SAP Meeting Minutes No. 2002-02

FIFRA Scientific Advisory Panel Meeting
April 30 - May 1, 2002

Held at the Sheraton Crystal City Hotel, Arlington, Virginia

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

CUMULATIVE AND AGGREGATE RISK EVALUATION SYSTEM (CARES) ™ MODEL REVIEW

Olga Odiott, M.S. /S/
Designated Federal Official
FIFRA Scientific Advisory Panel
Date: June 13, 2002

Stephen M. Roberts, Ph.D. /S/
FIFRA SAP Session Chair
FIFRA Scientific Advisory Panel
Date: June 13, 2002

Fred Jenkins, M.S. /S/
Designated Federal Official
FIFRA Scientific Advisory Panel
Date: June 13, 2002

NOTICE

These minutes have been written as part of the activities of the Federal Insecticide, Fungicide, and Rodenticide Act (FIFRA) Scientific Advisory Panel (SAP). These minutes have not been reviewed for approval by the United States Environmental Protection Agency and, hence, the contents of these 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 was established under the provisions of FIFRA, as amended by the Food Quality Protection Act (FQPA) of 1996, to provide 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 EPA, Office of Pesticide Programs (OPP) and is structured to provide balanced expert assessment of pesticide and pesticide-related matters facing the Agency. Food Quality Protection Act 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 meeting minutes 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 Larry Dorsey, SAP Executive Secretary, via e-mail at dorsey.larry@epa.gov.

CONTENTS

Federal Insecticide, Fungicide and Rodenticide Act
Scientific Advisory Panel Meeting
April 30 – May 1, 2002

CUMULATIVE AND AGGREGATE RISK EVALUATION SYSTEM (CARES) ™ MODEL REVIEW

PARTICIPANTS

FIFRA SAP Session Chair

Stephen M. Roberts, Ph.D., Professor and Program Director, Center for Environmental and Human Toxicology, University of Florida, Gainesville, FL

Designated Federal Official

Olga Odiott, FIFRA Scientific Advisory Panel, Office of Science Coordination and Policy, Environmental Protection Agency, Washington, DC

FIFRA Scientific Advisory Panel Members

Mary Anna Thrall, D.V.M., Professor, Department of Pathology, College of Veterinary and Biomedical Sciences, Colorado State University, Fort Collins, CO

FQPA Science Review Board Members

John L. Adgate, Ph.D., Assistant Professor, University of Minnesota School of Public Health, Division of Environmental and Occupational Health, Minneapolis, MN

Patrick Durkin, Ph.D., Vice President, Syracuse Environmental Research Associates, Inc., Fayetteville, NY

Bernard Engel, Ph.D., Purdue University, West Lafayette, IN

Natalie Freeman, Ph.D., Robert Wood Johnson School of Medicine, Department of Environmental and Community Medicine, Piscataway, NJ

Dale Hattis, Ph.D., George Perkins Marsh Institute, Clark University, Worcester, MA

Steven Heeringa, Ph.D., Director, Statistical Design and Analysis, Institute for Social Research, University of Michigan, Ann Arbor, MI

Peter Macdonald, D. Phil, Professor of Mathematics and Statistics, McMaster University, Hamilton, Ontario, Canada

Kenneth Portier, Ph.D., Associate Professor, Department of Statistics, University of Florida, Gainesville, FL

Thomas L. Potter, Ph.D., Research Chemist, Agricultural Research Service, U.S. Department of Agriculture, Southeast Watershed Research Laboratory, Tifton, GA

Nu-May Ruby Reed, Ph.D., Staff Toxicologist, California Environmental Protection Agency, Department of Pesticide Regulation, Sacramento, California

Lauren Zeise, Ph.D., Chief, Reproductive and Cancer Hazard Assessment, Office of Environmental Health Hazard Assessment, California Environmental Protection Agency, Oakland, CA

PUBLIC COMMENTERS

Oral statements were made by:

Christine F. Chaisson, Ph.D., on behalf of The Lifeline™ Project and The LifeLine Group.

Robert Sielken, Jr., Ph.D, Sielken Associates, on behalf of CropLife America.

Jennifer Phillips, Ph.D., Aventis CropScience, on behalf of CropLife America

Written statements were made by:

Christine F. Chaisson, Ph.D. on behalf of The Lifeline™ Project and The LifeLine Group.

INTRODUCTION

The Federal Insecticide, Fungicide, and Rodenticide Act (FIFRA), Scientific Advisory Panel (SAP) has completed its review of Version 1.0 of the Cumulative and Aggregate Risk Evaluation System (CARES) ™ Model. This review included discussions on the operation and documentation of the software, the model design and the model results. Advance notice of the meeting was published in the *Federal Register* on April 3, 2002. The review was conducted in an open Panel meeting held in Arlington, Virginia on April 30-May 1, 2002. Stephen M. Roberts, Ph.D. chaired the meeting. Ms. Olga Odiott served as the Designated Federal Official.

The CARES Model was developed through a cooperative effort between CropLife America, Government, and Academia. All Panel members were presented copies of the CARES User Guide, CARES Technical Manual, CARES Code Manual, CARES Program CD #1 and CARES Document CD #2 with Prototype CSU. Included in the CARES User Guide was a Tutorial in which the user was directed through five (5) case studies in which the capabilities of the CARES software were demonstrated. CropLife America, through InfoScientific and Sielken Associates, provided Panel members technical support in installing and running the software.

It was noted that FIFRA SAP meetings have been previously held to review different models for performing cumulative and aggregate risk assessments.

CHARGE

Documentation and Operation

CARES Documentation in the form of a User Guide, Technical Manual, and Code Manual were provided to the SAP to assist in their review of the CARES software.

Question 1: The User Guide provides installation instructions and 5 case study tutorials to illustrate operational features of the CARES software. Is the User Guide complete and understandable? Were Panel members able to load the software on to their computers and complete some or all of the case study tutorials?

Question 2: The Technical Manual provides an overview of the CARES model and detailed descriptions of key model components. Is the Technical Manual complete and understandable, and are the descriptions of specific model components scientifically sound?

Question 3: The Code Manual provides annotated code for select risk assessment algorithms used in the model. Do the algorithms in the annotated code perform the functions defined in the CARES Technical Manual?

Model Design

Question 4: The key data sources used in CARES include the US Census /PUMS and the CSFII/FCID. Is the methodology used by the CARES Population Generator to create the CARES Reference Population based on appropriate consideration of demographic factors (vector of individual characteristics) and statistical representativeness?

Question 5: The CARES reference population constructs a 365-day dietary profile by matching similar individuals in the CSFII database using specific matching rules. Does the CARES approach provide a reasonable and realistic construct with respect to temporal variability in magnitude and frequency of food consumption?

Question 6: What can the Panel say about how these databases were used together in estimating dietary exposure? Are there other publicly available sources of relevant data? What research opportunities would support a refined calendar based, probabilistic exposure and risk analyses using CARES?

Question 7: The CARES residential model uses an Event Allocation Module to create a temporal profile of residential-related product use occurrences, and includes a product co-occurrence matrix that can be derived from available survey instruments. Does the CARES approach provide a reasonable and realistic construct with respect to temporal variability in magnitude and frequency of residential use and the likelihood of exposure co-occurrence events?

Model Results

Question 8: Case studies described in the Users' Manual and provided on the CARES - CD illustrate the complexity of conducting Monte Carlo simulations in a PC environment (multiple parameters for multiple chemicals, where dose and risk from multiple vectors of exposure are calculated over 365 days). Does the SAP believe that an iterative testing strategy with CARES could be developed (conduct of multiple simulations with progressive refinement once risk drivers have been identified) that will permit the user to obtain information on source contribution necessary for decision-making?

Question 9: What types of contribution/sensitivity analysis are recommended by the Panel to be most useful in making scientifically sound regulatory decisions for one or more pesticides and their associated agricultural, professional and consumer uses? What should be routinely reported as part of a CARES assessment with respect to inputs and outputs? Are there certain key graphs and tables that should be reported? What types of model evaluation steps does the Panel recommend to further refine and advance models such as CARES?

SUMMARY OF PANEL RECOMMENDATIONS

- Panel members were able to install and run the CARES model program and some or all parts of each tutorial/case study, and generally found the User Guide well written and understandable. Several deficiencies were noted, many of which would be expected in a beta-test version of a program and User Guide. For example, one of the most commonly noted deficiencies was that CARES is not robust against user mistakes. Once it is in an error state, CARES will usually crash. Panel members suggested an "undo" feature to minimize system crashes occurring from user errors. Several recommendations to increase the functionality of the program were provided.

- The Panel found the Technical Manual to be complete, easy to understand, and well illustrated. Inclusion of a summary of key limitations and/or assumptions for each of the modules, and a clear separation of the material contained in the technical documents and the white papers were among the recommendations for enhancing the Technical Manual.

- The Panel noted that the annotated code provided in the Code Manual is not properly commented. The Panel suggested the development of commenting guidelines and the addition of a formal "Code Inspection" step to the validation process.

- The Panel agreed that the use of the US Census/PUMS and the CSFII/FCID databases to create the CARES Reference Population appears to be appropriate and facilitates statistical representativeness in the basic sub-groups. However, the Reference Population constructed from the PUMS demographic sample provides disproportionate representation for selected geographic and demographic subclasses. It was suggested that use of additional databases such as NHAPS would provide additional demographic factors without disrupting statistical representativeness.

- The SAP noted the need to periodically update the databases to assure that the model maintains its representativeness. The 1990 Census data currently used in CARES Version 1.0 will need to be updated from the 2000 Census due to documented changes in demographics since the 1990 Census.

- The SAP noted that several improvements should be made to the CARES model approach of matching dietary records of individuals in CSFII to records for the initial "reference person". It is not known at present how well the autocorrelation caused by this matching replicates the true autocorrelation of food consumption in actual dietary behavior. The Panel suggested evaluation of how the current matching system replicates the degree of autocorrelation indicated by an empirical analysis.

- The Panel noted that the lack of more extensive food consumption data is a key limitation of the CARES model. The Panel further expressed that it was not aware of any additional publicly available databases that are sizable and can be directly applied to formulating a reliable and representative food consumption pattern over an extended period such as 365 days. It was recommended that the CARES developers pursue an initiative to develop more realistic eating patterns by turning to research and expertise in the behavioral sciences,

marketing, nutrition, and medicine.

- The Panel agreed that based on the data available, the CARES model provides a reasonable approach in addressing temporal variability in magnitude and frequency of residential use. However, the Panel could not confirm whether the estimates derived from this approach are realistic. The SAP further explained that to determine how scientifically defensible the outputs of the model are, the model must incorporate additional residential monitoring data.

- The SAP suggested that an iterative testing strategy with CARES should focus on fine-tuning the high-contributing factors, which may be different for each exposure scenario and may need to be determined with preliminary runs.

- The Panel agreed that more experience with the CARES model is needed in order to determine refinements that would contribute to sound regulatory decisions on pesticides. This should include a variety of chemical/population/exposure scenarios, and data for currently registered active ingredients.

MINUTES OF PANEL DELIBERATIONS

Question 1: The User Guide provides installation instructions and 5 case study tutorials to illustrate operational features of the CARES software. Is the User Guide complete and understandable? Were Panel members able to load the software on to their computers and complete some or all of the case study tutorials?

The charge question addresses, essentially from the perspective of a potential user, the adequacy of the program and User Guide to enable the new user to install and run the program, to understand the basic program analyses and outputs, and additionally, to provide an entrée into possibly more sophisticated analyses and outputs. Panel members were able to install the program, run some or all parts of each tutorial/case study, and generally found the User Guide well written and understandable (albeit incomplete). Members noted several deficiencies, many of which would be expected in a beta-test version of a program and user's guide. An overall recommendation was to increase the functionality of the program, in addition to debugging.

The tutorials/case studies were found to be well constructed and quite useful. It was noted though that it may be difficult for many users to get started with only the program, tutorials and the other written materials, and that workshops or other approaches would likely facilitate the introduction and use for new users. The discussion below outlines the experiences of Panel members in installing the program, working through the case studies and working with program output.

Installation.

In general, Panel members with computers meeting the minimum specifications, and in some cases with less RAM than specified, successfully installed the program. One Panel member was only able to install the program on one of two systems tried, failing to install the program on a Windows 98 system apparently due to problems with ATL.DLL. In this regard, the Panel member attempted on several occasions to obtain assistance and explanation from Infoscientific but received neither a solution enabling the loading of the program on the Windows 98 system, nor an explanation. (Notitia™ is a registered trademark of Infoscientific and some program developers are from that firm.) It is noted however, that while the User Guide clearly instructed the user to follow the instructions for installation and set up (User's Guide, p. 9), the instruction on the first program to be loaded – dcom.exe – may puzzle some new users. The Guide instructs the user to load the CARES Program CD and execute dcom.exe, but on Windows 2000 machines the installation fails ("DCOM98 can only be installed on Windows 95 and Windows 98"). The novice is left to discover elsewhere that the program is only needed for Windows 98 installations and left wondering about the adequacy of the overall installation in the absence of the initial file.

With regard to recommendations, installation should be seamless on all systems with the minimum specifications, and efforts should be made to ascertain those systems for which this is not the case and the program updated accordingly. Any variations in installation procedures to be followed by machines meeting the specifications should be clearly noted in the User Guide. A proof reading of the User Guide should be conducted in tandem with installation of the

program on the variety of machines meeting the minimum specifications, with consequent correction as needed.

Running the Tutorials/Case Studies.

Crashes and Errors. Crashes were common and frustrated the tutorial exercises. On some systems, rebooting was required after each crash, whereas on others the program closed and had to be restarted. In either event crashes result in loss of work and are a great inconvenience to the user. Members had a variety of experiences and made various observations and suggestions regarding crashes and errors. First, CARES is not foolproof or robust against clumsiness. Once it is in an error state, as occurs if you try to open the wrong type of file, CARES will usually crash. Also, it is sensitive to the version of Excel™ data used as importing data. Second, the development of an "undo" feature would be useful. When an object was accidentally deleted, replacing it did not work. In the subsequent run, run time error 457 occurred, crashing CARES. This was noted by a number of Panel members, one of whom reported it occurring twice. An undo capability would have enabled the user to recover easily from the initial error. Third, the development of a facility to trap errors is clearly needed. As an example, in attempting to graph data, one Panel member reported selecting an individual who has "no data." A window popped up that indicated "no data found." The only way to recover was to stop the CARES run and restart. Further, more guidance on error messages is needed. The location of the error log generated is not given and the error message is sometimes difficult to understand. Finally, it is noted that crashes resulted from following the User Guide instructions, normal machine function (screen saver), and user error. On at least one system, activation of the screen saver resulted in a crash. Unspecified run time errors occurred on some systems running the canvas in Tutorials 4 and 5, requiring the system to be rebooted. This may have been due to inadequate RAM on those systems, but this could not be determined by the error messages given. In subsequent version of CARES, error trapping and softer crash and landings would surely increase usability.

File Management. An improved means of file management is needed. The location of files created by the User is not immediately obvious, nor are the names of the files. One Panel member produced four "Indiana Males" subsets, some of them invalid, but was unable to delete any of them. Attempts at using the delete button or otherwise trying to remove these files from the hard drive were unsuccessful and the files would reappear in subsequent runs. Multiple members experienced difficulties with file management.

Non-functioning Features. On some systems, the desktop CARES icon worked some, but not all of the time. When it failed, a run time error message was given. The Run Specifier - Advanced button appears to do nothing. There was no response from Help in several occurrences, including the Chart help button. The Help feature is clearly limited and provides no explanation or guidance regarding use of common features.

Feedback to User. On a number of occasions it was unclear whether a CARES operation was running – neither the cursor nor the windows provided an indication. The only indication of processing is the obvious activity of the hard drive. Further, there are several operations that require extended periods for completion. In some instances, an indication of the time required for completion of the operation is given by the current number being processed out of the total to be processed. However, this feature is not provided for a number of operations, leaving the user

to guess the total time remaining, and as indicated above, in some cases, left wondering whether the program is running at all. For all non-instantaneous operations, some indication of the time remaining in the operation window would be helpful. The message timer could be calibrated to the machine rate.

Case Studies and the User's Guide

The User Guide overall is of good design, clearly written, and well illustrated. The tutorials/case studies are well constructed and quite useful. More explanation in the User Guide of the objective of the case studies and operations within them, and ways the user can check and utilize results would substantially increase their usefulness. In this regard it would be helpful to include a road map to describe the analyses to be completed in the case studies. Placing this at the start of the case study would be helpful in understanding what the sequence of steps laid out in the User Guide is trying to accomplish. Inclusion of brief explanations along the way of what each step is trying to achieve would also strengthen the case studies. The tutorials stopped short of demonstrating all the graphical and analytical possibilities. More suggestions for "what if" and "what else" looks at the results after running each case study would be helpful. A further tutorial that continues with the drill down analysis would be quite instructive.

Each tutorial took a considerable amount of time to run. It would be useful to provide the user the option of working through the tutorial on smaller data sets. The User Guide does not indicate the length of time anticipated for the completion of each tutorial, nor for most operations performed. This information would be useful to the user, both in planning to learn the program and in gaining assurance and understanding in working through the exercises.

With regard to the specific case studies run, some errors were noted as well as formatting problems leading to confusion. For Case Studies 4 and 5 the accumulator routine did not appear to be aggregating exposures. For Case Study 4, a cursory comparison between exposure data output showed no entry for "Scenario 1 Route 1" and "Scenario 2 Route 3" for accumulated routes, while present on other data sheets. The reason for the apparent missing data is not clear to the user and should be explained. It was noted that the user is never quite sure when a spreadsheet can be closed. Some instruction and reassuring notes in the User Guide are needed.

The User Guide did not match the version of the program and was incomplete in that several screens encountered in runs were not included in the guide. Associated experiences of the members using the guide varied, perhaps due to varying levels of familiarity with such programs. Some found the mismatch problematic and an impediment to learning the program; others found it to be a minor distraction. Still, there was general agreement that the User Guide and program version released should match, and be fairly complete. Mismatches occurred between the screens and Guide illustrations for both entry and operation screens and results. The mismatches in the case study results displayed on screen and in the User Guide, including the CARES ID for the individual record, can be particularly problematic. The new user will be uncertain about whether entries erroneously entered or operations misapplied could explain the failure to reproduce the findings given in the User Guide. Examples of incompleteness in the Guide included: the Viewing Results windows shown in the User Guide never appeared on screen; and the Save Dataset windows following exposure calculations are not provided in the User Guide case studies documentation.

Further simple improvements to enhance the program would be to put the CARES ID on any individual plots and to limit the default number of decimal places displayed in numeric cells in output tables. In the results output data grids, the column headings cannot be fully seen without greatly extending the width of the columns, and this should be remedied. In plotting results to be compared, for example those on individuals, standardization of the axes would provide for comparisons to be made more easily. Various members were confused by the appearance that total exposure was less in some circumstances than maximum daily exposure. This was due to lack of comparability across columns in number formatting. For ease of reading, comparable formatting should be used. Units and headers for the database columns should be included in the data spread sheets for easy reference. Also, in Case Study 4, regardless of the chemical selected, chemical #1 was highlighted. Program data sheets and the User Guide should be carefully proofread to correct problems such as these.

The limitations of the basic databases included in CARES, such as the CSFII, and the analyses performed should be discussed to provide the new user with some understanding of the uncertainties associated with the CARES output. The Guide provides no guidance or discussion in this regard. Increasing the discussion in the Technical Manual, and cross-referencing it to the User Guide, is an obvious solution to making this information more available to the user.

Question 2: The Technical Manual provides an overview of the CARES model and detailed descriptions of key model components. Is the Technical Manual complete and understandable, and are the descriptions of specific model components scientifically sound?

The Technical Manual appears to be complete, is easy to understand, and is illustrated well. Some issues may be raised that may be useful to the developers of CARES in either clarifying the documentation and/or addressing some issues in areas where the Agency's approach to cumulative risk may allow for latitude or improvement. The following are specific recommendations for improving the Technical Manual.

A summary of key limitations or assumptions of each of the modules would be appropriate to include in their respective documentation section.

Much of what was contained in the technical documents was found in the white papers. A clearer separation would make both stronger. The technical documents should be reserved for discussing the specifics of the modules, providing enough detail on key assumptions and computational approaches that the reader has a clear understanding of what was done. The white papers should serve the purpose of reviewing and identifying the background to the effort.

In several places, the manual indicates specific features are not included or will be included in future versions of CARES. Some indication of the timeline for including such features would be useful.

Water Module Chapter

In the Water Module, the references used are not listed at the end of the chapter. Additional references should be included to allow readers to pursue additional scientific and other detail. For example, the Water Module briefly discusses PRZM/EXAMS but provides no reference to the model or work being done for the FQPA with this model. A second example from the Water Module where a reference would be appropriate is when water treatment effects are briefly discussed. The manual indicates CARES does not provide treatment effect capabilities. A reference or references that provide more details on treatment and its importance would be helpful to readers. A third example from the Water Module indicates monitoring data can be used in CARES but there is no indication of how or other requirements. Additional discussion or use of references would be appropriate.

The Water Module discusses geographic scale issues in detail but provides little or no guidance on the temporal aspects. What are the characteristics for a year of data that should be used?

Separation of content in the white paper and the technical chapter is needed. A simple statement like "All computations related to drinking water exposures must be made off-line and imported." is suggested for the water technical document. While the water documentation is problematic in this sense, the white paper provides a good discussion of important issues such as scale. However, there are some important omissions on the topic, including discussions at prior SAP meetings that are documented in Panel reports and represent current agency policy (e.g. use of index reservoirs, crop-area factors). A dialog on the feasibility of using water consumption patterns, bottled, office, home, etc; in the same way that food was handled in the dietary module would also be helpful.

Chapter 1

Section 1.4: CARES takes the same approach as the most recent version of the OPP Cumulative Risk methodology in generating route-specific MOS's and then adding these to get cumulative risk. In the last SAP review, the suggestion was made to consider an alternative approach – using route conversion to estimate "internal" dose. This approach could be considered in the CARES program and would lend itself to the use of PBPK models at some point. This is at least mentioned in some parts of the technical manual.

Section 1.7: At the top of page 14, there is a statement to the effect that the hazard index is somehow inferior to the MOS for chemical mixtures and that the hazard index can/should be used only if "... the magnitude of the uncertainty factors are the same". This does not seem to be correct. As the document clearly states elsewhere, the two approaches are essentially equivalent/reciprocal and neither is any better than the other.

Section 1.11: In separating the dietary and residential modules, it is not clear where or if the consumption of homegrown vegetation is considered. The treatment of home grown vegetation is clearly part of the residential model. There is also a pick your own fruits/vegetables scenario that handles dermal and inhalation exposure from commercial sources. It is unclear, however, where that consumption of homegrown vegetation is considered. This may be a very important route.

Appendix B

In Appendix B, text on page B-9, the authors first make a series of statements supporting the adequacy of samples of 5,000 individuals with a series of ordinal statements about the probabilities that various sample quantiles would be greater or less than larger or smaller population quantiles. Unfortunately, however, these calculations did not take into account the effect of the sample weights. Large differences from person to person in the population weights have the effect of reducing the effective number of independent observations below the raw number of people in the selected sample. Correct calculations of the probabilities listed in Table 1 and summarized on page B-9 would need to reflect this. However, even correct calculations of these ordinal characteristics are not of primary interest in judging the adequacy of the sample size. For purposes of making regulatory determinations, what matters is the breadth of the confidence intervals that can be expected for the absolute exposure levels calculated using the CARES system for upper percentiles of the population distributions. The last sentence on page B-9 seems to be making a statement about this when it says,

> "...a sample size of 5,000 essentially guarantees that the 99.9th sample will not appreciably underestimate the 99.9th population quantile and will not appreciably overestimate the 99.9th population quantile more than 5% of the time."

Unfortunately, in order to make a statement about the degree of uncertainty in estimating absolute exposure levels of the 99.9th or other population quantiles, the authors would have to utilize some information or an assumption about the overall amount of variability expected in the exposure distribution. A sample of 5,000 people from a lognormal population distribution with a geometric standard deviation of 10 would likely be estimated less precisely (have wider multiplicative 90% or 95% confidence limits above and below the central estimate) than if the geometric standard deviation were 3.

Appendix C

In the residential white paper there was an excess use of the term 'conservative' without clearly explaining how the term was being used. In one case the document describes an unclothed child as a conservative clothing scenario. This is a correct use of the term conservative, since the unclothed child is the upper limit on skin exposure. In contrast, the document claims that the Jazzercise exposure scenario is 'conservative' without explaining why. This is particularly a concern since there is lack of good data available to determine if this scenario is conservative or not when applied to children.

There are several ill-advised passages in the Residential White Paper on pages C-30 and C-31. The first of these says,

> "Consideration should be given to using the discrete values of a data set in place of a continuous distribution when it is not possible to determine how well the data conform to some theoretical distribution. When a continuous distribution is used that can mathematically continue to infinity, such as a lognormal distribution, the distribution must be truncated."

On the contrary, theoretical distributions are needed to make some estimates of upper percentiles of exposure distributions precisely when sample data are too limited to provide direct observations for relatively rare individuals in the tails. To assume that the highest directly observed person in a sample of 30 happens to correspond to the largest possible value of a continuous variable is clearly wrong. A continuous distribution should only be truncated if there is good reason to believe that values above some limit are physically impossible (e.g. because of a solubility limit, mass balance limitations, or some other known property of the system). To create an artificial truncation point introduces a bias toward underestimation of the upper percentiles of the exposure distribution relative to the likely actual exposures of people in the upper tail. This is true even for central tendency estimates of upper percentile exposures, and it would foreclose the EPA from effectively evaluating how confident it could be (that is the uncertainty) that upper percentile exposures were below some specific criterion level—e.g. the uncertainty in any conclusion that the 99th percentile or 99.9th percentile individual would have less than a "margin of exposure" of 100. Farther down on page C-30 it is reported that

> "During initial model case study simulations significant instability was found in the upper 10% of the distribution. Such instability was determined to result from not truncating the upper end of some of the parametric input distributions. Refinements such as placing bounds on the input distributions provided stability in the model output to beyond the 99th percentile."

The authors unfortunately do not report their criterion for judging "stability" —presumably it is a particular standard error of the estimated exposures on replicate runs with sample sizes achievable within the CARES system. In any event, if greater stability/narrower confidence intervals are desired, the solution is to increase the sample size. It is not acceptable to artificially truncate input distributions for this purpose. To do so does bring about more rapid convergence, but with incorrect estimates of the desired percentiles of the exposure distribution.

The final paragraph on page C-30 reads in part,

> "The EPA has generated significant controversy with its use of the 99.9th percentile of the dietary output distribution for acute risk assessment. It is unlikely that most of the data sets utilized in non-dietary exposure assessments will be of sufficient size to measure the 99.9th percentile of each input variable with any certainty. Therefore the 99.9th percentile will incorporate compounded uncertainty that surrounds the upper-ends of each input variable distribution. Decisions regarding the selection of an appropriate percentile for risk assessment must be made in the context of an understanding of the nature of non-dietary exposures and the populations being assessed.

One cannot quarrel with the last sentence in this paragraph. However, the other statements here invert the desirable relationship between risk assessment analyses and risk management determinations of the criteria on which EPA judgments of adequate public health protection are to be based. The limitations of particular existing residential exposure data sets should not be a prime consideration for EPA in selecting which values of x, y, and z fulfill its mandates to protect public health and achieve an appropriate balance between the benefits and risks of particular uses of pesticides. If current data are inadequate to provide the degree of confidence that is desired that exposures of upper percentile individuals are within acceptable limits, then

additional data should be generated.

Appendix D

Dermal 101, 102: As an input, the algorithm describes an absorption fraction but this does not appear to be used in the calculation of the output. Thus, the output appears to give dermal deposition rather than absorbed dose. If this is what is intended, the reference to fraction absorbed should just be removed. If internal dose is intended to be calculated, some better explanation should be given for fraction absorbed. This is not a real pharmacokinetic parameter. If one puts a dose on the surface of the skin, dermal absorption can be calculated from zero-order or first-order kinetics or in some more complicated way – mixed kinetics or a PBPK model that treats the skin as a multi-component compartment. In any case, one can express the amount absorbed after a specific time as a fraction of what had been on the surface of the skin. This is very situation specific, however, and is not a "property" of the chemical. The source of the "absorption fraction" should be clarified in this appendix.

Dermal 109: This is standard calculation for absorbed dose based on zero order absorption using U.S. EPA 1992 dermal guidelines.

Dermal 109, 110, Ingestion 105: The "reference duration" seems to be a way of converting to mg/kg day rather than mg/kg bw. This should be clarified. This may not be meaningful if the "reference duration" scaled the dose over several days.

Ingestion 102, 103, 104, 105: We are assuming that children will eat grass, dirt, paint chips, and swallow swimming pool water. This is fine. Where, however, is the scenario for the consumption of home grown vegetation? This is more plausible and likely to be a much more substantial source of contamination.

The technical manual (C-43) mentions pica along with soil consumption. Incidental soil consumption and pica are two very different things. Pica generally means the eating of any foreign substance (Halstead 1968), which includes the deliberate consumption of non-food items such as soil, paint chips, and plaster. Pica behavior is highly variable. A value of 10 g/day has been recommended by U.S. EPA/NCEA (1996 and 2000), a factor of 25 above the upper range of incidental soil consumption, 400 mg/day. It is not clear what value is used for pica by the program.

Ingestion 107, 108, 109: Guidance/discussion on when to use these three different approaches for hand-to-mouth transfer was not found. The only one that makes obvious sense is 109, from the EPA Standard Operating Procedures (SOP) for Residential Exposure Assessments.

Question 3: The Code Manual provides annotated code for select risk assessment algorithms used in the model. Do the algorithms in the annotated code perform the functions defined in the CARES Technical Manual?

The code is not properly commented. Although programming standards follow the guidelines in *Practical Standards for Visual Basic,* much more could be done to add comments that would make the code understandable. Five generally accepted rules for commenting are:

- Comment all of your code with the fewest number of words.
- Everything must be commented on.
- Use self-commenting code as much as possible – a lot of this was done.
- Use external references as much as possible – none of this was done, even though there are two standard sort functions listed in the code.
- Use comment blocks at the beginning of each function or subroutine that list, at a minimum, the programmer's name, the date, and the purpose of the code. Typically, a description of arguments and a statement of the return values from the function are also included.

There is very little block commenting in the code. With 49 public functions, 56 private functions, 21 public subs and 19 private subs it is very difficult to keep track of everything. There are some in-line comments but many of these are useless in helping the reviewer understand the code. (a good example is "yep" on page 89.) There are a number of "TO DO" comments in the code with nothing to indicate what there is to do. The Panel strongly suggests that the developers develop commenting guidelines and spend a few days adding in the comments.

There are places where the code does not seem to match the algorithm descriptions. The descriptions of the Residential Algorithms in Appendix D are easily compared to the associated code. In two places, the Residential Algorithm description included in the formula a Reference Duration, but the code does not make reference to this value. (Examples: dermal 105, dermal 106, ingestion 105 use a reference duration, user input, in the equation for hand exposure, but the equation, page 8 in the code section, does not.). In a number of places the Residential Algorithm incorporates some conversion constants that, when computed, typically divide the final calculation by some factor of 10. For ingestion 102 and 103 the residue numbers should be divided by 10, yet the code does not. This could be a significant factor in that the residue value then used in the exposure computation is 10 times larger than it should be. Finally, in Inhalation 103 there seems to be a factor of 10^{-3} left out of the code for computing the exposure from air concentration.

We were asked only to address the Exposure algorithms in the Code pages. Only the algorithms directly linked to Appendix D were understandable. This is only about half the code in the Code sections. There are a number of other functions and subroutines that do some of the other operations listed in Figure 3.1, page 50. Without documentation, the Panel was unable to determine exactly what the functions and subroutines were doing.

It does look as though the algorithms directly linked to Appendix D would be relatively easy for a user to modify a formula or add a new module.

On pages 47 to 49 in the code section, Private Function Get BNF and Private Function Get Areas have hard-coded specific "biometric" parameters. Not only is there no documentation of what and where these come from, but a user may want to play with these parameters when looking at the sensitivity of the system to assumptions. Again, these might need to be made "user input" or at least collected in a database.

The use of variant data types slows down Visual Basic programs greatly, yet there are many

Variant data types used in function calls and even as counters in FOR loops. In addition, many functions are declared with variant data types. It is much faster to use a strongly typed approach as in C++. In particular, counters should be typed as integers or longs rather than variants.

The programmers have elected to use ADO (ActiveX Data Objects) rather than DAO (Data Access Objects) which makes sense given that Microsoft does not have a strong commitment to DAO and is pushing ADO. Nonetheless, at this time, DAO is faster and easier to use for programs that are based on Access files, as CARES appears to be (Smith & Sussman 1999).

A formal "Code Inspection" step should be added to the validation process. This would involve a careful inspection for typos and verification that the computations all correspond exactly to the formulas in the CARES Technical Manual. It is well known that code inspection generally leads to cleaner, more transparent code even when there are no coding errors to pick up.

Question 4: The key data sources used in CARES include the US Census /PUMS and the CSFII/FCID. Is the methodology used by the CARES Population Generator to create the CARES Reference Population based on appropriate consideration of demographic factors (vector of individual characteristics) and statistical representativeness?

The ideal empirical reference population database (in CARES terminology, the vector of individual characteristics) for modeling cumulative exposure for individuals would at a minimum include:

1) Large representative sample of individuals that provides a high resolution demographic and geographic characterization of the U.S. population and its major subpopulations;
2) Data on these sample individuals' household or living arrangements that include drinking water sources, lawn, garden and household maintenance activities, activity pattern data for multiple days at multiple points in the year and reporting of all major household applications of products; and
3) Food and water consumption data for these individuals for consecutive day periods and at multiple points throughout the seasons of the year.

No existing database provides such a comprehensive set of data for a population-representative sample of individuals. To build a facsimile of this ideal sample-based reference population, the alternative is to use statistical models to assign the behaviors and characteristics in (2) and (3) to a geographically and demographically representative sample (1). In statistical terms this is exactly equivalent to imputation of unobserved or missing data. One model-based approach to creating the composite records is to statistically match (Rodgers, 1982) empirical data for similar (not identical individuals) from multiple census and survey sources. In statistical matching the "model" to impute food consumption or other attributes to individuals of certain age, gender, ethnicity and geographic residence is implicit in the distance criterion used to establish the pool of match candidates and any randomization applied in selecting a donor from the pool of identified matches. This is the approach that CARES has adopted for imputing food consumption profiles to reference population individuals.

The alternative approach is to impute characteristics to the base demographic population using conditional distributions estimated through regression models (linear, logistic, poisson, multinomial logit, etc.) fitted to survey data (e.g. CSFII/FCID) or possibly even posited from expert judgment or observational studies. This second approach could work well for quantities such as weight and height but is very awkward if the task is to impute a profile of food items and quantities consumed.

The 1990 PUMS 5% Census microdata sample is a representative sample of individuals and individual living arrangements in the United States on April 1990. Properly weighted, the CARES Reference Population is representative of the geographic and demographic characteristics of the 1990 U.S. population. Obvious demographic changes have occurred since 1990, such as the relative increase in Hispanic and Asian ethnicity populations that are not reflected in a proportionately weighted sample from the 1990 PUMS. A minor improvement would be to post-stratify the CARES subsample of 1990 PUMS individuals to current demographic distributions from the 2000 Census. This post-stratification would entail simple adjustments to current sample weights within geographic and demographic subgroups.

The CARES Reference Population that is constructed from the PUMS demographic sample provides disproportionate representation for selected geographic and demographic subclasses. The disproportionate subsampling of subpopulations (e.g. minorities, Census Divisions, infants) requires the application of a sample selection weighting factor whenever simulated results for the Reference Population sample members are pooled to create statistical summaries (e.g. 95^{th} percentile MOE for women age 20-39). The development of these weights for the chosen CARES reference population is described in the User Guide and Technical Documentation of the Model. Care must be taken in weighted estimation of population statistics. Means, correlations, regression coefficients have formulas based on weighted sums, sums of squares and sums of crossproducts. Weighted estimates of quantiles require construction of a weighted CDF. For small simulation samples where the true sampling variance of the RP subsample is nontrivial, the weighted estimation will also influence the variance of the sample estimates of statistics that are produced in the course of a risk assessment analysis. Current software systems such as SAS, STATA, SUDAAN, and WesVar PC include special programs that incorporate correct methods for estimated variances of weighted estimates of sample statistics.

In the CARES model (and all others involving dietary, residential, and water exposures), the characterization of the baseline population involves a modeling step. The process of developing the CARES Reference population involves a statistical match of a Census PUMS case to a CSFII individual respondent. The system of self-weighting via a calculated similarity index is one of the most innovative and creative contributions of the CARES system. However the current implementation of the system implicitly gives greater weight to foods that are consumed in larger amounts in determining the influence of various demographic factors on the similarity index. This is reasonable in cases where the distribution of pesticide residue levels among different foods is unknown. However, for cases where it is known that foods making up a modest portion of the total diet convey a major fraction of the total population dose of a particular pesticide, it would be desirable to have an option allowing the user to weight the relative importance of various foods in determining the similarity index by relative contributions to total population aggregate dietary doses.

Another improvement to the sampling would be a classification of the sampled individuals by urban vs. rural place of residence (or, more generally, by population density of the residential location at some appropriate level of aggregation—e.g. county). There is some doubt that the sample of 100,000 "hard-wired" people is adequate for all purposes. The choice of 100,000 as the nominal sample size is arbitrary after the choice of critical subpopulations and precision for estimated extreme quantiles in exposure distributions. Pushing the PUMS sample to sizes much greater (i.e. 2 to 3 times) than the CSFII sample sizes adds information only in the ability of analysts to specify geographically restricted populations. Nevertheless, it would be desirable to provide an option for the user to create an increased sample size by selecting additional individuals from PUMS if needed for a particular analysis.

As noted above, this is in fact a data imputation procedure. As the "vector of individual characteristics" is generated from this initial reference population match, the amount of new "statistical information" that is brought to our understanding of real world complexity is constrained by the size of the CSFII and NHAP samples and the strength of population associations between the Census characteristics in the base sample to the food and water consumption and residential characteristics of the individual (Little and Rubin, 1987).

How can we bring more "statistical information" to the characterization of the reference population? The short answer is that without additional new survey data collections it will be difficult to greatly increase the information contained in an empirical reference population data set such as that used by CARES. Information on demographic and geographic distributions can be increased by expanding the size of the population base sample (see 1 above). This could be achieved by increasing the size of the sample from the PUMS. It would also be possible to expand the base sample for the reference population by adding vectors of demographic and geographic variables from representative national sample programs such as the Current Population Survey (CPS) or the Health Interview Survey (HIS). Representative samples of special subpopulations that are very rich in covariate information could be obtained from special surveys such as the Early Childhood Longitudinal Survey (ECLS, children age 0-10) or the Health and Retirement Survey (HRS, adults age 50+). Moving to these other surveys as a source of representative sample data for individuals is only valuable if they provide additional empirical covariates (not found in PUMS) that can contribute to the modeling effort or to provide a better current representation of household and population characteristics that cannot be obtained through simple weighting adjustments to the 1990 PUMS or ultimately from the 2000 Census PUMS.

The use of the US Census /PUMS and the CSFII/FCID databases appears to be appropriate and facilitates statistical representativeness in the basic sub-groups. In the text it is stated that the vector of individual characteristics (VIC) used by the population generator is generated in a mutually consistent fashion for all modules. One database not used in CARES but which may be of value in conjunction with the Census/PUMS and CSFII/FCIF is the NHAPS. This can be easily linked to the Census data and would provide additional demographic factors of urban/suburban categorization without disrupting statistical representativeness.

There is a need to periodically update databases to assure that the model maintains its representativeness. The 1990 Census data currently implemented in CARES V1.0 will need to be updated from the 2000 Census because of the documented changes in demographics since the

1990 Census.

Question 5: The CARES reference population constructs a 365-day dietary profile by matching similar individuals in the CSFII database using specific matching rules. Does the CARES approach provide a reasonable and realistic construct with respect to temporal variability in magnitude and frequency of food consumption?

The consensus of the Panel is that considerable additional improvements are both desirable and feasible. The "secondary" matching of dietary records for other individuals in CSFII to records for the initial "reference person" is necessary due to the lack of longitudinal data. The weight of each matching parameter is dependent on an index of relative similarity estimated from a variety of data types (for which the Panel recommends some improvements—see below). An attempt is made to preserve the seasonality of food consumption habits through restricting matching neighborhood in time to 7-30 days from the calendar day that the CSFII information was collected. Preference is also given for matching the same day of week while weekdays are set to be more similar than weekends. Within this construct, each individual entry is treated as independent from day to day. However, in terms of consumption data, other factors that may have significant impact should also be considered. One such factor is the linkage between what is eaten on successive days of consumption due to purchasing patterns of perishable foods (e.g., large batches of fruits from roadside stands, e.g., watermelons) and holiday left-overs (e.g., apple pies from Thanksgiving). It is possible that factors such as these should transcend the preference for matching weekdays and weekends. What is important is to determine which factor could have greater impact on ultimate risk drivers, the current CARES default or linking days based on the types of foods, or both. There should also be some consideration in the future of procedures to enhance the ability to incorporate the second record for CSFII secondarily matched individuals by forcing out some previously selected records/individuals based on some statistical criteria.

The CARES model incorporates earlier Panel recommendations to depart from replication of single individual CSFII one or two-day food consumption reports. It does this at some cost in rendering the system relatively challenging to explain to other professionals, let alone people with less expertise in statistical modeling.

Despite the CARES authors' efforts to match records for reference persons with the most plausible available records for days beyond the two days available for the reference person, it is just not known at present how well the autocorrelation caused by this matching replicates the true autocorrelation of food consumption in actual dietary behavior. This subject, however, appears amenable to better empirical analysis through current efforts of Novigen researchers to examine the degree of autocorrelation in the consumption rates for specific foods seen in the two available days of data for individuals in the CSFII data. Once this autocorrelation information is available, the CARES authors should be encouraged to first evaluate whether their current matching system replicates the degree of autocorrelation indicated by the empirical analysis. If, as may be expected, it does not, the system should be adapted to change the frequency with which the food consumption records from the original CSFII reference person are selected in order to reproduce the empirical autocorrelation for relevant foods as well as possible. In any event it would be desirable to incorporate a feature into the modeling system that allows the user to vary the degree of autocorrelation in order to evaluate the effect of uncertainty in dietary autocorrelation on risk drivers relevant to decision-making—such as the Margin of Exposure

(MOE) for upper percentiles of the overall exposure distribution.

Exposures under specialty diets (fruit, vegetarian or ethnic specialty diets) of the types that we have discussed in previous SAP sessions on models are really testable hypotheses. CARES and other exposure models would accommodate a test panel of individuals with specialty diets that could be processed routinely in population exposure analyses. The food consumption profile for test panel individuals could be described based on model diets defined by nutritionists or food chemical exposure experts. Since little is known about the prevalence of these special diets in the general population, the results for the test panel would need to be segregated from the general population summary. The results would then need to be weighted with the aid of surveys designed to measure the frequency of each defined type of dietary behavior in the general population. Pending such a general population survey, the value of the test panel would lie in the ability to compare modeled exposures for these special cases to the general distribution of exposures estimated for the population at large.

One Panelist suggested that autocorrelation in the residue data might be just as important as autocorrelation in the amounts of different types of foods eaten. The assumption that the residue distribution is uniform throughout all regions and seasons should be rigorously examined by a systematic data analysis, especially for foods that may potentially show temporal and geographic variations and could significantly affect the dietary exposure for chemicals and chemical classes of interest for risk analyses.

Another Panelist expressed a desire to inspect some of the dietary profiles generated by CARES in the course of some future review. Presumably; this can be done easily, and it will make it possible for the CARES user to cross-reference the CSFII ID with the consumption details.

Finally, empirical studies of metabolic rates indicate that lean body weight is an important determinant of metabolic rates, and hence food consumption and total daily air inhalation rates (Ravussin et al. 1986; Ford, L. E. 1984). Lean body weights are readily estimatable from data in CSFII on height and weight. The CARES authors would do well to give further consideration to including estimated lean body weight in the criteria used for their secondary matching (to fill in the 363 days not already filled by the two days of data typically available for the primary-matched reference person).

Question 6: What can the Panel say about how these databases were used together in estimating dietary exposure? Are there other publicly available sources of relevant data? What research opportunities would support a refined calendar based, probabilistic exposure and risk analyses using CARES?

The comments on questions 4 and 5 above provide many suggestions that are also pertinent to this question. The lack of more extensive food consumption data is a key limitation in overcoming potential problems and concerns with this and other approaches in creating food consumption patterns.

There is a clear need for more longitudinal data on food consumption. The Panel is not aware of any additional publicly available databases that are sizable and can directly be applied to

formulating a reliable and representative population food consumption pattern over a prolonged period of time (e.g., 365 days). There are also concerns that large scale reliable data may be very difficult to obtain. One Panel member noted that there may be studies of the effectiveness of food frequency and diet record surveys, but there is one that comes on the data disk accompanying *Fundamentals of Biostatistics* (Rosner, B., 2000, Fifth Edition, Duxbury). This exercise compares a food frequency study with a diet record study on the same people over the same time period. It is interesting that only alcohol consumption shows any real correlation between the two studies and even then the heavy drinkers could not recall accurately how much they had consumed. This exercise gives students a good introduction to reality and leaves one somewhat cynical about the possibility of getting good longitudinal data on large numbers of subjects.

Thus, the main focus under this discussion is on how to best use the individual's short-term (1 - 2 days) survey data to realistically construct long-term dietary intake patterns. CARES uses a sophisticated approach to match individuals in the population to food consumed. This represents an interesting approach in creating longitudinal food consumption patterns. However, many questions remain and additional validation of this approach is needed.

In terms of constructing a reference population, the procedure used for producing daily profile of an individual's consumption for each day throughout the year does not address several key behaviors that one would expect to be important in reproducing eating patterns, such as habituation of diet, development of taste preferences, product loyalty, and recurrent menus. It would be instructive to examine the daily menus throughout the years associated with any given CARES run and would be useful if the facility to do so were provided. Because eating behavior is not specifically addressed, one would expect the matching procedure to fall far short of being able to capture the recurrent eating for an average individual, let alone those at the upper percentiles. Matching on the basis of raw agricultural commodities rather than food types (e.g., tomato, rather than tomato paste or fresh tomato; wheat and not pasta or whole wheat bread) results in heavily weighting the large consumption foods, such as starches and meats, and lessens the chance that the constructed profiles for an individual will be realistic, especially for raw agricultural commodities that are not a considerable fraction of the diet. It is therefore doubtful that average and upper percentile distributions will be realistic.

An effort to reconstruct more realistic eating patterns for chronic and sub-chronic analyses is clearly needed. It is noted that the development of protocols to sort food as consumed into raw agricultural commodities was a resource intensive process using empirical information on recipes and judgment to develop generic recipes for the analysis. An initiative to develop more realistic eating patterns could turn to research and expertise in the behavioral sciences, marketing, nutrition and medicine to explore different techniques for patterning an individual's diet. Appropriate experts from these fields may also be useful in developing procedures to address how food consumption patterns for an individual change with age and other characteristics. A solution for these generic problems would require a commitment within the Agency and considerable time to address, and is not particular to the CARES system. Nonetheless, it would be useful for CARES to develop a framework that would allow the user to make certain assumptions about recurring menus for an individual in constructing the food profile for the year. For example, a day's diary of eating could be assigned a certain probability of recurrence on each day of the year. This could be further refined to enable the user to assign

certain probabilities to specific meals, or within those meals commodities consumed with differing degrees of regularity. Correlations between each day in the two-day consumption in the CSFII may help in assigning probabilities between adjacent days, but provide little insight beyond. Information may be available from food product consumption studies. One suggestion is to look into the possibility of utilizing food purchase databases collected by supermarkets for creating the link for food consumption patterns.

Question 7: The CARES residential model uses an Event Allocation Module to create a temporal profile of residential-related product use occurrences, and includes a product co-occurrence matrix that can be derived from available survey instruments. Does the CARES approach provide a reasonable and realistic construct with respect to temporal variability in magnitude and frequency of residential use and the likelihood of exposure co-occurrence events?

The documentation for the Event Allocation Model appears to be clear and complete. Section C provides a clear description of how the module works. Section 4.2, discusses data sources used and the quality of the data – specifically the National Human Activity Pattern Survey (NHAPS), the National Home and Garden Pesticide Use Survey (NHGPUS), and the National Human Exposure Assessment Survey (NHEXAS). The descriptions were very helpful to Panel members unfamiliar with these surveys. The descriptions of the Exposure Factors Handbook (EFH) and the Child-Specific Exposure Factors Handbook (C-SEFH) are reasonably complete. [The link to the C-SEFH is not current. The current link is: http://cfpub.epa.gov/ncea/cfm/ efcsefh2.cfm].

As with other aspects of aggregate and cumulative assessments the ideal data to develop a believable probabilistic modeling construct for residential exposures is a longitudinal survey that records the co-occurrence in use of pesticide products over time. The Residential Exposure Joint Venture (REJV) will eventually provide an estimate of the variability in product use over time as well as estimates of the co-occurrence probabilities. This will be a step forward, but won't be perfect and will raise new issues.

Until these data are available the method proposed in section 4.3 is sound as it uses the existing data sources in a reasonable manner with two caveats. To briefly summarize, CARES generates a random number of scenario events over a year, then randomly distributes these events over the year, then randomly generates the number of products used per event, then randomly picks the product. The process has appropriate constraints, i.e., it appears to use a logic tree that accommodates dependencies introduced by demographics, season, use intervals, etc. The distributions should not be artificially truncated by existing empirical data with relatively small sample sizes. Instead the bounds (if not also the shape) of the distributions should be based on the physical limits on parameters where information is sparse and/or based on small data sets. The example presented in section 4.3 is otherwise sound as it takes a scientifically defensible approach to the important issues, accommodates and adapts the existing data, and appears to match on appropriate variables, such as gender, age, SES, location, etc. Some clustering of uses are more likely than others. For example joint occurrences of lawn, garden and tree care are more likely than joint occurrences of lawn and pet care.

The questions of "reasonable and realistic" are quite different. The Panel does not see anything

"unreasonable" in the methods used to estimate temporal variability in magnitude and frequency of exposures, although caveats were noted above. On the other hand, we have no idea if the estimates are "realistic". To determine how scientifically defensible the outputs of the model are the model needs to incorporate additional residential monitoring studies and these are clearly beyond the scope of the CARES project. All that can be said of the Event Allocator Module is that it uses the available data in a reasonable way.

The Panel notes that the Residential white paper has considerable overlap with the example in section 4.3, and has a number of shortcomings, i.e., it mixes policy and science issues unnecessarily in places, and cites non-existent documents (e.g., page C-27 refers to "section VIII" and page C-44 refers to appendix H).

In the residential exposure model, point estimates for body weight for different age groups are used. On the otherhand, in the dietary model, based as it is on the CSFII, individual weights are used. This can lead to a situation whereby the joint dietary/residential exposure assessment may be using different weights for the same individual for the two parts of the assessment. This seems an untidy way to handle the joint assessment and could lead to problems for some scenarios.

Question 8: Case studies described in the User Manual and provided on the CARES - CD illustrate the complexity of conducting Monte Carlo simulations in a PC environment (multiple parameters for multiple chemicals, where dose and risk from multiple vectors of exposure are calculated over 365 days). Does the SAP believe that an iterative testing strategy with CARES could be developed (conduct of multiple simulations with progressive refinement once risk drivers have been identified) that will permit the user to obtain information on source contribution necessary for decision-making?

Sensitivity analysis options should be designed to meet the most common needs of the exposure analysis and the risk assessors who will be the primary users of the CARES model (this is discussed further in the Question 9 response). An iterative testing strategy should focus on fine-tuning the high-contributing factors. These factors may be different for each exposure scenario and may need to be determined with some preliminary runs.

The user can start by exploring the model and trying out different scenarios. With so many factors of possible interest, a more systematic approach, such as that taken in experimental design (e.g. fractional factorials) should be used. The order in which different scenarios (runs) are considered could be determined by examinations of computational efficiency; that is clustering scenarios for which the results of some of the model steps are the same and hence can be reused without being recomputed. This approach is addressed in Welch *et al.* (1992).

A scripting language would be very helpful, allowing the user to automate the process of running and re-running scenarios as part of a sensitivity analysis.

Because of the complexity and large numbers of factors, the CARES model seems to have characteristics in common with current climate change models as well as with process control models. It was suggested that approaches to sensitivity analysis used with these models might find good use with the CARES model and eventually identify which factors are the major model

driving factors. Refer, for example, to Kharin & Zwiers (2000) or to some of the dimension reduction methodologies found in the *Journal of Chemometrics.*

One concern is that some potentially significant parameters may be hard-wired into CARES (refer to Question 3 comments about pages 47 to 49) and not subject to sensitivity analysis. Thus, the future versions of CARES should consider allowing the user more capability to directly specify some of these parameters and set weights for some exposure vectors such as those that would impact the definition of the reference population (e.g., VICs).

Efforts should be directed at maintaining the highest quality in data inputs, exposure estimation, and assessment of uncertainty in the results of the CARES run. By making it so easy for the user to proceed without questioning all of the data used, greater responsibility for maintaining the integrity of the database falls on the software developer.

Question 9: What types of contribution/sensitivity analysis are recommended by the Panel to be most useful in making scientifically-sound regulatory decisions for one or more pesticides and their associated agricultural, professional and consumer uses? What should be routinely reported as part of a CARES assessment with respect to inputs and outputs? Are there certain key graphs and tables that should be reported? What types of model evaluation steps does the Panel recommend to further refine and advance models such as CARES?

There was consensus among Panel members that more experience running CARES is needed. This should include a variety of chemical/population/exposure scenarios and, to the best extent possible, use data for currently registered active ingredients. The results will help identify the most sensitive parameters and pathways, which are likely to contribute the largest exposures. Panel members also emphasized that the effects of model assumptions on the estimates of upper percentiles of exposure and reproducibility of exposure estimates under replicated simulations need to be explored.

Several suggestions were made which may help delineate exposure distribution tails in dietary exposure scenarios. Specific recommendations were:

- Determine whether the tight restrictions on dates in the matching used to build food consumption profiles are really needed. Multiple replications of the model varying only the restriction on dates (e.g. 7 days, 30 days, quarter) that define the pool of CSFII records that can serve as donors for reference population individuals would provide insight on how important tight restrictions on the donor time window are to the exposure analysis.
- Evaluate how diet re-insertions can be used to track recurrent exposure. Some individuals may eat large quantities of fruits or vegetables that are "in season" and available at low prices. This could contribute to "upper-limit" exposures, which are not captured using the current dietary matching procedures.
- Expand the number of simulation trials by independently repeating the model assessment multiple times. For example, when the selected population is small (e.g. Indiana males, n~1100) the number of simulation trials could be increased by

independently repeating the model assessment multiple times (5-20 times). Each replication would produce exposure estimates for each individual in the Reference Population and a distribution of exposure for the population. Averaging over the distributional statistics for each replication's exposure distribution could produce final measures of population distributions. Variability in exposure within each replication and between replications could be computed using the same formulas used in multiple imputation analysis.

To facilitate sensitivity and contribution analyses, modifications to the user interface were recommended. Performing these analyses in the "point and click" environment in the currently available interface will be time consuming and tedious. An alternative is to provide the opportunity to create an input/settings file in conjunction with the creation of a log file describing the parameters of a run. If provisions are made so that this file can be directly modified (e.g. a parameter changed directly in the file) the simulation can easily be rerun with all other parameters fixed. It was noted that this is the essence of what is normally done in a sensitivity analysis and that use of scripting language may prove helpful.

With regard to reporting, CARES developers were commended for providing access to reports (input/output tables) that display or summarize inputs, generated data, and outputs of model runs. It was recommended that the system maintain available graphical displays for individual exposures, route- and chemical-specific contributions. Some suggested input-output enhancements included summary tables of input parameters and key assumptions with each scenario. Currently values are embedded in tables, which are not easily accessible from the data analysis module. Making them readily available with reports will contribute to transparency required when the model is used for regulatory assessments. To improve readability, the number of "significant figures" in model output tables should be reduced. The number currently provided is unreasonably high and may create confusion among users regarding the "certainty" of results.

Finally, model developers were cautioned that the usability of the system requires balancing the complexity of internal display and analysis capabilities with export of analyzable data tables to other software systems (SAS, STATA, etc.). To aid in their use, export features in a variety of standard file formats in addition to the currently available comma-delimited ASCI files may be helpful.

REFERENCES

Ford, L. E. (1984) Some consequences of body size. *Am J Physiol*. 247(4 Pt 2):H495-507].

Kharin, V.V. & F.W. Zwiers (2000). Changes in the extremes in an ensemble of transient climate simulation with a coupled atmosphere-ocean GCM. *Journal of Climate* 13:3760-3788.

Little and Rubin, Statistical Analysis with Missing Data, John Wiley & Sons, 1987).

Ravussin E, Lillioja S, Anderson TE, Christin L, Bogardus C. (1986) "Determinants of 24-hour energy expenditure in man. Methods and results using a respiratory chamber." *J Clin Invest*. 1986 Dec;78(6):1568-78.

Smith, R & D. Sussman. (1999). Beginning Access 2000 VBA. Wrox Press Ltd., Birmingham, U.K. 864 pp.

Welch, W.J., R.J. Buck, J. Sacks, H.P. Wynn, T.J. Mitchell & M.D. Morris (1992). Screening, predicting and computer experiments. *Technometrics* 34:15-25.