimedia version 3 to address multiple chemicals in version 4.

The logic of extending the application of SHEDS from single to multiple chemical assessments is plausible and warranted. However the nature of such proposed work would be complex, time and resource consuming, and the process itself may pose significant challenges. The Panel expressed several concerns about the planned methodology

As the document notes, the move from aggregate to cumulative involves relatively few changes to the code/structure/interface. The Panel believes this is what the previous SAPs envisioned when they recommended stochastic models for estimating cumulative risks. Overall, the concept of going from aggregate to cumulative as described is useful. One important feature that is not mentioned in the document is the importance of input from risk managers and Agency scientists about the chemicals that are subject to cumulative assessments and the process used to make these determinations. Once this decision is made the focus shifts to model structure and outputs, and the Panel focused its comments on mass tracking, product co-occurrence, pesticide metabolites, programming issues, scenarios, and absorption and bioavailability.

The process of tracking mass is unclear in Version 3, and many Panel members felt that future versions of the model should track mass in media because, for example, it makes it possible to track the mass that is lost through touching and subsequent washing of the hand, and allows tracking of mass in specific media over time. The Agency is encouraged to explore the feasibility of this mass balance step. The Panel noted that, in the document, the definition of "carrier" in SHEDS is somewhat overlapping with "media." For example, air (either indoor or outdoor, or in the treated or untreated rooms) is considered a carrier and a medium in some places in the manual. The manual is also unclear as to why "residue" itself would be considered as a carrier. Residue has to be associated with or adhered to the carrier before it can reach the simulated individuals in any of the media. As stated in the documentation, "When a person comes into contact with a contaminated MEDIUM, a certain amount of one of more carriers will

be transferred onto or into the body" (page 7, SHEDS Version 4 Planning document, June 12, 2007 version). Care must be taken to correctly define and use the terms "carrier", "media" and "residue."

A scenario (a child playing with uncontaminated soil) provided by the Agency further complicated this matter. As stated in the documentation, by taking into account such activity, the mass of the carrier, which is soil, will change on the hands, the body, and possibly in the GI tract. However, if the soil is not contaminated, the change of soil mass would not affect the model outputs. Since each diary is being considered as an independent event (with some autocorrelation), even if this child might come in contact with contaminated soil on a different occasion, the mass of the soil that will be transferred this time would not affect the mass of soil that is uncontaminated.

As stated by the Agency, the run-time and the storage space for the model output may significantly be hampered by the proposed approach, if all the chemicals in the formulation are included in developing the variables. However, if the intent of developing SHEDS version 4 is to cumulatively assess exposures to a group of chemicals that has the same toxicological endpoint, the concern about resources may not be warranted. One Panel member noted that it is uncommon to have a single product containing more than 2 active ingredients that cause the same health concern. So instead of n=10 as proposed for use in the algorithms in version 4, the actual sample size might be no more than 3.

The document available for the Panel to review is brief and lacks detail on a number of issues. The Panel noted that the current product-related co-occurrence system may be problematic if there are substantial future changes in product formulation. This also implies that intentional product co-occurrence is more important than unintentional "random" product co-occurrence, an assumption that needs close inspection. The document was unclear, but during the presentation the Agency indicated that it will consider addressing both multiple products that contain multiple chemicals and multiple chemicals with a common mode of action that are not necessarily present in a single product. The Agency should explore using different data sources for assessing multiple chemical co-occurrences: for example, pesticide use data could potentially be useful for assessing multiple chemicals for different purposes (e.g., garden versus indoor versus lawn). In addition, the considerations for co-occurrence for residue in food is another complex issue to be explored (e.g., number of pesticides found in one food commodity versus multiple chemicals in a salad) that is not fully explicated in the document.

The Panel noted that the ability to model compound metabolites should be included in the multiple chemical model, but this will require development of a module to quantify metabolite production. One Panel member noted that one important lesson from the Agency's CTEPP (Children's Total Exposure to Persistent Pesticides and Other Persistent Organic Pollutants) Study is that tracking of metabolites is necessary for mass balance. In the absence of full accounting for metabolites, evaluation of model performance will be limited to a small number of compounds, such as PCP and 2,4-D, for which the urinary biomarker is the parent compound. The current document states that using a single efficiency for washing removal of soil/dust from the skin regardless of compound is an advantage of the multiple chemical model. Uniform washing efficiency may be reasonable for soil or dust-borne chemicals. However, a single removal efficiency is also apparently assumed for "residues" (i.e., for free chemical on the skin).

Since chemical washing efficiency is likely to be tied to the physical-chemical properties of a given chemical, this could be a disadvantage. Relatively lipophilic compounds are less likely to be efficiently removed by washing than relatively water soluble compounds.

While this move from aggregate to cumulative assessment provides an opportunity to use an "object oriented" programming approach for encapsulation of variables, it is also possible that users can be overwhelmed by the number of parameters that they have to provide. Therefore, the Panel envisions that GUI modifications may be substantial (e.g., moving from pesticide-specific interface to general multimedia interface) even though the underlying code changes may be relatively minor. Therefore, it is desirable to have one underlying code base and different interfaces. One Panel member noted that the memory requirements of such a system may be large due to the increase in the number of chemicals to be simultaneously modeled. The temporal resolution of the simulation will potentially be dictated by the fastest simulation time and the lowest time scale to be modeled.

An additional feature that is desirable but not included in the written plan is scenarios for home applicator exposure. This was presented at the meeting as one potential added component for version 4, and includes both professional and home applicators, and the Panel felt that this was an important addition that needs careful consideration.

One Panel member noted that the description of absorption efficiency and bioavailability for non-dermal pathways in the Technical Manual is unclear, and therefore the state of science cannot be evaluated because equations and model inputs are not available. The Panel cautions that, like dermal absorption, GI and lung bioavailability estimations and assumptions are dependent upon factors such as medium, particle size, material thickness on exposure surfaces, and other factors that need careful attention in a multi-pathway, multi-media, and multi-chemical model.

Question 2-2:

Please comment on the technical aspects and usefulness of the planned methodology for incorporating a fugacity-based source-to-concentration module into SHEDS-Multimedia version 4. Does the Panel recommend additional efforts with the fugacity module (e.g., modeling more realistic multi-room dwellings) given available information?

The Panel agrees that chemical redistribution following residential pesticide application is scientifically well established, that the fugacity approach is well founded and that it is an appropriate method for addressing this phenomenon. Providing this methodology as an option within SHEDS, without removing any of the existing chemical fate approaches, can only make the code more valuable even if only a subset of potential users avail themselves of the fugacity option. The Panel also noted that a fugacity module, based on the work of Bennett and Furtaw (2004), has already been developed and programmed in SAS. Therefore, no question regarding relative value versus resources required is presented by mere inclusion of a fugacity module. The Panel further believes that the fugacity approach is well suited to treatment of multiple chemicals and is therefore compatible with other development goals for SHEDS.

The Panel expressed less unanimity with respect to the actual likely utility of the fugacity option and expressed a desire to see example calculations that might better illustrate the performance of

this option. Only through sensitivity analysis can the question of value be determined. In particular, the effect of time scale being simulated should be examined. The Panel also asked for clarification regarding whether the fugacity approach will be limited to indoor environments.

With respect to the second part of the charge question, the Panel agrees that simulation of a multi-room house is certainly feasible mathematically. Mixed opinions were expressed regarding the value of moving beyond a two-zone (treated/untreated) model. Clearly there is some point at which lack of data necessary to describe more complicated systems will prevent derivation of added benefit by increasing complexity. Data needs would include both architectural features and room-specific pesticide application practice and human behaviors. Behavioral data are often a limiting factor in using current models. The Panel therefore suggested that EPA start with a simple (two-zone) version and more thoroughly investigate and document performance at that level before considering multi-room systems.

The Panel notes that EPA proposes to evaluate the fugacity module using chemical transport data from a 1-house study. While this is a reasonable place to start, the Panel encourages more extensive testing including application to cases in which both environmental data and biomonitoring data are available. One member expressed the opinion that fugacity modeling could lead to greater appreciation of the significance of surface contamination on human exposure indoors.

Under this charge question, there was discussion of whether SHEDS should account for removal of contaminant residues from environmental compartments via absorption by occupants. It was suggested that EPA compare predicted mass absorbed to assumed mass applied for various scenarios to provide more information on this topic. Although the term did not arise explicitly in the discussion, in effect the Panel suggested that EPA include routine computation of Intake Fraction (Bennett et al., 2002a,b) in SHEDS output.

Question 2-3:

Please comment on the technical aspects, potential utility, and added value of the planned methodology for longitudinal diary assembly in SHEDS-Multimedia version 4. Does the Panel believe that this new method will create an assemblage of diaries that better simulates reality and provides more accurate estimates of exposures related to within-individual time-activity patterns? Please suggest procedures and/or longitudinal data which could be used to select factors (the "D" factor intra-class correlation coefficient, and the 1-day lag autocorrelation) or refine/ evaluate this method in SHEDS.

The Panel appreciates the clarity with which the approach is presented in the paper by Glen et al, 2007, and was encouraged that the work underway to analyze for patterns in the data will possibly shed light on what may be reasonable key diary variables and their bounding statistics.

The Panel believes that the concept behind the new longitudinal diary assembly method has scientific merits, and the technical aspects of the proposed method based on the D-statistic and 1-day lag autocorrelation is sound. However, the application of this method in the interest of predicting aggregate human exposures to environmental chemicals may be somewhat limited. This is particularly true for activities that may lead to potential exposures that are either difficult

or implausible to track as continuous "key variables." For example, it is useful to use the D-statistic for tracking how long a person spent outdoors for studying human exposure to air pollutants (like the "Southern California Ozone Exposure Study" cited by the Agency) because the time spent outdoors is correlated with the amount of ozone being inhaled by the subjects. However, it might not be as feasible to implement the D-statistic approach in compiling longitudinal exposure data sets, such as the frequency of hand contact on a contaminated surface, when multiple factors need to be considered before the exposure scenario can be confirmed. Such a limitation is inherited from the concept of incorporating the within and between-subject variation in order to better control the longitudinal properties of the final assembled diary. The key variable(s) also needs to be ranked in order to generate the x-score. Those theoretical requirements constrain the utility of this proposed method for predicting longitudinal exposures in which inhalation exposure is not the predominate pathway. In other words, unless the longitudinal diaries do not encompass a great number of human micro-activities (such as potential exposures other than inhalation) in which time spent in the area where the contaminant is present is the only and the ideal key variable, the incorporation of this new longitudinal diary assembly method in SHEDS version 4 may not add significant value. Based on Glen et al. (2007), it appears that there is also a possibility to choose more than one key variable for creating a longitudinal diary and therefore the Agency is encouraged to explore this possibility.

The Panel believes that it is likely that the current activity database (CHAD) may not contain enough diaries to allow random draws using the new longitudinal diary assembly method. Since a significant number of CHAD diaries were excluded for various reasons, decision criteria that further reduce an already small dataset might be problematic for accurately determining the within-subject variation. Furthermore, as stated in the documentation, the two most important pieces of information, chemical usage patterns and contact possibilities, are not included in the diary database. These concerns dampen the enthusiasm of the Panel for implementing this new longitudinal diary assembly method in the upcoming revision of SHEDS.

The Panel concurs that including a lag autocorrelation function (currently set at 1-day) would be necessary to eliminate repetitive activities from the longitudinal diaries that are likely unrealistic (such as pumping gas into an individual's car on successive days). However, the Panel suggests that the Agency's efforts should focus on behaviors relevant to potential residential pesticide use, such as frequency of gardening activity, rather than examples related to air pollution exposure assessment.

The Panel does not doubt the Agency's assertion that improving the construction of longitudinal diaries can lead to more reliable exposure estimates. But use of realistic and well documented scenarios in the model is, in the Panel's judgment, as high or higher in priority as further refinement of the diary assembly and autocorrelation methods. For the explicit purpose of estimating the individual and population exposures, it may be more constructive for future versions of SHEDS to suggest defaults for key variables and allow for deviation from these defaults as needed by informed users to display the strengths and limitations of the model.

One Panel member wanted to see how the longitudinal diary development using the D-statistic and the autocorrelation would impact exposure prediction. This individual felt that the Agency should either initiate longitudinal behavioral data collection (such as dietary consumption or time-activity surveys) within specific population subgroups, or look for other data sets that have information that would be better suited for this proposed method and therefore alleviate the

limitations inherent in using the D-statistic. As presented by the Agency, there are several data sets that have the potential to be used for this purpose, and the Panel recommends that the Agency contact the PI of those studies to obtain permission to access their data sets.

Question 2-4:

Please comment on the technical aspects and usefulness of the planned methodology for utilizing Sobol's method for sensitivity analysis in SHEDS-Multimedia version 4, and whether Sobol's method would be a useful supplement to the existing sensitivity analysis methods used for the SHEDS-Multimedia version 3 model.

Several Panel members expressed concern that the required number of simulations (2N + 2) for sensitivity analysis involving N parameters appears to be very low. For example, if a single parameter has 10 levels of values, doing a sensitivity analysis with just 4 simulations (2*1+2) is impractical. The Agency's view that a single sensitivity run can be considered to represent multiple "individual-days" of simulations is not correct in this case, because a single variability run will ultimately produce a single population metric (e.g., the 95^{th} percentile for population exposures).

It appears that the proposed use of Sobol's method will allow the computation of "local sensitivity," without addressing the corresponding uncertainty. Global sensitivity analysis, on the other hand, allows the estimation of sensitivity as well as uncertainty.

The Panel thus wonders whether the new sensitivity analysis method is needed in this case. Based on the classification by Saltelli, et al., (2005), Sobol's method can be seen as a variant of the Fourier Amplitude Sensitivity Test (FAST), which is a variance based approach. Since ANOVA is another commonly used variance based approach, the Panel's observation was that such well known methods should be evaluated before using relatively less known techniques such as Sobol's method. One Panel member noted that this approach appears to be reinventing experimental design. Many of the Panel members were not convinced about the utility of the Sobol's method for SHEDS, and were also perplexed by the complexity of the method.

Since Sobol's method has not been used with a large scale probabilistic model before, the Panel urges caution in using the method. Furthermore, the number of changes to the model structure and code (e.g., by making a majority of the SHEDS code deterministic, with random parameters provided in one module) appear to be large, further reducing the usefulness of incorporating this methodology in SHEDS.

However, some Panel members felt that since the Sobol method is an accepted method for sensitivity analysis it should be included, since in general it is a good idea to provide new techniques as alternatives to variability and uncertainty analysis.

Question 2-5:

a) Please comment on (and prioritize, as appropriate) the technical aspects and usefulness of planned changes to the SAS code and GUI for SHEDS-Multimedia version 4 that are listed items in Section 5 of the above-referenced background document.

b) Please comment on any additional modules, features, or capabilities that the Panel feels should also be high priorities for the next version of SHEDS including issues associated with the code, user interface/user friendliness, input, and output/output display. Are there modules, features, or capabilities of other human exposure models that should be considered for inclusion in SHEDS-Multimedia version 4 (e.g., simulation of individuals; longitudinal diary assembly)?

The Panel was not comfortable choosing from the list as a dialogue with the developers is needed to better inform the Agency's priorities. As a first suggestion the Panel felt that the developers should do everything that is simple and does not involve a major reorganization of the code. At the same time, a lot of effort should not be placed into options that will not make much difference in the overall conclusions drawn from the model. Most Panel members thought that PBPK modules should not be a high priority and should remain independent modules from the SHEDS model output.

The Panel felt that the Agency should get more experience with the application of this model before getting into a major reorganization of the code. It is important that the developers anticipate future data sets, especially longitudinal, and try to keep the code scalable. Autocorrelation can have an important effect on the tails of the exposure distributions. Anything that helps the user understand uncertainty and variability is important. Several Panel members emphasized that improved graphics were important for wide model comprehension and acceptance. They also felt it was useful to add utilities to make the model more accessible and useful, e.g., code to review input values and note extreme or illogical combinations, or a script generator to run designed experiments on the model. This is similar to sensitivity analysis but could be more general. The user could ask a series of questions about what factors and interactions are of interest, and then plan and implement a fractional factorial or response surface design to test hypotheses. Prior to adding additional functionality (modules, features, or capabilities), sensitivity analysis with the specific modeling assumptions should be conducted. This involves defining the population exposure metrics, and performing sensitivity analysis. It is possible that some proposed features can be discarded after this analysis, based on lack of sensitivity of the exposure metric to the new feature.

The Panel liked the idea of users being able to re-parameterize distribution mean, variance and shape so they can conveniently move between distributions without having to refer to statistics texts (e.g., Johnson, et al., 1994, 1995, and 2005) for the parameter definitions. Experience has shown, however, that unless a distribution has an extreme right tail, the choice of distribution does not make much difference.

The Panel did have a number of suggestions for improvements, and issues that the Agency should not put as high priorities. As additional features, the Panel felt that the model should include home applicators. Some users of the model will not be constrained by EPA practice and will want to model non-compliant application, i.e., not following the pesticide label. It is important that the model provide an "application programming interface" definition that describes the format and the type of intermediate data, in a simple and easy to use manner. This will lead to the following advantages: a) users and developers can independently create input/output analysis modules, provided there is adequate documentation for developing such tools; b) this will overcome current issues with the quality of plots and graphs; and c) provide facility to input empirical distributions for specifying uncertainty and variability; d) plus provide

sufficient number of examples, so the novice and expert users can utilize the system to the fullest extent. The Panel felt that it would be useful if SHEDS provided better context specific help files, the ability to calculate margins of exposure (MOEs), and other options for selecting and matching CHAD and food consumption diaries (e.g., by age, gender, season, weekday, region, race, METS/caloric intake).

Issue 3: An Update on the Development of the SHEDS-Dietary Model

Question 3-1:

Eating Occasion Analyses.

As described in the SHEDS dietary background document, the timing information available in CSFII can be used to model food and indirect water intake throughout the day. With the ability to incorporate the timing of eating occasions in dietary exposure assessments, it is possible to assign either the same residue or a different residue to foods consumed on multiple eating occasions. In certain instances, the former seems logical (e.g, consumption of leftovers) while in other instances the latter appears appropriate (e.g, hash browns at home for breakfast and fried potatoes away from home for dinner).

Please comment on developing simple decision rules - as described in the document - for assigning residues to commodities eaten on multiple eating occasions.

The Panel noted that the Agency's document did not include a clear list of rules for review. For future presentations, a decision diagram or flow chart should be used to describe the decision rule process and allow for substantive critical review.

This issue question is relevant to acute exposure scenarios. The Panel noted that there are two opposing default rules, i.e., assigning either the same or different residues to the same foods consumed on multiple eating occasions in a day. Realistically, the possibility exists for residue linkage between eating occasions. The same batch of foods can be consumed in consecutive eating occasions, such as leftovers. Raw agricultural commodities can also be purchased in quantity and eaten over a period of time, such as a batch of berries or a bag of apples.

According to the Agency's presentation, the current practice by the Agency is to assign the same residue to food commodities eaten on the same day. Theoretically, adding inter-day variability in residues for the same commodity by assigning different residue levels is likely to increase the variance of the distribution, but possibly more so in the lower tail than in the upper tail. The median is likely to shift toward lower exposure. On the other hand, assigning the same residue level for multiple eating occasions in a day will likely result in a higher upper tail and thus capture the possibility of high end exposure, especially if there is a spike at the high end of residue distribution. Since the upper tail is where the main interest lies in risk assessment, it is unlikely that there will be significant gain in changing into random draws of residues for multiple eating occasions. Nevertheless, the Panel understood that the Agency may wish to offer different yet viable options in the SHEDS model which tracks time sequence of event and activities in a much finer increment than a 24-hour day.

The Panel discussed the added value of fine-tuning this residue assignment assumption. Assigning the same residues to leftovers and different residues to the same food commodity, but served as different dishes, appears to be too labor-intensive and subject to bias. Such level of refinement may not have added value especially when resolution in the food intake diary and residue data (each item in composite data does not contain the same level of residue) is lacking. In addition, the Panel was concerned that making further assumptions without proper support data may introduce additional uncertainties in the exposure analysis. For example, the only way to determine whether the same food forms eaten at lunch and dinner are separate dishes or the same dish as leftovers for dinner may only be mere speculation. Thus, it is not known if assigning different residue values for 5,000 iterations yields a more realistic residue distribution that is consistent with the stochastic nature of the SHEDS model.

As to setting rules for assigning residue levels, the Panel was in general agreement that simple rules have the advantage over complex rules for having clarity, uniformity of application, and transparency in operation. From the stand point of conserving resources by reducing unnecessary complicated steps in data analysis and interpretation, a simple set of rules in this case can be somewhat akin to a tier approach that is informed by a series of stepwise refinement analysis, progressing toward increasing complexity only when supported by data and the need for fine-tuning the analysis.

For this issue, the first consideration is whether dietary exposure constitutes a major contributor to an overall multi-media exposure that exceeds a regulatory threshold (e.g., having lower than desirable MOE). When dietary exposure is the focus for refinement, high contributing commodities are identified in the initial assessment and the impact of link-day residue for these commodities can then be assessed.

Alternatively, a more generic decision process may begin with the consumption-residue pairs that can potentially be major contributors to the exposure. In this evaluation, the Agency may focus the initial analysis on foods most frequently consumed (e.g., FDA's top 10 list). Commodities having a fairly high ratio between the high end (e.g., 90th percentile) and the central tendency (e.g., mean or 50th percentile) residue levels may be the first group for investigation. This type of summary data may already be available (e.g., similar to Appendix M in the 2005 PDP summary report). Once the potential high contributing commodity-residue pairs are identified, the possibilities for repeated eating occasions could then be further investigated. Another generic decision process may be to evaluate how often people eat the same foods within a day or on different days based directly on food intake diary, e.g., CSFII data. With either approach, the initial focus can be placed on raw agricultural commodities unless food processing would dramatically alter the residue contents.

The timing of food consumption within a day becomes more prominent as the Agency prepares to include dietary exposure estimates into the time-series exposure analysis of SHEDS, and especially when simulated exposure from SHEDS is ready for input into PBPK modeling. With chemicals that are non-persistent (e.g., metabolized and excreted in a matter of hours), retaining the timing of consumption information in the diary will greatly facilitate the subsequent PBPK analysis.

The Panel cited several potential sources of nationally-representive time use and activity pattern data that the Agency should be familiar with as it considers enhancements to within-day modeling of exposures. The **Child Development Supplement (CDS)** to the Panel Study of Income Dynamics, http://psidonline.isr.umich.edu/CDS/time_diary/readme.html, collected time diary data from the children or from the children with input from their caregiver, or, for very young children, solely from a caregiver report. For each child, a random school day (Monday through Friday) and a random weekend day (Saturday or Sunday) was recorded in a diary. The diaries were originally mailed out and the difficulties respondents had in completing them led to follow-up telephone interviews in which trained interviewers recorded most of the diary answers given by the caregiver/child.

The **American Time Use Survey (ATUS),** http://www.bls.gov/tus/#overview, sponsored by the U.S. Bureau of Labor Statistics measures the amount of time people spend doing various activities, such as paid work, childcare, volunteering, and socializing.

In addition, the University of Maryland maintains a large archive of recent and historical data sets on inidvidual time use and activity patterns. These data sets and documentation may be accessed at the archive's URL: http://www.webuse.umd.edu/ .

Question 3-2:

Please comment on the 8-record approach in SHEDS-dietary and the selection of age group, gender, season and day-type from which to create the "diary pools". What other approaches does the Panel recommend? Can the Panel suggest any "bounding approaches" that may - based on knowledge of actual eating patterns - provide upper and lower limits for longitudinal exposure estimates (e.g., yearlong consumption of the same diary throughout the year vs. random daily selection of CSFII diaries).

The majority of the Panel is convinced that given the data and analysis presented by the Agency, it is not sufficient to construct the longitudinal dietary consumption pattern based on the 8-day eating occasions. A recent paper published in Environmental Research (Givens, et al. 2007) concluded that the variability of daily average dietary intake estimates, as well as the within and the between-child variability, is significantly decreased when individuals are followed longitudinally. This finding suggests that a longitudinal dietary survey (a minimum of 6 consecutive days' dietary consumption in each of the 4 seasons) would be adequate to be representative of one-year dietary consumption pattern at the individual level and would offer substantial improvements for assessing dietary exposures compared to a standard cross-sectional study design with repeated sampling, like CSFII.

The Panel recommends that the Agency should conduct analysis similar to those already performed with subsets of the CHAD database, and see the pattern of D-statistic that may be useful for longitudinal implementation. Essentially, there are two approaches to this proposed investigation. One is to use longitudinal measurement data to derive some understanding of patterns, and the key variables that may arise from these types of analysis. The other is to do iterations of analysis based on the variables that may have high likelihood to be key variables for characterizing the dietary patterns, and create D statistics from these data. These analyses may be the beginning of understanding of the important features in longitudinal data use and exposure assessment. Some Panel members suggested that the Agency pay attention to the tails, as well as

the median of the distribution in terms of consumption patterns. The Agency should treat with skepticism apparent good prediction of upper tail without concurrent good prediction of median. Such a situation assumes that the tail was accurately estimated in the absence of accurate prediction of the distribution or sources of variance.

For the bounding estimates, one Panel member suggested that to determine the largest contributions by various food types it would be useful to examine the highest level of food consumed by individuals on the bases of age and gender in the CSFII data set. This bounding might include all of the same foods – for example, the three-year old who only eats apples and macaroni & cheese - on the basis of consumption (g/body weight/day). In the process, it should be made certain that for the bounding estimates, the caloric contributions/day are consistent with the records.

In conclusion, the Panel generally agreed that significant value would likely be added to the SHEDS model development by collecting longitudinal dietary consumption, as well as individual behavior/activity data, particularly the activities at the microenvironment level from a small group of study participants stratified by age and gender. This study could be conducted in a way so that the data collected from this small study can be used for establishing a larger longitudinal data set using the D-statistic approach without introducing additional substantial variation and measurement error. Technology (both the survey tools and data download/upload capacity) has improved in recent years to facilitate conducting such a longitudinal study without significant investment in resources and time. Dietary consumption information, as well as time-location-activity, can be collected remotely using the current wireless Internet connection where the bulk of the data can be processed in a timely manner (Lu et al. 2006). Other Panel members have suggested the possibility of nesting a separate study within the proposed National Children's Study, and that this option should be considered by the Agency. Several Panel members raised concern about the dependence on the CSFII database to provide reliable and timely data to represent the US population in terms of current dietary patterns regardless of sophisticated statistical analysis.

Question 3-3:

Please comment on the advantages and disadvantages of providing an option to use the Bayer DWCS data in SHEDS-Multimedia v. 4. Please include in your comments any statistical concerns or issues associated with the design and conduct of the DWCS study.

EPA Clarification of Charge Question 3: EPA recognizes that it has provided each of the Panel members with only background information on the Bayer DWCS as part of the document "An Update on the Development of the SHEDS-Dietary Model" and has not provided Panel members with the Bayer submission itself. The intended purpose of the above charge question was less to receive specific comments on the conduct and analysis of the Bayer DWCS, but rather to introduce this as a possible source of data and obtain the Panel's thoughts and ideas on the conceptual utility of this kind of study, in principle. More specifically, we are looking for advice from the Panel on ways studies conducted in this way could be interpreted and potentially used and any suggestions the Panel might have on ways data collected in this manner should be examined and interpreted. EPA has, to date, only performed exploratory analysis of the data provided by Bayer and has not yet formally reviewed this submission or used it as part of a regulatory decision. EPA will use

the thoughts and ideas presented by the Panel in its discussion of this topic and, if warranted at a later date, may present the study and our analysis to the Panel, along with our proposed use of the study as an available option in the SHEDS model.

With the Bayer DWCS water consumption data, EPA has the opportunity to change and possibly improve the drinking water model in the SHEDS model. The current SHEDS approach is to assign to each individual a total amount of water consumed each day according to the daily water consumption amounts in the CSFII diaries and then distribute this amount in 6 equal amounts to 6 time points spread evenly through the day (method 1a). The FIFRA SAP (2005) suggested allocating this same total water amount to five events – with three occasions during meals (25% of total per occasion) and two occasions in between meals (12.5% of total per occasion) (method 1b). The data from DWCS could be used to either assign new total daily consumption amounts, use the drinking events data to create new allocations patterns throughout the day or both. EPA is exploring an approach that allocates total direct drinking water amount using CSFII data and assigning distribution throughout the day using a randomly selected DWCS diary matched to sex, age, season and total reported amount for the corresponding CSFII food diary event (method 2). A summary of the Bayer DWCS study was made available to the Panel on which the discussion and comments are based.

In attempting to answer this question the Panel raised questions of its own.

- *Do DWCS drinking water patterns differ very much from the current implementation?* The limited analysis made available to the Panel indicated that most individuals have 6 or fewer drinking water events during the day with a modal value of 3 (Figure 4). Distribution is very uniform throughout the 16 hours measured (Figure 5). This would seem to suggest that the approach currently implemented in SHEDS is representative and there is little here to support implementing the 2005 FIFRA SAP recommended allocation.
- *Can a minor modification of the current implementation move the drinking water pattern closer to the DWCS without having to fully implement the DWCS into the simulation?* Allowing the number of drinking events in a day to vary would be an obvious first modification. More analysis would be needed to see if direct drinking events are correlated to eating events, with the eating events occurring more uniformly throughout the day.
- *Will allocation of direct drinking water consumption via an empirical use of the Bayer DWCS data, say via some form of Bootstrap sampling methodology, produce results that are markedly different from a modified fixed drinking water pattern?* Preliminary simulation results seem to suggest that it does not (page 28 of SHEDS diet model modification document). An analysis by one Panel member indicated that the DWCS option would provide similar, but slightly higher peak exposures at the per capita 99.9th percentile than the fixed option with 6 equal allocations.
- *Can individuals be grouped into drinking water pattern types that are internally homogeneous and externally varied, allowing individuals to be selected to a drinking water pattern "type"?* No data were presented to answer this question. Knowing what type of drinking water pattern produces highest exposures might be more useful than knowing which individuals drank the most. While the DWCS has measures on drinking water amounts on 18 hourly intervals, it is probably not necessary to model drinking pattern into 18 hourly intervals.

- *Can "typical" drinking water patterns be discerned from the data that would lead to pattern types that would be easier to model and interpret?* Typical at work patterns, typical at home patterns, typical patterns for children, etc... for example, Figure 5 in the Agency's document, "An Update on the Development of the SHEDS-Dietary Model," demonstrates that consumption is nearly uniform during the day. Is this true for all individuals or is it uniform because it represents a mixture of different patterns (such as, meal drinkers and meal never-drinkers in equal proportions)?
- *Can the DWCS database be properly weighted to produce population relevant information?* This question was voiced by a number of Panel members who were concerned about a number of the sampling aspects of the study. In particular, there is the potential for bias in study results because of the low response rate and high fraction of incomplete diaries. An attempt should be made to compute sampling weights to better match study results (the demographics of study responders) with the (demographics of the) US population. In particular, the Panel wondered which age, race, ethnicity, geographical or other groups were not adequately sampled or whether some of these groups are not represented in the dataset at all?
- *Because exposure from drinking water sources is a very small fraction of total exposure, is there a need to further model drinking water in the SHEDS model?* A number of Panel members felt that because of its low contribution to total exposure, further modeling of drinking water consumption would not noticeably change model results. This view was also mentioned in the previous SAP consultation on this topic. The Panel has seen a number of cumulative risk assessments that have the contribution from drinking water being one to two orders of magnitude below diet levels. Still, other Panel members felt that in some exposure scenarios the contribution from drinking water could represent a large fraction of total exposure and in those cases different water allocation models could produce quite different results. The importance of water consumption as a pathway in chemical exposure will depend on the properties of the chemical in the environment and presence in the drinking water supply (groundwater or surface water) in addition to the consumption rates.
- *Is it true that all water consumed in a day must have the same concentration?* A note in Section 3.5 suggests that all water consumed in one day has the same concentration, regardless of source. This may have been true in the past but can this really be the case when many individuals drink water from multiple sources each day (bottled, tap, filtered, etc)? One situation mentioned was the scenario where contaminants in the water occur only at one specific time in the year (a pulse) with concentrations the rest of year being zero. Another situation mentioned is where the water used in cooking food is the contamination pathway. It is possible that factors related to drinking water source, municipal versus private, ground versus surface, chlorinated versus not, may influence consumption rates. Individuals may not be fully aware of many of these factors, but they will know where the water they pay for comes from and this may also affect consumption. Finally we should not lose sight of the information available on residue data in drinking water as available in the Office of Drinking Water.

Previous SAP consultations on this topic remarked at the lack of adequate information on drinking water consumption patterns to properly analyze the fraction of exposure from this source. A good portion of the Panel do not think that the issues related to exposure from drinking water sources are likely to be less in the future and would like to see drinking water's impact examined in more detail. Issues such as water source, use in cooking, temporal changes in

contamination levels, methods of disinfection, etc., all seem to need further discussion and examination. Water consumption patterns in the US seem to have changed since the CSFII data were collected and, due to the wide use of bottled water, water is now a consumer product. It is hoped that the DWCS data, once properly peer reviewed, will be able to answer some of these questions. The current PDP data on finished water can be used in the reanalysis for drinking water contribution when needed.

Question 3-4:

Sensitivity and contribution analyses are a routine part of OPP risk assessments. These analyses help inform the risk manager how exposures may change when certain model inputs are modified. These modifications to the model inputs are typically performed "one at a time" to permit isolation of the effect. In a typical risk assessment, all the dietary consumption data (i.e., reported CSFII diaries) are used along with the best available pesticide residue data. OPP risk assessors specify a sufficiently large number of Monte-Carlo iterations such that exposure estimates are stable with respect to the random seed.

The Agency has not conducted formal quantitative uncertainty analyses. The Agency presented a simple bootstrapping procedure for conducting uncertainty analyses, utilizing only a subset of the consumption and residue data inputs. That procedure was designed to provide some insight into the question, "How much better would our exposure estimates be if we had more data?" by conducting the uncertainty analysis in the other direction.

a) Please comment on the scientific soundness and utility of the proposed bootstrap uncertainty approach.

The Panel was asked to comment on the uncertainty approach of daily dietary exposure presented in the SHEDS Dietary Model FIFRA SAP 08-15-07.pdf document. In this approach, the SHEDS model is run multiple times, each time with a different subset of the CSFII person-day and/or commodities containing residues datasets, each subset constructed via bootstrap sampling. The results from these multiple runs are used to help answer the question: "How much better would the exposure estimates be if we had more data?" This is presented as one way of addressing Scenario Uncertainty (defined as descriptive errors, aggregation errors, judgment errors, and errors from incomplete analyses, see EFH Volume I, General Factors, Chapter 2, pp.2-5, 2-6.).

The approach proposed (described in the document SHEDS Dietary Model FIFRA SAP 08-15-07.pdf) creates limited CSFII person-day datasets each containing a fraction of the full dataset and a fraction of the commodities-containing-residues database. The full SHEDS (1-d Monte Carlo) model is run separately for each subset with a response distribution obtained for each. The proposal suggested doing this for a moderate number of times (n=200), with each replicate producing a response distribution. For each response distribution, an index of uncertainty would be computed and the distribution of this index examined. One index of uncertainty suggested, the ratio of the 95th percentile to the 5th percentile, is actually a measure of distributional spread. The median of this measure was proposed as one way of quantifying the potential impact on model results of inadequacies in the diet or residue data. An alternative measure suggested is the median of some important upper percentile; say the 99th percentile, across the set of simulated response distributions.

Exactly what was being proposed and how to interpret the results was not uniformly clear to all Panel members. In general, those Panel members commenting felt that using the median tail spread or median upper extreme percentile of the response distribution seemed reasonable indices to use. It was mentioned that EPA would always have the problem of choosing and defending the particular percentiles that make up the indices.

A couple of Panel members indicated concern about bootstrap sampling using the same data that is used in model development/calibration. Sampling from the calibration data does little to tell us what might be expected using an expanded dataset, especially where the expansion of the dataset might produce diets/concentrations that are more extreme than those in the calibration dataset. Using this logic, the sampling approach here only addresses part of the question pertaining to access to more data. In a sense the proposed approach attempts to answer the question of what might happen with "more data" by looking at what happens when we use "less data". One can reasonably expect an expanded database to include clusters of individuals that differ from those in the existing database in unpredictable ways. In that sense, the proposed method will not indicate the true value of enlarging the database. It will be useful for showing the uncertainty associated with re-sampling from a finite set of consumption and residue inputs.

One Panel member indicated that the figures used in the report are transposed with respect to what is typically seen in publications, in this case with the probability axis on the horizontal and amount on the vertical axis. From a communication standpoint it would be preferable to make that change.

b) Can the Panel recommend alternative approaches - and how they might be interpreted and used - for conducting uncertainty analyses of dietary exposure estimates?

Working with the indices and statistics (median and upper percentiles of the 95/5 percentile ratio distributions), members of the Panel had suggestions for other analyses approaches.

The uncertainty estimates as presented involve four variables: the percentage of CSFII food consumption data used (denoted p_1), the percentage commodities used (denoted p_2), the number of samples generated in one run of the SHEDS model and from which the response distribution was constructed (referred to as *b*), and the number of SHEDS runs for the uncertainty analysis (denoted as *r*). In one example provided to the Panel, these values are $p_1 = 1/20$, $p_2 = 1/4$, $b = 200$, and $r = 200$. In this example, *b* and *r* are fixed at 200 and a total of 40,000 (200x200) simulated person diets were created and used. One Panel member wondered what would be the best choices for *b* and *r* if they were allowed to vary but the total number of simulated person diets was held constant, (e.g., 40,000) for specific values of p_1 and p_2.

Measuring the impact of changes to p_1 and p_2: A relatively straightforward approach was proposed to analyze the impact of changes to p_1 and p_2 on the 99th percentile of the CI ratio. For a fixed food percentage, say, p_1, let (x_i, y_i) denote the point for the x_i percentage of commodity with the corresponding 99th percentile of the CI ratio y_i, $(i = 1,2)$. Suppose $x_2 > x_1$, then the slope of the line connecting the two points (x_1, y_1) and (x_2, y_2), $s = (y_2-y_1)/(x_2-x_1)$, might be used to quantify the reduction of uncertainty of the confidence limit ratio for the change in food percentage. For example, using the data presented in Table 9 (*An Update on the Development of the SHEDS-Dietary Model* July 25th, 2007, USEPA) for $p_1 = 1/8$ the slope of the line connecting

the two points (1/30, 4.47) and (1, 1.73) for the 99th percentile is -0.0283. Thus an increase of 1% more commodity results in a decrease in the confidence limit ratio of 0.0283. Similarly, for fixed residue percentage p_2 at 1/10, the slope for the two points (1/8, 2.52) and (1/4, 2.43) is -0.0007. Thus, a 1% increase in the percentage of the diary samples has less effect in reducing the uncertainty ratio at the 99th percentile than a corresponding 1% increase in food percentage.

Post hoc power analysis: One Panel member suggested that the sensitivity analysis proposed looked similar to a post-hoc power analysis for determining sample size (b, p_1 and p_2 relate to sample size) when designing a study. Power analyses are typically made based on certain model assumptions. In this case, it is not clear what those assumptions are and it is unlikely that any assumptions made could be justified using available information. In this case, one is assuming that a bootstrap sample of sample size larger than the original database is representative of what might happen if one were actually able to get more diet or concentration data. Since there is actually no new data here, it is quite likely that the analysis will overestimate the power and underestimate the needed sample size.

Recent developments with the bootstrap method related to the analysis of large genomic datasets might be of value here. For example, simulating 365 days and a full 24 hours per day produces a dataset of over 10,000 values per individual. It is likely that the planned extensions in SHEDS 4 will require higher levels of modularity and, with the need to perform more missing value imputation, will drive the development of new resampling designs.

Inter and Intra run variability: Even when using the whole food and residue datasets, the Monte Carlo nature of the SHEDS model will ensure that each run will produce a slightly different response distribution. As the number of runs (r) increases, the differences in the response distribution (as measured by the indices discussed previously) among runs may be small. One would need to first quantify the range or variability in response distributions (Intra-variability). Next, model runs would be performed using a fraction of the food and/or residue datasets ($p_1 < 1$ and/or $p_2 < 1$) as described above. Again, the range or variability of the resulting response distributions would be quantified (Inter-variability). One recommendation is to normalize the Inter-variability by the Intra-variability, holding total simulation size (40,000) or n and b fixed for each situation. Another recommendation is to scale the Inter-variability values to their theoretically equivalent Intra-variability value. That is, normalize Inter-variability based on amount of additional information going into the model, so that for example, if $p_1 = ¼$ s then the full dataset would be 4 times larger so Inter-variability might be multiplied by 4 for comparison purposes. It was noted that 200 individuals may be too small of a sample to produce reliable results. It is important to understand the metric one is looking at, realizing that for one metric one might have a sample size equivalent to 200 values while some other metric has 40,000 values.

Alternate Indices: As mentioned previously, the ratio index suggested in the proposed approach is actually a measure of the (extreme percentage) range of the response distribution. It would be preferable if an index could be used that measured not only the range but more of the full form of the distribution. One Panel member made a presentation suggesting a number of indices that measure distributional differences. Unfortunately, the presentation was based on a misunderstanding of what was being presented in the plots accompanying the proposed approach. The discussion below takes the basic ideas presented by that Panel member but extends them more appropriately to the data that are actually available in the uncertainty simulations.

Assume n simulations are run with the full diet and concentration datasets, each simulation being run with b individuals. Let P_s be the cumulative distribution function for run s=1…n. Let P_0 represent an estimate of the cumulative distribution constructed using the median for each percentile value from .1% to 99.9%. The basic idea is to use as an index of uncertainty for each simulation run s, a function of the area between P_s and P_0. One could then compute the standard deviation or extreme percentile range of these run indices to measure the uncertainty of changes in p_1 or p_2. Smaller values mean that the curves are quite close to each other indicating little "overall" uncertainty from the limiting factor, say, for example, diet. The indices proposed are simply goodness of fit test statistic originally proposed by Anderson and Darling (1952, 1954). Mathematically this area can be represented as an integral.

$$I_{1s} = \int_{.001}^{.999} \left(P_s(x) - P_o(x)\right)^2 \Psi[P(x)] dP_0(x)$$

The weighting function, $\Psi[.]$, allows one to differentially emphasize areas of the differences between the full data distribution, $P_0(x)$ and each simulation sample distribution, $P_s(x)$, and in particular can be used to focus the index on comparisons in the tails of the distribution. The Cramér-von Mises statistic does not do this weighting and uses $\Psi[P(x)]=1$. The original Anderson-Darling test statistic uses $\Psi[P(x)]=[P(x)\{1-P(x)\}]^{-1}$ placing more weight on the tails of the distribution. Many consider this the reason that the regular Anderson-Darling test is more powerful than the Cramér-von Mises test. Other simpler forms of this index can be found in Hall (1989) and Cook, Buja and Cabrera (1993). Any of these indices could be used in place of the two indices proposed in the EPA document. A good reference on goodness of fit tests would be Thode (2002).

Question 3-5:

Please suggest statistical or other methods that might be used to determine the extent to which region- and season- specific dietary consumption amounts and patterns might be important in developing dietary exposure estimates. Please consider in your response whether and how quantitative uncertainty methods could be used in addressing this issue.

The NHANES (National Health and Nutrition Examination Survey) database does not include information on season nor region because the Center for Disease Control (CDC) did not release that information in order to protect the respondents' privacy. Food consumption may vary with season, and from region to region, and the differences in dietary consumption scenarios may result in different exposure estimates. The NHANES data may be sufficient to capture the general statistics at the national level, but they may not be adequate for use in regional or seasonal scale assessments. Panel members are not aware of any existing statistical methods that might be directly used to determine the extent to which region- and season- specific dietary consumption amounts and patterns might be important in developing dietary exposure estimates.

One Panel member raised a question of whether commodity consumption differs regionally, pointing out that diaries might be different from region to region but commodities are likely to be from national or international sources. Exposure analyses that were presented at the previous SAP meetings indicated that seasonal and regional variations do not generally contribute

significantly to the high end dietary exposures when a large number of commodities are included in the dietary exposure estimation. However, without seasonal- and regional- specific analyses, this conclusion is not apparent in light of the intuition that seasonal and regional variations of consumption likely exist.

Statistical methods alone cannot resolve the uncertainties that arise due to regional and seasonal variations. The issue of the regional and seasonal differences can be addressed through analyzing the population and region- or season-stratified datasets, such as data from, for example, CSFII. For example, the CSFII dataset has the information on season and region, and CSFII was designed to represent the whole population. Using bootstrap methods, the percentiles of the exposure estimates can be computed for each region, each season, and region and season combination. The 50^{th}, 95^{th}, and 99^{th} percentiles can be compared to evaluate the impact of regional and/or seasonal differences in diaries, consumption, and exposure on the exposure estimates.

If different regions or different seasons would result in different exposure estimates, then the question is the utility of NHANES dietary consumption data given the lack of access to information on the season and region. Panel members suggested that it may be possible to generate region-specific or season-specific food lists. For example, Floridians consumed much more marine fishes than their near neighbors by looking at sales data for these commodities. A simple analysis of variance (ANOVA) can be applied to generate region-specific commodity lists. In NHANES data, there are approximate 6,000 food items consumed by respondents. For a given food item, the regional difference in consumption can be tested using the ANOVA. The ANOVA is performed for each of the 6,000 food items resulting in 6,000 statistical tests, which can be described by the 6000 individual p-values measuring the extent of the difference in consumption. Application of appropriate multiple test procedures could then be applied to obtain the region-specific food list. Furthermore, these food items can be used as surrogates for the region to compute total exposure. This method can also be used to generate season-specified food consumption lists.

NHANES protocols were initially designed for the interest of collecting information on population nutrition and heath status. Some of the data, but not all, have implications of interest to the SHEDS development. The data missing from the NHANES are critical to the Agency to further develop SHEDS. There are other issues, such as fasting over night or over a period of 8 hours before the appointment, associated with the NHANES protocols that would further limit the use in SHEDS development. The Panel strongly recommends that the Agency should be involved in the study design stage for the future NHANES to address those missing data/information issues. The Panel supports a better and stronger sharing of these data among the Agencies.

The Agency is encouraged to continue looking into collecting more updated dietary survey data, especially for building longitudinal patterns, with the idea that new data can be added to it on a regular basis. A Panel member recommended that the Agency conduct a longitudinal dietary and time-location microactivity study on its own. The Panel member wants to emphasize the value of collecting longitudinal data, so it can facilitate the development of the SHEDS model. The timing of conducting such a study is impeccable, considering the availability of technology that can be deployed to facilitate the data collection and the better understanding of both

chemical use and human exposure patterns. If possible, the further study that the Agency will conduct should be reviewed by a separate SAP prior to the initiation of data collection.

References:

Anderson, TW and DA Darling (1952), Asymptotic theory of certain "goodness of fit" criteria based on stochastic processes. Annals of Mathematical Statistics 23, 193-212.

Anderson, TW and DA Darling (1954). A test of goodness of fit. Journal of the American Statistical Association 49, 765-769.

Bennett DH, Furtaw EJ Jr. (2004) Fugacity-based indoor residential pesticide fate model. Environ Sci Technol. 38(7):2142-52. Erratum in: 38(18):4912.

Bennett DH, Margni MD, McKone TE, Jolliet O. (2002b) Intake fraction for multimedia pollutants: a tool for life cycle analysis and comparative risk assessment. Risk Anal. 22(5):905-18.

Bennett DH, McKone TE, Evans JS, Nazaroff WW, Margni MD, Jolliet O, Smith KR. (2002a) Defining intake fraction. Environ Sci Technol. 36(9):207A-211A.

Cook, D, A Buja and J Cabrera (1993) Projection Pursuit Indexes Based on Orthonormal Function Expansions, Journal of Computational and Graphical Statistics, Vol. 2, No. 3 (Sep., 1993), pp. 225-250.

Fenske RA, Lu C. Determination of handwash removal efficiency: incomplete removal of the pesticide chlorpyrifos from skin by standard handwash techniques. Am Ind Hyg Assoc J. 1994 55(5):425-32.

Givens et al. (2007) Estimating dietary consumption patterns among children: A comparison between cross-sectional and longitudinal study designs. *Environmental Research*, 103:325-330.

Gurunathan S, Robson M, Freeman N, Buckley B, Roy A, Meyer R, Bukowski J, Lioy PJ. Accumulation of chlorpyrifos on residential surfaces and toys accessible to children. Environ Health Perspect. 1998 106(1):9-16.

Hall, P (1989). On Projection Pursuit Regression, The Annals of Statistics, Vol. 17, No. 2 (Jun., 1989), pp. 573-588.

Johnson N.L., Kotz S., Kemp A.W. (2005): Univariate Discrete Distributions. 3rd edition, Wiley Series in Probability and Mathematical Statistics, 672 p., New York: J. Wiley Inc.

Johnson N.L., Kotz S., Balakrishnan N. (1994): Continuous Univariate Distributions. Vol.1, 2nd edition, Wiley Series in Probability and Mathematical Statistics, 784 p., New York: J. Wiley Inc.

Johnson N.L., Kotz S., Balakrishnan N. (1995): Continuous Univariate Distributions. Vol.2, 2nd edition, Wiley Series in Probability and Mathematical Statistics, 752 p., New York: J. Wiley Inc.

Lu et al. (2006). A novel system for collecting longitudinal self-reported dietary consumption information: the Internet Data Logger (iDL). *J. Exp. Sci. Environ. Epidemiol.* 16(5): 427-433.

Saltelli, A., Ratto, M., Tarantola, S., and Campolongo, F. 2005. Sensitivity analysis for chemical models. Chemical Reviews 105 (7):2811-2827.

Thode, HT (2002). Chapter 5 Goodness of Fit Tests in Testing for Normality, Marcel-Dekker, Inc., NY, NY.

USEPA/OPP/HED, Health Canada/PMRA, CA EPA/DPR/WS&HB, Review of Worker Exposure Assessment Methods, see http://www.epa.gov/scipoly/sap/#january