SAP Minutes No. 2005-04

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

PRELIMINARY N-METHYL CARBAMATE CUMULATIVE RISK ASSESSMENT

AUGUST 23 - 26, 2005
FIFRA Scientific Advisory Panel Meeting, held at the Holiday Inn - Rosslyn at Key Bridge, 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 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 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 Myrta R. Christian, SAP Designated Federal Official, via e-mail at christian.myrta@epa.gov.

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

TABLE OF CONTENTS

PARTICIPANTS..... 5
INTRODUCTION..... 7
PUBLIC COMMENTERS..... 8
SUMMARY OF PANEL DISCUSSION AND RECOMMENDATIONS 8
PANEL DELIBERATIONS AND RESPONSE TO CHARGE..... 11
REFERENCES.....60

SAP Minutes No. 2005-04

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

PRELIMINARY N-METHYL CARBAMATE CUMULATIVE RISK ASSESSMENT

AUGUST 23 - 26, 2005
FIFRA Scientific Advisory Panel Meeting, held at the Holiday Inn - Rosslyn at Key Bridge, Arlington, Virginia

Stephen M. Roberts, Ph.D.
FIFRA SAP, Session Chair
FIFRA Scientific Advisory Panel
Date: October 13, 2005

Steven G. Heeringa, Ph.D.
FIFRA SAP, Session Chair
FIFRA Scientific Advisory Panel
Date: October 13, 2005

Myrta R. Christian, M.S.
Designated Federal Official
FIFRA Scientific Advisory Panel
Date: October 13, 2005

Federal Insecticide, Fungicide, and Rodenticide Act
Scientific Advisory Panel Meeting
August 23 - 26, 2005

PRELIMINARY N-METHYL CARBAMATE CUMULATIVE RISK ASSESSMENT

PARTICIPANTS

FIFRA SAP Chairs

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

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

Designated Federal Official

Myrta R. Christian, 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 & Director, Center for Environmental Health Sciences, College of Veterinary Medicine, Mississippi State University, Mississippi State, MS

H. Christopher Frey, Ph.D., Professor, Civil Engineering, North Carolina State University
Raleigh, NC

Kenneth M. Portier, Ph.D., Associate Professor, Statistics, Institute of Food and Agricultural Sciences, University of Florida, Gainesville, FL

FQPA Science Review Board Members

W. Stephen Brimijoin, Ph.D., Chair, Department of Molecular Pharmacology, Mayo Clinic and Medical School, Rochester, MN

John R. Bucher, Ph.D., D.A.B.T., Deputy Director, Environmental Toxicology Program, National Institute of Environmental Health Sciences, National Institutes of Health, Research Triangle Park, NC

Deborah Cory-Slechta, Ph.D., Director, Environmental and Occupational Health Sciences Institute, Robert Wood Johnson Medical School, University of Medicine and Dentistry of New Jersey and Rutgers State University, Piscataway, NJ

Bernard Engel, Ph.D., Professor, Agricultural & Biological Engineering, Purdue University, West Lafayette, IN

Dale Hattis, Ph.D., Research Professor, Center for Technology, Environment & Development (CENTED), George Perkins Marsh Institute, Clark University, Worcester, MA

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

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

Robert W. Malone, Ph.D., Agricultural Engineer, USDA-ARS, National Soil Tilth Laboratory, Ames, IA

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

P. Barry Ryan, Ph.D., Professor, Environmental & Occupational Health, Rollins School of Public Health, Emory University, Atlanta, GA

Tammo S. Steenhuis, Ph.D., Professor of Watershed Management, Cornell University Ithaca, NY

INTRODUCTION

The Federal Insecticide, Fungicide, and Rodenticide Act (FIFRA), Scientific Advisory Panel (SAP) has completed its review of the preliminary N-methyl carbamate (NMC) cumulative risk assessment. Advance notice of the meeting was published in the *Federal Register* on June 29, 2005. The review was conducted in an open Panel meeting held in Arlington, Virginia, on August 23 - 26, 2005. Dr. Stephen M. Roberts and Dr. Steven G. Heeringa chaired the meeting. Myrta R. Christian served as the Designated Federal Official.

The FIFRA SAP met to consider and review the preliminary N-methyl carbamate cumulative risk assessment. The Food Quality Protection Act of 1996 amended both FIFRA and FFDCA. One of the major changes is the requirement that EPA considers risk posed by pesticides acting by a common mechanism of toxicity. For such groups of pesticides, EPA's Office of Pesticide Programs (OPP) has treated cumulative risk, under FQPA, as the risk of a common toxic effect associated with concurrent exposure by all relevant pathways and routes. The N-methyl carbamate pesticides were assigned priority for tolerance reassessment early during the process of FQPA implementation. OPP established the N-methyl carbamate pesticides as a common mechanism group in February 2004 based on their shared ability to inhibit AChE by carbamylation. Those pesticides included in the cumulative risk assessment were announced in the February FR Notice. OPP has proceeded with the development of the cumulative risk assessment in a step by step process including review of a case study for the N-methyl carbamate risk assessment in February 2005 by the FIFRA SAP. Based on the comments from the SAP, the Agency made appropriate revisions. The Agency released its preliminary cumulative risk assessment for the N-methyl carbamate pesticides in late July 2005. The hazard assessment for these chemicals involved empirical dose-response modeling of the available red blood cell and brain cholinesterase inhibition and recovery data. The exposure assessment utilized probabilistic approaches in all pathways considered: food, drinking water, and residential/non-occupational for various population subgroups and regions. These aspects were incorporated into a preliminary cumulative risk assessment document, which the SAP reviewed in August 2005.

The agenda for this SAP meeting involved an introduction, background, and detailed presentations of the several issues related to the preliminary N-methyl carbamate cumulative risk assessment. Issues related to the hazard assessment were provided by Dr. Anna Lowit (Health Effects Division [HED], Office of Pesticide Programs [OPP], EPA), Dr. R. Woodrow Setzer (Office of Research and Development [ORD], National Center for Computational Toxicology, EPA), and Dr. Stephanie Padilla (ORD, National Health and Environmental Effects Research Laboratory, EPA). Issues related to water exposure assessment were presented by Mr. Nelson Thurman and Dr. Dirk Young (Environmental Fate and Effects Division, OPP, EPA). Dietary Assessment presentation was provided by Mr. David Hrdy (HED, OPP, EPA). Residential Assessment issues were presented by Mr. Jeff Evans, Dr. Steve Nako, and Mr. Philip Villanueva (HED,

OPP, EPA). Model Results Comparison and Cumulative Risk Assessment presentations were provided by Mr. Alan Dixon and Mr. David Hrdy (HED, OPP, EPA) , respectively. Finally, the Risk Characterization presentation was provided by Dr. Anna Lowit, Mr. Nelson Thurman, and Mr. David Miller (OPP, EPA)

Dr. Clifford Gabriel (Director, Office of Science Coordination and Policy, EPA), Mr. Jim Jones (Director, Office of Pesticides Programs, EPA), Dr. Tina Levine (Director, Health Effects Division, Office of Pesticide Programs, EPA) and Dr. Steven Bradbury (Director, Environmental Fate and Effects Division, Office of Pesticide Programs) offered opening remarks at the meeting.

In preparing these meeting minutes, the Panel carefully considered all information provided and presented by the Agency presenters, as well as information presented by public commenters. This document addresses the information provided and presented at the meeting, especially the response to the Agency's charge.

PUBLIC COMMENTERS

Oral statements were presented as follows:

On behalf of Bayer CropScience:
Iain Kelly, Ph.D., Gary Mihlan, Ph.D., and Abraham Tobia, Ph.D.

On behalf of Carbamate Working Group:
Harvey Clewell, Director, Center for Human Health Assessment, CIIT Centers for Health Research
Jane D. McCarty, DABT, Chair, Toxicology Sub-team of CWG, Technical Leader for Toxicology, FMC Corporation

On behalf of DuPont Crop Protection:
Ralph L. Warren, Ph.D.

On behalf of the American Bird Conservancy:
Michael Fry, Ph.D.

SUMMARY OF PANEL DISCUSSION AND RECOMMENDATIONS

The Panel addressed a total of eleven questions regarding the preliminary cumulative risk assessment for the N-methyl carbamates. In its responses, the Panel repeatedly commended the Agency for its progress in developing a cumulative risk assessment based on good empirical data, sound technology and proper statistical methodology. Overall there was strong support for the Agency's approach and for the document that was under review. Although there were many suggestions for improving

the clarity and transparency of the document, the Panel raised no issues that represent major stumbling blocks in the path towards a final risk assessment. Nonetheless, a number of points were felt to deserve serious consideration or reconsideration. These points are elaborated in later sections of this report but the most significant of them are highlighted here under the headings: Hazard, Water, Food, Residential, and Integration.

HAZARD

EPA's choice of oxamyl as the index chemical provoked extensive discussion. The Panel recognized that the diversity of chemical structure, toxicity, and metabolic half-lives among the N-methyl carbamates makes it nearly impossible to identify one single compound that might be considered truly representative of the group. Many of the Panel members agreed with the selection of oxamyl as the index chemical. However, several reasons for choosing carbaryl over oxamyl were articulated. These included carbaryl's wider use, including residential applications, the availability of metabolic studies, and the progress toward a PBPK model for this compound that will inform future risk assessments. Because the Panel did not reach a true consensus on the issue, the EPA is advised to consider carefully the discussion summarized later in this report, and thoroughly reconsider its choice of index chemical.

WATER

The Panel was in favor of the Agency's plan to account for variable rates of pesticide degradation at different soil depths and with setback distances from field to well. As for comparisons of the three models under consideration as predictors of pesticide residues in drinking water, the Panel favored evaluating models in light of the mass balances of water and pesticide and to consider the hydrology carefully, especially the effect of wells used for irrigation. The Panel also urged the Agency to consider scenarios in which a pesticide application is quickly followed by rainfall that leads to preferential flow to shallow groundwater. Suggestions were offered on several other points, including the potential use of informed individuals to identify small local areas of high use and great leaching potential that might lead to elevated risks of exposure.

FOOD

Panel recommendations included the following. The Agency is urged to make a more detailed analysis of food exposure and identify specific food-pesticide combinations that are major contributors to risk. A sensitivity analysis should be conducted to determine how the choice of assumed values for "non-detects" affects the estimated exposure. The Agency should evaluate the Carbamate Market Basket Residue Monitoring Study and its implications for cumulative risk assessment (particularly with respect to single item vs. composite samples). The Agency also should investigate the effect of seasonal residues and consumption patterns on the cumulative assessment, especially with regard to individuals whose diet is heavily weighted towards certain food

sources.

RESIDENTIAL

Several comments were offered on limitations in the REJV database, on ways to supplement or improve this database, and on ways to use a larger fraction of the data in it. Panel statisticians were united in opposing the creation and use of uniform distributions and to data truncation except in extraordinary cases, as when physical factors constrain the range of possible values. Another issue that received much discussion was the uncertainty associated with assessment of exposure from hand-to-mouth behavior in small children. The Panel agreed that the macro-activity approach in the current document will overestimate exposure, but there was no consensus on the proper solution to this problem. At least one Panel member argued for deleting this component from the assessment. Reasons given for such a step were that overall exposure from food and water is more important, that the dermal route already accounts for residential exposure in part, and that the hand-to-mouth data are more likely to propagate uncertainty than reduce it. This argument should not be dismissed "out of hand". In any case the Agency can consider the Panel's constructive suggestions for mitigating the problems involved in properly assessing children's exposure risks from mouthing behavior (see detailed response to **R2**).

INTEGRATION

With regard to this topic the Panel again gave the Agency high marks for the great improvement in its latest cumulative risk assessment document. Two integrative aspects inspired extensive thought and discussion at the present meeting. The first of these was the question as to whether BMD_{10} values should be based exclusively on estimated peak levels of AChE inhibition. The Panel considered an alternative strategy to accommodate the possibility that duration of action also is important, particularly in regard to developmental toxicity. An approach recommended for further consideration would multiply the standard relative potency factor (RPF) for each carbamate by the associated half-time for recovery of brain AChE with that agent. The effect would be to increase the RPF values for compounds with relatively slow metabolic clearance.

While considering the recovery half-life for inhibition of brain AChE in rodents, the Panel also recognized a problem that may arise in extrapolating to humans. When standard inter-species scaling factors are applied to some compounds, the resulting half-lives may violate one of the basic assumptions of the cumulative risk assessment now envisaged by the Agency. That is to say, cholinesterase inhibition by the N-methyl carbamates in humans may reverse slowly enough to cumulate from one day to the next.

Inhibition half-lives also were considered in a final Panel discussion focusing on the timing of water consumption events. The present practice of lumping such events into a single occasion appears to be conservative. When water consumption is distributed

across the day, and half-lives for reversal of inhibition are taken into account, peak inhibition of brain AChE is predicted to be substantially smaller than currently estimated. Since exposure to water-borne pesticide is the major contributor to risk in certain geographical settings, the regulatory impact of a decision to distribute or lump consumption could be significant.

PANEL DELIBERATIONS AND RESPONSE TO CHARGE

The specific issues addressed by the Panel are keyed to the Agency's background documents, references, and the Agency's charge questions.

Questions

HAZARD

EPA's hazard and dose-response chapter (I.B) and associated appendices (II.B.1-6) of the Preliminary Cumulative Risk Assessment describe the application of the Relative Potency Factor (RPF) method to the N-methyl carbamate pesticides. These documents a) outline the steps in developing the dose-response relationships for each pesticide and its capacity to inhibit AChE in rats; b) describe the data used in the assessment; c) summarize the empirical dose-response modeling which provides the basis for the relative potency factors (RPFs), points of departure (PoDs), and estimates of AChE inhibition half life; and d) provide the rationale for selecting oxamyl as the index chemical.

HAZARD QUESTION #1

Empirical Dose-Response and Time Course Modeling

At the February, 2005 meeting of the FIFRA SAP, EPA proposed an empirical model for use in the cumulative risk assessment of the N-methyl carbamates. This model contains a dose-response and a time to recovery component. Based on the comments from the Panel and following experience with its application EPA made some modifications to this proposed model. EPA has applied this revised empirical model to the available RBC and brain cholinesterase data for the N-methyl carbamates. BMD and BMDL estimates provided in the preliminary assessment were derived from cholinesterase data from multiple studies and in some cases, using different cholinesterase measurement techniques.

H1a. Please comment on the mathematical/statistical approach to modeling cholinesterase data used to estimate benchmark dose values and time to half-life recovery in the preliminary cumulative risk assessment. Please address biological and mathematical/statistical considerations in your response.

Response

The Panel was in consensus that EPA has used best available statistical methodology to fit the proposed empirical model for cholinesterase concentration and time course data. It was also agreed that the presented analysis demonstrated implementation of the comments suggested in the February 2005 SAP.

There was significant discussion about whether the model-estimated BDM_{10} (or $BMDL_{10}$) values are the appropriate values to be used in the determination of Relative Potency Factors (RPFs). A proposal that the RPFs be modified to take into account the apparently substantial differences in estimated half-lives for reversal of brain AChE inhibition among the carbamates was discussed several times in the Panel's three-day deliberations. At this initial stage of discussion Panel members generally agreed that the peak inhibition level is likely to be the most relevant measure of internal dose for producing gross acute toxic effects. It was understood that the current model aggregates all of the exposures projected for an individual in the course of a day by summing the RPF converted concentrations. There was concern that aggregating chemicals with very different recovery half-lives may not properly capture true exposure. Final Panel recommendations on this topic are included in the response to question I1.

A number of other observations relating to the mathematical/statistical aspects of the exposure model were discussed and are summarized below.

- There is uncertainty associated with the model building process and decisions made regarding model parameterization. The model used is non-linear and the available data are often inadequate to estimate all model parameters or the available data do not support the more complex model forms. In a number of situations the model is simplified by either simplifying the parameterization or by specifying some parameters as known constants. While these changes are documented, the impacts of these changes on the final estimate of the BMD_{10} or $BMDL_{10}$ are not discussed. It was recognized that this may be part of the sensitivity analysis performed on the final estimating models.
- The estimating model incorporates random effect terms to account for study-to-study and animal-to-animal differences in background cholinesterase levels. Histograms of the estimated random effects (the Best Linear Unbiased Predictors) would help demonstrate the appropriateness of distributional assumptions for the random effects as well as help identify influential ("outlier") individuals or studies.
- The initial fit of model parameters was via a graphical technique followed by a few maximum likelihood iterations. Given the expected correlation among model parameters, minor changes in initial estimates can result in significantly different final parameter estimates. Assurances need to be provided that the estimates presented are the maximum likelihood values or are very close to them.

- At least one Panel member suggested that in addition to helping implement the constraints on model parameters, the transformation of the parameters to log scale is likely to reduce the correlation among parameters and make for a model that is easier to fit.
- EPA chose BMD_{10}, which is the central tendency estimate of the benchmark dose at which there is a 10 percent response level, as a reference point for developing Relative Potency Factors (RPFs). The lower limit is not clearly defined but is presumably the 2.5^{th} percentile of the confidence interval. This lower limit is used as the point of departure (PoD) for extrapolating risk. Most studies are reported as able to detect a benchmark response (BMR) of 10% that is significantly different from zero. There was a question about what other measures of effect can be examined by the model (e.g., a BMD_{20} or peak concentration) and how such a choice would change the RPFs.
- The measurements of cholinesterase inhibition may tend to underestimate the actual inhibition, because recovery can occur during the dilution and prolonged incubation of various methods. Reducing dilution, shortening the incubation time, and lowering the assay temperature are known to limit the decarbamylation of inhibited enzyme. These modifications are not yet universal, although EPA scientists have published them in the open literature (Nostrandt et al., 1993). At the previous meeting, EPA reported that their inhibition measurements with the modified Ellman method and with the radiometric method generally agreed with registrant data on the N-methyl carbamates. The Panel encourages EPA to publish these results to reduce the remaining uncertainty on this issue.
- The Panel discussed whether the available dose-response data actually capture the peak cholinesterase inhibition. While researchers target the peak, it is typically not known if the first measurement is taken at the peak or shortly after. The question is how much error is introduced into estimates of BMD_{10} and recovery rates when a dose-response model is fitted to data obtained after the moment of peak inhibition.
- The main focus of the modeling was on the oral exposure route. To the extent that data are lacking for the other routes of exposure, one Panel member wondered how one might characterize at least semi-quantitatively the dose-response and time-response relationships. Another Panel member suggested that some sort of Bayesian analysis, using prior distributions developed via information gleaned from experts, might allow extension of results to other groups such as the very young or the elderly for which appropriate data are not available. This issue is partly revisited in the Panel's response to question I1.

H1b. Please comment on the adequacy, clarity, and transparency of the documentation provided for the empirical dose-response and time course modeling.

Response

In general the Panel found the documentation of the model a little too terse and overly dependent on a reader's previous experience with the model and associated topics. The recommendations listed below summarize the Panel discussion and should help clarify model components and improve transparency.

- The report only *presents* the model without justifying or supporting its choice.
- Key assumptions of the model should be listed along with arguments that support them. In particular, error distribution assumptions should be documented.
- Dr. Setzer's slide presentation was better organized and more useful in many ways than the written report. The written presentation of the full model would benefit from a reversed order, beginning with an overview of the full inhibition model (Eq. 5) and then proceeding to particular components in more detail. This was the order of the oral presentation, which the Panel found easier to grasp.
- More graphics would help to demonstrate the forms possible for the model components. For example, to illustrate the non-linearity of the dose-response model for inhibition as a function of γ, the Panel recommends Figure 1 (see below), similar to one of the graphs shown by EPA.
- Inhibition as a function of time as given by Equation 2 could be usefully illustrated as in Figure 2. This figure also demonstrates that inhibition reaches a peak at time T*, representing a competition between an exponential increase in inhibition after dosing and an exponential decrease in inhibition associated with recovery.
- A detailed explanation and justification for the full inhibition model should be included in the report.
- The simplified model (Eq. 4) could be illustrated with the graph given in Figure 3. Use of the $(t-\delta)$ term presumes that a unit response occurs at time $t = \delta$. This model seems to presume that cholinesterase inhibition occurs instantaneously, followed by rapid recovery.
- In comparing the two models it can be shown that if T_A (half-life for the process of inhibition) is very short compared to T_R (half life for the process of recovery), the full model and the simplified model give essentially equivalent results. This can be illustrated with Figure 4.
- The document needs to indicate clearly which parameters have been estimated and which have been held constant. It should also show whether or not the values applied for constants in cases where a parameter cannot be estimated are at least somewhat similar to the values obtained when the parameter can be estimated.
- Documentation of R scripts is in the help files but absent from the scripts themselves. One Panel member suggested that this documentation should also be included as comments in the scripts themselves.
- It is suggested that, when examples are given in the document, the background and justification for the example should also be provided.
- Although the Panel recognizes that, empirically, there may be a level below which cholinesterase activity cannot drop, regardless of pesticide dose, some comment is needed to justify incorporating this feature into the model (p 33).

- The second bullet on page 33 could be written in a less confusing manner to explain how the model can account for the possibility that data were not collected at the time of peak inhibition.
- Clarification should be given as to whether the distributions described on pages 36 and later are intended to represent uncertainty or inter-individual variability. For example, when there is more than one study, for each ID there is assumed to be a specific, D_R (or $\ln(D_R)$), and that these values are assumed to follow a normal distribution. The text indicates that this distribution may be different for the different sexes. Is this true and what are the implications for the model?
- The modeling steps are summarized on pages 37-38, but quantitative examples of how these steps are implemented would really help the reader.
- More discussion is needed on the potential dependence of recovery half-life on dose. The implications of this dependence for the final use of the model should be discussed.
- During the public comment period, a representative of the American Bird Conservancy raised the concern that half-lives associated with dermal exposure could be much longer than those for the oral route (Table I.B.6). EPA should consider this comment and possibly revise the dermal exposure assessment appropriately.
- There is some confusion as to when female data are used in the model and when not. Specifically, what is done when there is a significant gender-specific sensitivity, and how does this decision affect the overall model uncertainty?
- An argument is made that BMD_{10} values for brain AChE are more suitable health endpoints than RBC-based BMD_{10}'s, (Figure I.B.3 – the label for this figure should identify the oral route of exposure). For many of the lower potency chemicals, including thiodicarb and carbaryl, it does appear that brain AChE data provide endpoints as conservative or more conservative than RBC data. For the potent agents aldicarb and carbofuran, however, the BMD_{10} based on RBC enzyme is lower than one based on brain AChE. If that is truly the case, what does it say about the overall conservativeness of the process?
- In presenting a series of equations it would be helpful to number them all, to define the x and y side of each one, to specify units, to include a nomenclature box, and to minimize confusion by using brackets.
- When considering alternatives, it would be useful to specify the criteria for a good, adequate, or acceptable model, specify what is expected from each model, and specify how the performance of each model will be assessed.
- The main document should include better references to and descriptions of the material in the appendices.
- A flow diagram is needed to show how the final models are derived. An example could be given, followed by a reference to the appropriate appendix (II.B.2). There are many files in this appendix but they appear to follow a common methodology. Some explanation of the purpose of each step in the analysis would help. Also, there should be a summary of the results for each of these cases, so that a reader can reproduce the final dose-response and time response equations.

- Graphical or other appropriate quantitative information is needed to support the assertion that the "nonstandard" modified Ellman method is reliable (p.30).
- On pages 32-33 several key assumptions for the empirical dose-response and time course models are listed. This is very helpful. However, to be fully convincing, empirical observations that support these key assumptions should be provided to show that they were derived based on interpretation of data.
- The documentation of the dose-response model for inhibition could be more clearly explained, as for example, by defining the response g(d).
- Units should be given for all variables.
- Equation 1 lacks clarity. For example, is the log in the exponent a log 10? Does this log take as its argument both terms, including the one raised to the γ power? Probably not. This could be made clearer with an appropriate equation editor.
- It also would help to have an explanation up front as to what general modeling approaches were used with respect to pesticide and exposure route. For example, after slogging through Equations 1-5, the reader is told that other modeling approaches were used. Each and every model actually used should be documented, not just alluded to. A sensitivity analysis of each generic type of model should be provided to illustrate (and explain/justify) the key behaviors of these models.
- The final parameterized version of the model should be given explicitly, e.g., on page 36.
- Discussion of the biological aspects of inhibition and recovery would be very helpful – to establish, perhaps, the biological plausibility of the models used.
- Clear discussion of data quality objectives would help – for example, what constitutes a "good model" in terms of precision or accuracy relative to available data, and are the models intended to be somewhat qualitative in simply describing trends rather than providing exact estimates?
- A footnote on table I.B.6 would help the reader understand what the values in parentheses are for some of the chemicals.
- The point of Figure I.B.3 would be better illustrated if RBC and brain cholinesterase BMD_{10}'s were shown separately for each chemical, side by side, rather than in the overlapping graph presented.
- A user's guide to DRutils should be provided.
- The issue of what happens to model estimates when data are collected after the peak should be addressed explicitly. It is worth trying to detect differences in slopes calculated with and without the first two points. One must determine the sensitivity of bias in this estimation and also deal with the issue that the third data point may not be well estimated.
- There should be a discussion of how often the gamma parameter (γ) <u>does</u> differ statistically from 1.
- It should be pointed out that the present model is rooted in the mechanistic construction of previous models for the OP anticholinesterases.

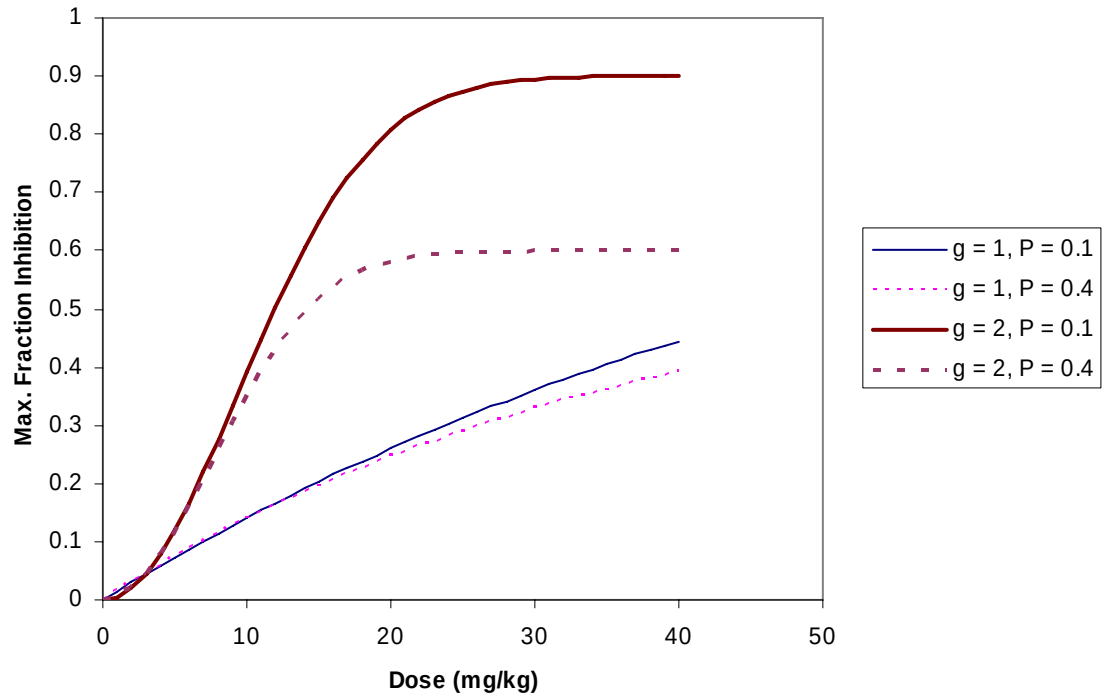


Figure 1 Sensitivity of the dose response model equation to different levels of γ for D_R=3 and R=0.1.

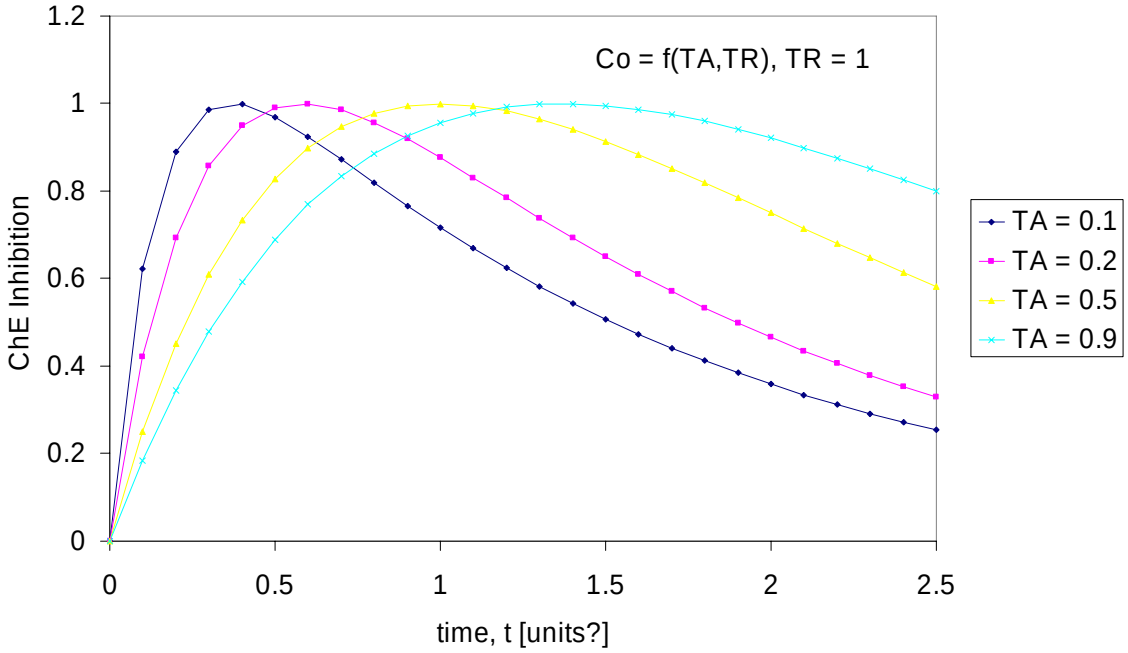


Figure 2 Impact of different Ta values on the time pattern of response.

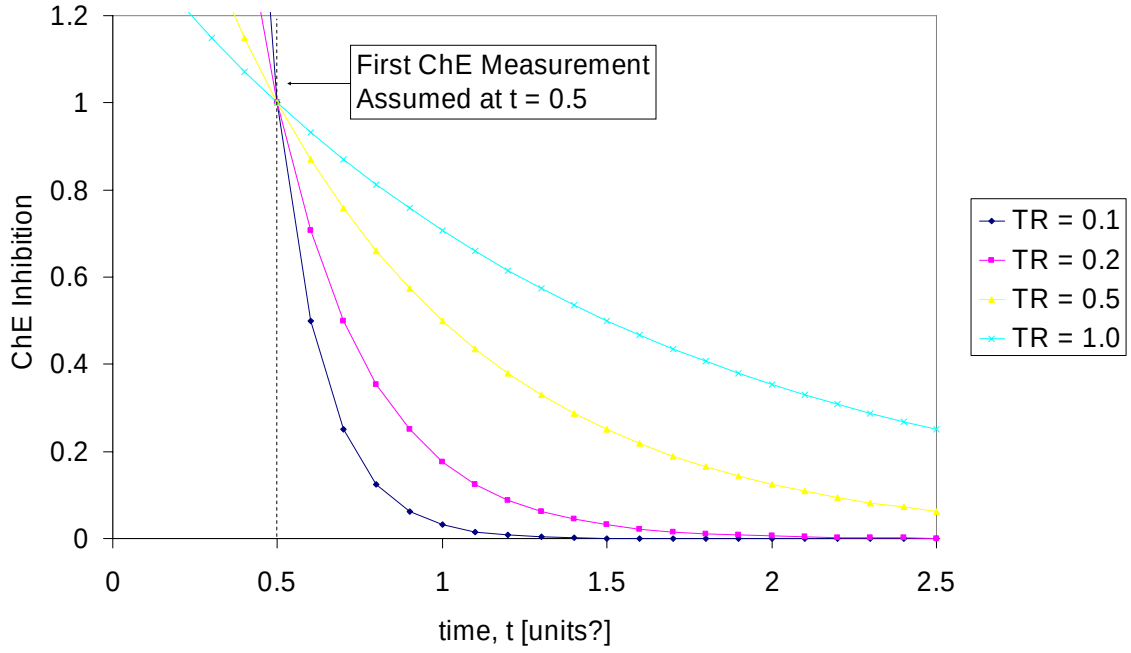


Figure 3 Impact of different Tr values on the time pattern of response using the simplified recovery model.

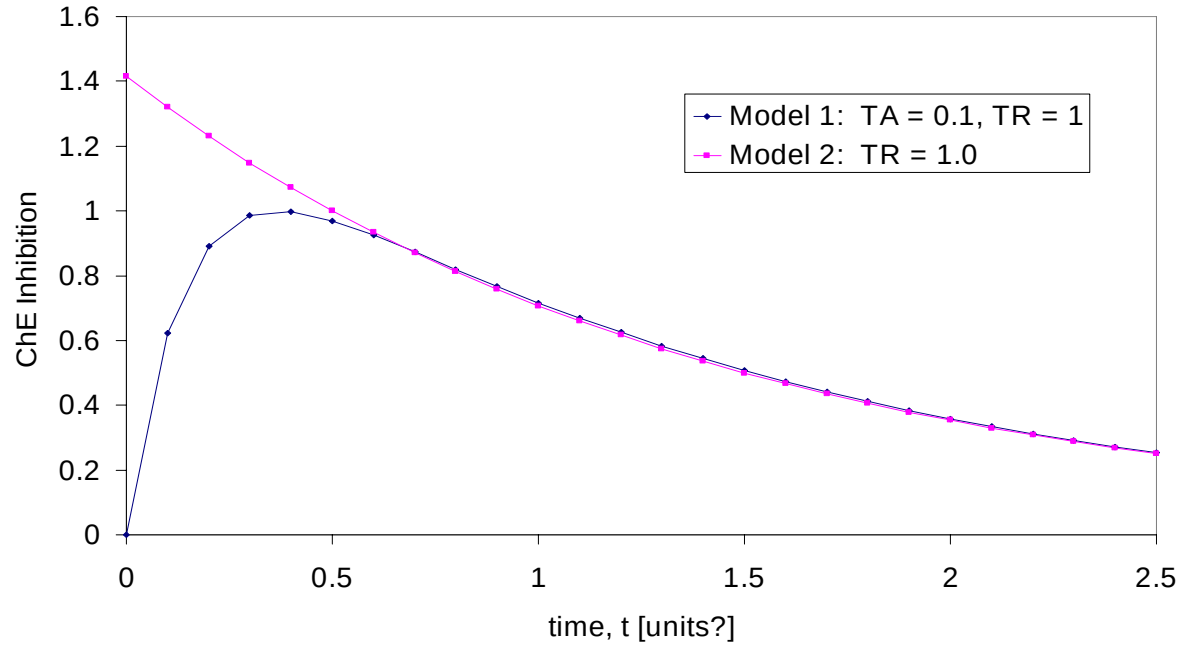


Figure 4 Comparison of the full model and simplified model for recovery over time.

HAZARD QUESTION #2

Selection of the Index Chemical

EPA's cumulative risk assessment guidance indicates that the index chemical should be selected based on the availability of a high quality toxicity database for the common mechanism endpoint. The selection of the index chemical is an important step in the cumulative risk assessment; the BMD for oxamyl was used to calculate RPFs and the BMDL for oxamyl was used as the PoD for extrapolating cumulative risk.

H2. Please comment on the rationale provided for the selection of the index chemical. Should any additional factors be included in the rationale for the selection of oxamyl as the index chemical?

Response

The preliminary N-methyl carbamate cumulative risk assessment designates oxamyl as the index chemical for calculating Relative Potency Factors (RPFs) for the other nine pesticides and for pesticide mixtures using an additive approach. In principle any of the ten pesticides could be selected as the index chemical, but the EPA proposes that the optimal choice is a pesticide with robust experimental data on an endpoint of interest (in this case brain AChE activity) from studies with all relevant routes of exposure.

The Panel recognized the strengths of the oxamyl database and generally agreed with the rationale for its selection as the index chemical. Oxamyl is one of only two chemicals for which the database includes studies with all three major routes of exposure: oral, dermal, and inhalational. Nonetheless, several Panel members argued that carbaryl should be considered for reasons addressed below.

The oral exposure database for oxamyl comprises three acute rat dosing registration studies and the EPA NHEERL rat dose-response and time course studies. The doses used covered a wide range, and whole or half brain ChE data were available from all four studies. Recovery data also are available although it is not stated how many studies included this endpoint. Calculated BMD_{10} values for brain ChE activity differed statistically between the sexes, but the differences were not biologically important, and the 95% confidence intervals were small. The relative potency of oral oxamyl for blood and brain ChE inhibition is fairly high among the 10 pesticides (Fig. 1.B.3) a possible negative for its selection if this introduces a systematic bias when used to estimate RPFs for pesticides with markedly different pharmacokinetics. Adding to this concern was that the acute oral dose response curve for brain ChE inhibition for oxamyl did not fully parallel those of the other pesticides in the NHERRL data. With regard to the chemistry of the pesticides, six of the compounds have an aromatic or heterocyclic ring whereas four do not, oxamyl among them. Thus, although hardly an outlier, oxamyl differs chemically from 60% of the compounds in the common mechanism group. In contrast,

and a potential plus, the water solubility of oxamyl is intermediate for the set of common mechanism compounds. Oxamyl uses are limited to agricultural applications, and human exposure is projected through food or drinking water. Oxamyl has no residential use.

For carbaryl, five oral administration studies are included in the database. Recovery half-life estimates for brain ChE after oral dosing with carbaryl were somewhat slower than for oxamyl, and were shown to increase with dose. The availability of information about the dose dependency of recovery with carbaryl contrasts favorably with the single dose recovery data for oxamyl. However carbaryl's relatively low toxicity was considered a drawback for an index chemical in a group of pesticides that are generally one or two orders of magnitude more toxic. The low toxicity suggests that detoxication and clearance factors are more important for carbaryl than for the other pesticides. Therefore using carbaryl as the index chemical might distort the assessment of the more toxic compounds whose detoxication and clearance are likely less important.

Two dermal studies were available for oxamyl, both in the rabbit. The BMD_{10} values were, as expected, much higher than those from the oral studies, probably owing to kinetic and/or species differences. Unfortunately the studies on rabbits used only dermal exposure, so that we cannot assess species differences in response by comparable routes, or compare dermal exposure with oral exposure in a single species. This deficiency weakens the case for oxamyl as the index chemical. For carbaryl, on the other hand, the one dermal dosing study was performed in the rat, allowing direct comparison with oral gavage studies in that animal. The document states that the resulting data were sufficient to calculate BMD_{10} values for RBC and brain ChE inhibition.

One study of acute inhalation exposure is available for oxamyl and propoxur (appendix II.B.2). None exists for carbaryl, which is unfortunate since this agent has residential uses that may generate exposures by that route. A PBPK model under development might eventually help to estimate BMD_{10} for inhalational exposure to carbaryl. Meanwhile, only oxamyl and propoxur have been studied by all routes and the latter suffers from an unusually large confidence interval in the BMD_{10} data for brain ChE in acute oral studies.

In light of all available data, the Panel generally agreed it is appropriate to use oxamyl as the index chemical. If the Agency goes forward with this choice, however, there are a number of underlying assumptions that should be pointed out, succinctly and explicitly, as an aid to understanding.

All ten pesticides in this class act on ChE by carbamylating its active site serine and modifying the enzyme in exactly the same way. Therefore, the intrinsic recovery rate should be identical for each agent, quite independent of the leaving group. Differences among recovery rates experimentally observed in vivo must reflect differences in the persistence of residual unreacted pesticide, available to inhibit the enzyme. To approach this issue mathematically requires knowing the time course of the pesticide levels. Some of the necessary information may be available in the absorption, distribution, metabolism,

and excretion (ADME) data developed for these agents. Alternatively, rough estimates of the residual active pesticide may be obtained from the degree to which enzyme recovery is delayed when compared to the recovery from oxamyl, the agent with the shortest recovery half-life in vivo. It should be remembered that the recovery rates reflect an integration of the parent pesticide elimination kinetics and the constant rate for reversal of enzyme carbamylation. This is why "recovery half-life" has a complex meaning. None of these considerations, however, suggest that the recovery information has been incorrectly incorporated into the current assessment.

The calculated BMDs for brain ChE inhibition are based on applied doses. While this approach is fine for the purposes of this risk assessment, it may again assist in understanding if the BMDs derived by the three routes of administration are put into perspective. The BMDs for oxamyl are essentially the same for the oral and inhalation routes of exposure, but are much higher for the dermal route. This is what would be expected if acute oral exposure and inhalation exposures both resulted in rapid uptake and distribution to the brain. The much higher BMD for dermal studies probably reflects the slower kinetics of absorption, although the species problem cited above makes it impossible to say this with certainty based on these data. These arguments clarify why it is best to have information from all three routes of exposure when selecting an index chemical, rather than relying on a possibly superior dataset in studies of oral exposure alone, as with carbaryl. These considerations also support the use of the data from oral studies for pesticides for which dermal or inhalation data are absent, as an adequately protective, or conservative approach to the cumulative risk assessment. In this regard, it is better to use the oral RPFs, rather than the BMD_{10} (as done for Table 1.B.8) to derive RPFs for other routes of exposure on which we lack data.

Other arguments were put forth and supported by some of the Panel, in favor of selecting carbaryl as the index chemical. These points follow.

The cumulative risk assessment guidance document (USEPA, 2002a) states that the criteria for an index chemical should be high quality dose-response data on the common mechanism endpoint, preferably with each exposure route, and a toxicology that resembles other agents in the common mechanism group. Even so, the selection of the index chemical should consider real-world uses. It is understood that use of an index chemical with imperfect toxicology data may introduce error and uncertainty into the estimation of cumulative risk. However, the PoD for the index chemical will be used to extrapolate risk to exposure levels anticipated in the human population. Therefore the selection guidelines for the index chemical should take into account both the potency of the agent and the range of uses for which it is registered.

Oxamyl has very limited use in agriculture, and has no residential applications. Regardless of their quality, the toxicological data from dermal and inhalational studies of oxamyl, and to a lesser extent the oral studies, indicate that this compound contributes little to the cumulative carbamate risk estimates in the human population. Almost all the NMC cumulative risk would result from exposures to pesticides other than oxamyl, over

dose ranges such that the responses, either brain or RBC ChE effects, would need to be extrapolated from the oxamyl dose-response model. Thus, greater error and uncertainty may actually be introduced into the cumulative risk estimates if oxamyl is chosen as the index chemical.

The wide use of carbaryl, compared to other NMC members, both in agricultural and residential environments should make carbaryl a favorable candidate for the index chemical. Because of its wider use, selecting carbaryl as the index chemical would require less data conversion in the cumulative carbamate exposure assessment, which should minimize the inherent error and uncertainty associated with the cumulative risk assessment involving relative potency factors. In addition, the development of a PBPK/PBPD model for carbaryl has been encouraged by the Agency and is currently being undertaken. It was intuitively apparent to some Panel members that the selection of the index chemical for the derivation of relative potency factors should parallel the effort of PBPK/PBPD model development.

HAZARD QUESTION #3

Selection of Brain ChE data for developing RPFs and PoDs

EPA has used data for brain ChE as the basis for the RPFs and PoDs. The rationale for this selection was provided in I.B.

H3. Please comment on the rationale provided for the selection of the brain ChE as the basis for RPFs and PoDs in the preliminary cumulative risk assessment. Should any additional factors be considered?

Response

The Panel found a compelling case for using brain cholinesterase as the endpoint in determining relative potency factors. Brain AChE is abundant and critical to normal physiologic function. Brain tissue is readily removed and homogenized in a reproducible manner. Lastly, inhibition of this enzyme is not simply an index of exposure but is an integral portion of the common mechanism of toxicity for N-methyl carbamates.

The alternatives, to mention them briefly, would be to use 1) AChE inhibition in blood (i.e., RBCs); 2) AChE inhibition in a peripheral tissue such as muscle, nerve, gut, or heart; or 3) clinical signs and behavioral disturbances. Each of these alternatives has serious drawbacks. RBC AChE is difficult to assay and has no known physiologic function. Inhibition of this enzyme is therefore at best a surrogate for the actual mechanism of toxicity. Likewise, although AChE inhibition in peripheral tissues might ultimately provide a sensitive and direct index of toxicity, there are no extensive data to support this concept as yet, and the accurate dissection and assay of such tissues requires care and skill. Finally, when it comes to clinical and behavioral observations, although these measures are relevant and qualitatively informative, as endpoints they are more

subjective and, generally speaking, less sensitive than biochemical determinations.

Further support for focusing on brain AChE comes from NHEERL data summarized in the current document. These data showed that BMD_{10} values for AChE inhibition in the brains of carbamate-dosed rats were, on average, as low as those derived by measuring the enzyme in RBCs. Furthermore, as an Agency expert explained at the meeting, confidence limits on a brain BMD_{10} are as tight as those on red cells, despite the advantage of repeated measures analysis in the latter case. This statistical advantage is probably offset by the low activity of AChE in rat red cells and the problems of accurate quantitation in the presence of hemoglobin. In any case, methods now under development may further increase the precision of brain AChE determinations with rapidly reactivating inhibitors and increase the utility of this metric.

One Panel member suggested establishing BMD_{10} values for inhibition of AChE in multiple sub-regions of brain, rather than in whole brain or brain hemispheres. Such information could lead to RPFs and PoDs based on the most sensitive target area. If brain regions are eventually found to differ sharply in vulnerability to pesticides, a regional analysis will become essential. At present, however, there are good reasons for going forward with whole brain AChE. First there is little evidence to suggest large regional variations of AChE inhibition in brain after systemic exposure to a carbamate anticholinesterase. In one of the few papers to address the issue (Hammond et al., 1996) brains were micro-dissected 1.5 hr after rats were gavaged with carbaryl (50 mg/kg). The results showed a small variation from the most sensitive regions (cortex and striatum at 55% inhibition) to the least sensitive regions (inferior colliculus and hippocampus at about 40% inhibition). It is worth noting that the more sensitive areas together account for a large fraction of the total brain AChE (the total mass of cortex is large, and striatum contains disproportionately high concentrations of the enzyme). Thus, AChE inhibition in whole brain will not be drastically lower than in the “sentinel regions”. A second reason for accepting measurements in whole brain samples is that, even in skilled hands and within a single laboratory, regional dissection can lead to data with much greater variability than data from whole brains. Variability across multiple performance sites is likely to be unacceptably high, even if EPA’s time frame allowed for a data call-in. In summary these considerations speak decisively in favor of the Agency’s plan to use existing data on whole brain AChE as its basic metric. The only qualification to be added is an assumption that the database is confined to studies for which the Agency has documented Standard Operating Procedures (SOPs) that demonstrate care to minimize distortions caused by rapid reactivation of carbamylated enzyme.

Turning to issues of communication and presentation, the Panel finds a need for more detail in the Preliminary Cumulative Risk Assessment and also a need for additional references to support the Agency’s rationale and conclusions. For example, part of the rationale given on page 27 for choosing brain AChE inhibition as the basis for the point of departure is that this metric is as sensitive as behavioral measures of toxicity, or more sensitive. That statement is justified with reference to internal EPA studies summarized in Appendix II.B.5. These studies concluded that traditional clinical

measures of cholinergic signs (salivation, lacrimation, urination, and defecation, abbreviated as SLUD) are less sensitive than effects on motor activity, and that these in turn are less sensitive than inhibition of blood and brain ChE. The underlying data directly support the concept that adverse effects of ChE inhibition in the peripheral and central nervous systems are adequately prevented by preventing effects on brain ChE. Appendix II.B.5, however, reports only locomotor activity and toxic signs for each of 7 carbamates. It states that correlations between behavioral outcomes and changes in cholinesterase were analyzed but never shows them. Explicit presentation of this information would strengthen the Agency's case for its choice of metric. The case would be further improved by fuller reference to the EPA data in Appendix II.B.5 in which the levels and variance of RBC and brain AChE inhibition are directly compared.

Elsewhere the document needs better referencing. For example, page 27 provides a general discussion of endpoints in toxicology studies with N-methyl carbamates. Part of this discussion simply states without reference that behavioral measures often lack standardization, are variable, and less sensitive to disturbance by carbamates than are measures of peripheral or central AChE activity. Also largely unreferenced is the discussion of the difficulties associated with measuring ChE inhibition in the peripheral nervous system, and the problems with assays of whole blood that do not distinguish activity from AChE in red blood cells and butyrylcholinesterase in plasma.

WATER

WATER QUESTION #1

Revised Conceptual Model for Ground Water

Based on recommendations of the February 2005 SAP, OPP revised its ground water modeling approach to estimate pesticide concentrations in the upper meter of a fixed saturated zone (ground water) that starts at 3.5 m below the surface. The Agency has included two additional adjustments to the original conceptual model since the earlier SAP. The models consider variable degradation rates with depth and account for setback distances between the well and the application area by using lateral velocity to estimate the additional travel time for a pesticide to reach the well.

W1. Please comment on the Agency's revisions to the ground water modeling approach to account for variable degradation rates with depth and varying setback distances between the well and treated fields.

Response

EPA has made great progress in modeling pesticide movement in the vadose and ground water zones. The revised ground water modeling approach provides a more realistic representation of conditions than was used for the February 2005 SAP. The

revised modeling is still adequately conservative. Although the model designers did not correct all errors found in the computer code, remaining problems are small and overall results are impressive.

For the simulated profiles (high conductivity soil with no ponding of water on a restrictive layer) a degradation rate established for aerobic metabolism in lab studies is reasonable in the top 25cm of soil. For the zone from 25 cm to 1 m the Panel also agrees that, unless there are data to the contrary, the best assumption is a linear decrease from the aerobic metabolism rate to the abiotic degradation rate. On this point, however, the Agency should check the report of Ou et al. (1988), who measured degradation rates with depth in Florida.

The concentration profile predicted by the model is quite "peaky". Data from Long Island (Steenhuis et al., 1987) indicated quite a bit of spatial variation in the movement of the chemical with slow and fast paths, giving a less peaky profile. This means that the average of the predictions is likely more realistic than the temporal concentration profile. Smoother profiles can also be simulated by running the model several times with different fluxes, each passing through a portion of the soil profile.

The ground water routine is refreshingly simple and represents processes well. However, EPA is advised to consider the following points.

A value of 0.15 m/day seems to be at the high end of groundwater velocities in "real" aquifers for water flowing with a natural gradient, but it probably underestimates the velocity in aquifers in valley bottoms with rivers. The assumption is sound that the presence of domestic wells does not affect travel time of pesticides. However, wells used for irrigation might alter flow patterns in surrounding ground water to a larger degree, and the induced velocity fields could be many times greater than the natural gradient-induced flows. This might explain the high pesticide levels in the Florida data set, even in wells with a substantial offset. Since the MOE for central Florida sites is strongly influenced by the carbamate concentrations in drinking water, it is important to decide how to handle the effect of irrigation wells on groundwater concentration.

Although the proposal to average pesticide concentration over a 1-meter depth interval is technically correct, a plan to screen for pesticide just below the surface of the ground water is physically unrealistic for wells with an offset. Consider a residential drinking well that is located 300 ft from the edge of a field, with water flowing under a natural gradient of 0.15 m/day. At this rate, water will take 600 days to travel from field to well. Along the way, clean water recharges and will push the pesticide deeper. Assuming a recharge rate of 36 cm/year, a porosity of 0.4, and no diffusion, chemicals eluted from the field will be 2.4 m below the top of the aquifer by the time they reach the well. Therefore pesticide levels should be screened at least 2.4 m below the ground water. Because the current model assumes that the wells are in fact screened deeper than1 m, only the wording in the document needs to be changed. A problem that is based more in reality, on the other hand, is that some diffusion, dispersion, mixing, and dilution

with recharge will occur throughout the setback distance. To deal with this problem EPA may want to use a simple model (e.g., KYSPILL - developed by Sergio Serrano, sserrano@temple.edu) to estimate changes in pesticide burdens along the path from source to well. This will assure that the equations proposed are reasonable. In addition the model can be used to assess the effects of pumping-induced flows by irrigation wells.

In summary, the Panel supports the approach to identifying the spatially variable nature of pesticides reaching groundwater. OPP should continue to pursue this approach.

Other minor comments.

It was assumed that Aldicarb degrades by a first order process. In reality Aldicarb degrades to some byproducts that are just as toxic. Therefore the latter process should be taken into account and can be simulated by using the first order degradation rates of the byproducts. Liu et al (2003) have shown that Aldicarb degradation is much faster than the degradation of its byproducts. Therefore an assumption in the model that Aldicarb instantaneously is converted to its byproducts would likely not introduce large errors

The set back distances and travel times to them should be checked.

On page 4 of section II.D.7, it is stated that in order to implement the irrigation routine and obtain correct irrigation rates, the depth of the root zone in the PRZM model was decreased. Reducing the depth of the root zone might have an unexpected effect on the amount of evaporation and thereby increase the amount of recharge compared to the other models that have the correct root depth. See the response to question W2 for further discussion of the accuracy of water balance.

The Agency's working document defines C_0 as the concentration of pesticide at the point of application. Instead it should be the concentration at the point where pesticide enters the ground water.

Tables II.d.7.1 through II.d.7.4 use a mix of English and metric units. In some cases the same numerical values appear with different units. Copy editing and proofreading is advised.

WATER QUESTION #2

Comparisons of the Three Models

The three models used by the Agency (PRZM, RZWQM, and LEACHP) provided predicted concentrations that were similar on average, but short-term concentration differences among the models varied considerably. Differences in peak concentration estimates ranged from a factor of 2 to 5 in Florida to as much as a factor of 20 in North Carolina; however, there was no consistency with regard to which model gave the highest

or lowest predictions. Some of these differences may due to differences in the way the models handle degradation-temperature relationships, evapotranspiration, and weather generation.

W2. Given that no model stands out as superior when compared to the monitoring data evaluated so far, can the SAP suggest criteria for further evaluation of the models?

Response

To facilitate comparison of the three models (RZWQM, PRZM, and LEACHP), a common weather file should be used as input for all of them. Historical weather data is preferable to CLIGEN (climate generator), which may not simulate subtropical weather accurately. An appendix summarizing the major input parameters and the rainfall characteristics would help readers understand the modeling scenarios more fully.

Before predicting pesticide losses, it is important to investigate the hydrology. The most important hydrologic consideration is an accurate water balance. For that purpose the estimate of evapotranspiration is crucial. One study on Long Island measured recharge over a sixth month period (Steenhuis et al., 1985; Steenhuis and van der Molen, 1986). Information from this study can be used to check the water balances of the current simulations. Another hydrologic consideration is the expected water flow out of the unit area (or unit volume) given the lateral groundwater velocity and aquifer porosity. Consider calibrating the three models so that predicted flow from the unit area is similar with each. With RZWQM, tile flow calibration involves adjusting drain diameter, lateral hydraulic conductivity, and "effective porosity" or porosity minus field capacity (Singh et al., 1996). With PRZM the calibration parameters to equate predicted and observed flow from the unit volume may be saturated hydraulic conductivity and/or porosity of the aquifer. By using the calibrated models, the simulations of pesticide fate can be assessed without confounding influences from uncertainties of water transport. In any case, understanding the hydrologic balance of the three models will greatly help model comparison.

As always it is important to consider how the chemicals are sampled. The sampling protocol is questionable in light of the curves in Figure II.D.7.17. Assuming that the bromide flux is the same for the model and the observations, then the mass under the simulated and observed curves should be the same. The mass for the observed data is considerably less (by about half) than for the simulated data. This discrepancy points toward a sampling problem because the model should be able to simulate the correct total mass. The difference between observed and predicted bromide flux may be due to the physical sampling process. In particular, pumping of wells used for irrigation could cause mixing from a depth increment that would not be included in the model predictions.

The peak pesticide concentrations in groundwater differed considerably between

the three models. Peak exposures are important for this risk assessment. To identify the source of the model differences, including peak concentrations in groundwater, a pesticide mass balance and a hydrologic balance are both important. Factors to consider with both pesticide and water include: application, runoff, percolate (or lateral flow) out of the profile, tile-drainage, pesticide degradation in profile, plant uptake, and evapotranspiration. Without completing a hydrologic and pesticide mass balance on the three models, it is difficult to compare the models.

The Panel suggests the following ways to evaluate the models under consideration. The current simulations are primarily from areas of low soil carbon, sandy soil, and low sorption coefficient (Koc). These conditions are consistent with "…drinking water that is expected to be among the most vulnerable…" (p. 88). Vulnerable sources, however, should include high intensity rainfall within a few days of pesticide application where macropore flow occurred. The current version of the NMC Cumulative Risk Assessment does not specify if macropore flow was simulated to occur and does not specify when intense rainfall occurred in relation to application and other rainfall. Even if macropores were parameterized for RZWQM, rainfall must be intense enough and soil properties such that macropore flow actually occurred and reached groundwater. On structured soils, low intensity rainfall after pesticide application reduces pesticide transport in macropores during subsequent rainfall (Shipitalo et al., 1990).

Pesticide leaching under some circumstances can be greater on structured soil than on sandier soil. Sadeghi et al. (2000) concluded that intact and repacked silt loam soil laboratory columns leached more atrazine than a sandy loam soil that had less carbon and less clay content because of more macropore flow from the silt loam soil.

For the high conductivity soils in Florida, macropore flow may be less important than on structured soils because few storms will exceed the saturated conductivity of the soil. Preferential flow, however, is still important on high conductivity soils because of fingered and/or funnel flow (Kung, 1990; Glass et al., 1989). This is consistent with the conclusion of Jarvis et al. (1994) that it was important to model preferential flow in order to accurately predict herbicide leaching on sandy textured soil. One method to predict the portion of preferential flow within the profile of sandy soil is to divide the maximum intensity of the rainfall or irrigation (on an hourly basis or less) by the saturated conductivity of the soil (information about this approach can be found in Darnault et al., 2004; Kim et al., 2005; Selker et al., 1996). Another way to estimate the preferential flow area of the soil is to calibrate the model with observed data using an inert non-adsorbing tracer such as bromide.

In summary, the Panel strongly favors comparing the hydrologic and pesticide balance of the three models. Another important step to consider, focusing on the most vulnerable groundwater, is to simulate preferential flow shortly after pesticide application. Many points must be considered in choosing which model to use for the present purpose. Among these is the question of which model most reasonably

represents the expected processes (pesticide, hydrology, and preferential flow). If thorough comparison doesn't identify a superior model, however, it may be best to use the simplest model.

WATER QUESTION #3

Evaluation of the Ground Water Model Estimates

The Agency compared NMC concentrations in ground water estimated with the three models (PRZM, RZWQM, LEACHP) to results of available prospective ground water monitoring studies (oxamyl in NC and MD and methomyl in GA), two well-monitoring studies along the central ridge of FL, and published literature on in-field monitoring studies. Using the FL well monitoring data, known fate characteristics of the NMC pesticides, and soil and hydrologic data, the Agency identified the conditions under which exposures similar to that estimated in the NMC CRA may occur: private wells drawing from shallow, acidic ground water with high to very high saturated hydraulic conductivities in the soil and vadose zone. This has allowed the Agency to move toward a spatially-explicit characterization of potential high exposure areas.

W3. Please comment on the performance of the models against the available monitoring data. What additional considerations should be taken when applying modeled estimates to risk assessments for areas where monitoring data are not available?

Response

Model Performance

There are no real surprises in performance of the models relative to the available monitoring data. The magnitudes of the differences in model predictions are expected given the differences in the models and the input data they use. The performance of the models is reasonable. Statistical measures, such as Nash Sutcliffe coefficients, R^2, and RMSE, describing model performance relative to observed data would be helpful in assessing and comparing model performance.

It would be useful to know which of these models is best at estimating the hydrology for the cases being modeled. The model that best estimates the hydrology would have the potential to perform the best in estimating pesticide concentrations in shallow ground water. If the hydrology is incorrect, it is difficult to estimate the pesticide concentrations correctly given that the movement of water is the transport mechanism for the pesticides. In such a case pesticide concentrations may at best be correct for the wrong reasons.

Since the models were validated for only two locations, it may be worthwhile to

compare and validate the models with additional observed data. Observed data from eastern Canada could be one of these data sets (http://www.hc-sc.gc.ca/ewh-semt/pubs/water-eau/doc_sup-appui/aldicarb-aldicarbe/index_e.html). This reference indicates that the following data are available:

> "In a survey of 317 wells in eastern Canada in September 1986, aldicarb was detected in 167 of 782 samples; concentrations were above 10 ppb in only 9% of the 167 samples. In surveys of private and municipal drinking water supplies in five Canadian provinces, conducted from 1980 to 1986, aldicarb was detected in 111 of 1017 samples (detection limits ranged from 0.01 to 3.0 µg/L); the maximum concentration was 28 µg/L. In Prince Edward Island during 1985 and 1986, 77 of 96 samples (80%) in two areas contained aldicarb residues above the detection limit of 0.1 µg/L, with a maximum of 16.4 µg/L. In the same province, ground water quality was monitored between 1985 and 1988 near two potato fields to which aldicarb was applied at planting once or twice between 1983 and 1986. In May 1988, concentrations of aldicarb plus its degradation products exceeded 9 µg/L in 12% of 48 well samples. Residues of aldicarb and its sulphoxide and sulphone have also been reported frequently in water samples in a number of U.S. states; concentrations are typically in the range 1-50 µg/L, and a maximum of 400 µg/L was recorded in one case from Long Island, New York."

Models should produce distributions similar to those in monitoring data. The ability of PRZM, RZWQM, and LEACHP to produce distributions of pesticide concentrations in ground water relative to monitoring data should be explored and characterized.

From the information provided, the Panel concurs with EPA's statement in the ESTIMATION OF CUMULATIVE RISK FROM N-METHYL CARBAMATE PESTICIDES: Preliminary Assessment document that "There is no clear "best" model to use to assess pesticide concentrations in ground water." However, this conclusion could change after EPA considers the Panel comments from questions W1, W2, and W3.

Additional considerations when applying modeled estimates to risk assessments for areas where monitoring data are not available

Monitoring is necessarily limited. Therefore, models are used to extrapolate beyond the monitoring data. The key factors that identify areas with the greatest potential for pesticides reaching ground water have been captured in the process that is being used. However, some additional factors and steps should potentially be considered.

The models can be readily extended to locate regions likely to have high concentrations of carbamates in ground water, although some further validation is desirable. The model simulations, as verified by experimental data from Florida and experiences in Long Island, show clearly that elevated levels of carbamate pesticides (and especially high aldicarb concentrations) are probable in locations with sandy soils of

high conductivity, low organic matter, and ground water with a pH< 7. These sites have travel times on the same order of magnitude as the abiotic degradation half-life of carbamates.

To predict oxamyl concentration in ground water, the dissolved oxygen in the ground water also might be an important factor. The number of pesticide applications is important too.

It is unclear if the monitoring data to which the model outputs were compared included intense rainfall shortly after pesticide application. A worst-case modeling scenario should include a high intensity rainfall shortly after pesticide application where macropore flow occurs. Research suggests that preferential flow occurs on sandy soil. On structured soil, the most significant preferential flow event is generally during the first intense rainfall after application. Therefore, the Agency should consider a scenario in which intense rainfall shortly after application leads to macropore flow to shallow ground water. Additional discussion, rationale, and references concerning macropore and preferential flow are provided in the Panel response to the W2 question.

Experts may be able to identify other regions in which high pesticide concentrations might reach ground water. It may be desirable to consult with informed individuals in various regions where carbamates are applied, since limited data are now available on ground water quality in these areas. In particular, the quality of available spatial data does not allow us to identify very small areas in regions that have just a few fields with high carbamate use. Such areas may not be important to the national level assessments but could present high exposure risk for a small number of people.

Miscellaneous comments on water questions

Although it is assumed that the private wells are the most sensitive, it is not unlikely that the municipal wells can have also carbamates in the drinking water. For example in the city of Owen Sound in Ontario a sample taken in late summer of 2000 had a Carbofuran concentration of 2 ppb as the only pesticide (**http://city.owen-sound.on.ca/water/2000-thirdquarter.pdf**).

Isolated cases of ground water pollution of domestic wells below pesticide treated fields will only affect a few locations in certain regions but possibly cannot be ignored in the aggregate.

FOOD

FOOD QUESTION #1

The food portion of the N-methyl carbamate cumulative risk assessment used similar data sources and techniques to those used for the organophosphate pesticide for

estimating cumulative risk from food. This included use of both the USDA's Continuing Survey of Food Intakes by Individuals (CSFII) as a data source for food consumption and Pesticide Data Program Data (PDP) as a data source for food residues.

F1. Please comment on the planned intermediate- and longer- term activities associated with sensitivity analyses identified in Section I of the document. Does the Panel have any suggestions for other or additional activities which the Agency should consider?

Response

General Comments

In a broad sense, the Panel agrees with the Agency's intermediate and long-term activities, and appreciates the Agency's responsiveness to the Panel's comments in the previous SAP meeting regarding the food component in the N-methyl carbamate cumulative exposure and risk assessment. Now the Panel would like to see the Agency prioritize its activities so that the key outcomes that may help improve the dietary component of the cumulative exposure assessment will be available in the near future. A series of recommendations are offered here.

Intermediate Term

- Conduct a more detailed analysis of food exposure to identify major contributors to risk, identifying specific food-pesticide combinations.

Identifying individual foods and food classes likely to make major contributions to dietary exposure is an excellent next step, and one that can be accomplished using the databases described. It also is of interest to identify major contributors that are age-specific, for example, the differential sensitivity of children to dietary intake of carbamates.

Of further potential interest are NHANES data on biological markers of exposure, e.g., urinary 1-naphthol, that can be correlated with certain intake data. EPA may wish to explore the use of such markers as part of the overall risk calculations.

One Panel recommendation is that the Agency plan to gather longitudinal dietary consumption patterns from individuals living in different regions of the country. Autocorrelation and "anti-autocorrelation" are likely in dietary intake and cannot be obtained from cross-sectional data. For example, some individuals may tend to eat foods drawn from a relatively small fraction of the total possible items due to preferences that may be purely personal (e.g., a vegetarian diet) or based on culture, ethnicity, and geographical location. Such data differ from the data compiled by the CSFII, which is a cross-sectional survey with (one-time) repeated sampling within a 10-day period. The

CSFII lacks information on each individual's long-term dietary consumption pattern. This limitation can be problematic since certain food commodities are more likely than others to contain NMC residues. Furthermore, such differences may have seasonal components that are overlooked in rolling, cross-sectional studies. A longitudinal study of consumption is a daunting task; however, recent improvements in survey tools should facilitate the process.

- Conduct a series of sensitivity analyses for input parameters that are most likely to impact the outcome of the assessment and determine their effects. The effects of deleting earlier years and of using PDP data translation protocols are worth study.

Sensitivity analyses were discussed during this meeting in other contexts and are of particular interest here. Several kinds of sensitivity analyses can be envisioned including various omissions from data sets to assess the resulting change in estimated parameters. Small changes will increase confidence in the parameter estimates, while large changes, suggesting that such estimates are not robust, may focus attention on the quality and comprehensiveness of the data.

Data omissions to consider in a sensitivity analysis include elimination of earlier years from the PDP data. This elimination may be desirable because pesticide registrations and use patterns have changed and concentrations measured in earlier years are no longer appropriate for modeling current exposures. Sampling strategies for foods have not changed. Thus the data from early CFSII collections can be used and may be indicative of long-term trends in dietary change. Examination of the PDP data for carbamates will indicate numerous foods that do not contain measurable quantities of these materials. It may be worthwhile to test the effects of eliminating such foods.

In general, sensitivity analyses are to be encouraged as they give insight into the potential impact of dietary trends. The long-term change that led from high-fat to lower-fat diets and then to the "low-carb craze" could be modeled. One could look forward and test the effects of possible dietary changes in the future. For example, it would be interesting to know if carbamate intake would rise or fall substantially if Americans began to eat less fast food.

With regard to translation protocols, the Panel finds it acceptable to use residue data from one commodity as a surrogate for another when both commodities are in the same tolerance crop group published in the CFR. Moreover, within a given food form, the processing factor for one pesticide *should* be used for another pesticide when data for all pesticides are not available in the NMC CRA analysis. This is especially true when the processing factor appears to reflect dehydration, e.g., from apple to dried apple.

- Determine how the choice of assumed values for "non-detects" affects the estimated exposure.

Choosing different values for non-detects may well affect estimated exposures and risks. Some compounds could pose significant residual risk at or below LOD because certain groups consume large amounts of material (e.g., children and milk). Even though a given carbamate may not occur at levels above the LOD for a given analytical technique, the total exposure to pesticide, calculated as concentration in food multiplied by the quantity eaten, may nonetheless be significant. Exposures from foods eaten in large quantities may be underestimated if concentrations in samples recorded as "below LOD" are arbitrarily set at zero.

The Agency has apparently conducted a preliminary analysis that showed little impact of LOD values on the final dietary exposure analysis. However, it is important to qualify "little impact." We do not advocate abandoning the assumption of zero for non-detects, and we do not argue for any other value in particular, but we would like better communication from EPA. We recognize that in a cumulative risk assessment, a zero residue for one NMC may not necessarily mean zero residues for all. And we understand that any assumption about the values associated with "non-detects" will have little effect on high-end exposures. Nonetheless, the Panel points out that scoring non-detects with zero values will most certainly distort the shape of the exposure distribution at the lower end. That must happen because the ultimate exposure will then be zero regardless of consumption rate, not the low level that would be computed from any other finite value assumed for the tested samples.

In the long run the best plan of attack is to ensure that the limits of detection (LOD) for the residue-monitoring program are toxicologically relevant (i.e., that a residue at the LOD does not contribute significantly to overall risk). This plan can be accomplished with the aid of better methods in laboratory analysis. Alternative strategies include combining information from the datasets described above and evaluating the most important foods using both presumed residue data and overall intake in a combined fashion. One can then determine in a more quantitative manner how varying LOD values affect total exposure.

- Evaluate the Carbamate Market Basket Residue Monitoring Study and its implications for cumulative risk assessment (particularly with respect to single item vs. composite samples).

This proposed work fits in well with the suggestions made above. Single item samples are of interest since only a few food items may be expected to have levels of contamination deemed "large" by a given metric. Composite samples dilute the effect of the more contaminated samples and may completely mask them by dropping concentrations below LOD. An example cited at the Panel meeting was a sample of 15 apples, one with carbamate residue at 15 x LOD and the rest at zero. A composited sample would still be at LOD and would score as zero. However, an individual who eats the 15 x LOD apple would receive a significant exposure. The most significant point is that individual foods with high levels must be identified and not composited with other,

uncontaminated foods. At the very least, the fraction of high-concentration items must be assessed.

Market Basket Monitoring: Since this survey is based on a single unit of each commodity, it is not "just another dataset." The Agency is encouraged to evaluate market basket data for input into the dietary exposure assessment. Such data should be compared to the PDP composite residue and single unit residue data. The Panel does not know if the Agency has already conducted studies showing that replacing composite data with single unit data does not significantly impact the final outcome of the analysis. If that is true, however, the effort and its outcome should be communicated clearly in the document, in the context of the percentile and population basis that will be used to characterize the final risk.

Long-term:

- Investigate the effect of seasonal residues and consumption patterns on the cumulative assessment.

The Panel disagreed on this question. Most of the Panel considered it is almost essential that the Agency investigate further the effect of seasonal residue and consumption patterns in the cumulative assessment. In fact they recommended that this be done immediately and not deferred as a long-term plan. In a previous SAP meeting, one Panel member raised the seasonality issue and asked why the cumulative exposure assessment model passed along the seasonal effect for water but not for dietary consumption. According to the Agency's own assessment, 65% of total NMC dietary exposure comes from citrus fruits, including orange, tangerine, and grapefruit, whose availability in the US is seasonal. At least one Panel member was concerned that the cumulative exposure model did not transmit this seasonal effect of dietary consumption to the higher end of the total exposure profiles. The majority of the Panel had the sense that the failure to treat seasonal-regional level effects reflects Agency policy, not methodological constraints in the exposure model. The Agency's plan to conduct further investigation to identify specific food-pesticide residues-consumption combinations and their contribution to risk is welcomed and should be pursued as soon as possible.

Another Panel member of the Panel suggested that investigation of seasonal residues and consumption patterns would be interesting but not an immediate priority. According to this Panel member, seasonal effects are likely to be second order. Specific crops are likely to be treated with a specific carbamate, and then stored. The likely exposure will be modified by second-order effects like seasonal changes in food sources, e.g. from locally grown items to imports from other parts of the country or abroad. However, the primary effect is still likely to be food-item-specific. Hence, a study of seasonal effects is appropriately deferred while attention focuses on identifying those foods likely to have high contamination.

- Evaluate the tails of the food exposure distribution to verify that unusual consumption patterns are not inappropriately impacting the results of the assessment.

The CSFII is designed to be representative of the population as a whole. Hence the "tails" of the distribution are still part of the distribution and, therefore, cannot be said to impact the results of the assessment inappropriately. An individual whose diet consists of nothing but the single most contaminated food item may be unusual, even extreme, but is still relevant. Appropriate statistical analysis may be all that is needed to identify such individuals. The analysis of unusual but reasonable eating patterns can begin with the high-contributing commodities identified as recommended under the first bulleted item above. For evaluating the high-end consumption pattern, and with a high contributing commodity, the analysis should look at "user-only" distribution and not "per-capita" data. The joint probability of finding an individual who eats large amounts of heavily contaminated foods is likely to be small, but should also be reflected accurately in the assessment. In any calculation of risk, including this type of exposure assessment, a significant fraction of the burden is often carried by a small number of individuals.

The above comments address the question in the form it was submitted, but probably not its intent. It is an excellent idea to look for unusual patterns of individual consumption in the tails of the distribution. As with mercury in tuna, individuals who eat large quantities of foods that are known to contain a specific carbamate should be made aware of the risks. This is even more appropriate when one examines the cumulative exposure group. While one particular food is unlikely to contain large concentrations of multiple carbamates, several foods, each with modest amounts of several carbamates could generate exposure to the whole class of compounds.

RESIDENTIAL

RESIDENTIAL QUESTION #1

Use of REJV Data and Professional Judgment

To generate estimates of exposure from residential use of NMC pesticides, the probabilistic models use a variety of inputs to address potential exposure from multiple use scenarios. Critical inputs include the percent of households applying the various pesticide products, and the timing of those applications. These two inputs, coupled with potential exposure from pesticide residues in drinking water and the diet, directly impact per capita estimates of cumulative exposure. In its February Case Study, the Agency presented background information on the Residential Exposure Joint Venture (REJV) survey. The Agency used this database as the primary source for data on the inputs relating to timing of applications and percent of households using NMC products. Details regarding the empirical data of the REJV survey are presented in Appendix II.E.1.

In February 2005, the SAP expressed reservations regarding the REJV data. In response to SAP concerns, EPA used other non-survey information in this preliminary CRA, in addition to estimates from REJV, to develop use/usage inputs and seasonal timelines of pesticide use which were representative of the Southern region of the U.S.

As previously mentioned, the REJV survey can be used to generate empirically-based estimates of percent of household use and the frequency of product specific applications. But, because the REJV did not collect information regarding the reason for the reported pesticide use (pest treated) or how much of the product was used, the empirical timing and frequency information (based on a national survey) may not provide a clear picture of regional use. Therefore, to establish the timing of pesticide applications for the scenarios likely to result in the highest exposure, the Agency made these estimates based on a combination of REJV data, product label information, professional judgment, and pest pressure information available from the Cooperative State Extension Services. Specific examples of how these sources were used to determine timing and frequency of pesticide use for PNMC residential assessment are presented in Section E of the preliminary NMC CRA document.

R1. Please comment on the use of information sources other than REJV to establish periods of pesticide use and other use/usage information. Does the Panel suggest an alternative method to improve the use of REJV in the NMC assessment? Does the Panel know of other data sources that may be available?

Response

The Panel is pleased to see more description of the REJV data pertaining to residential exposure of NMC, especially considering that the Panel had been unable to comment on the use of REJV data at the February 2005 SAP meeting due to lack of information on this proprietary dataset.

The Agency should explore the possibility that other proprietary data might be available. Nevertheless, based on the information in Appendix II.E.1 of the Agency document (August 2, 2005), the REJV dataset appears to contain much useful information regarding residential exposure to NMC, and in many respects is more useful than the National Home and Garden Pesticide Use Survey (NHGPUS), which is dated. The Agency also appears to appreciate the limitations of this database and has articulated a plan to consider the impact of these limitations on assessing the cumulative risk of NMC.

It will be almost unavoidable to use information sources other then REJV to supplement the estimation of exposure from residential applications of NMC, considering that the REJV database contains insufficient information for cumulative exposure assessment. The Agency's general principles and approach toward using other necessary data (e.g., information on pesticide use pattern and available formulation, maximum

application rate) are reasonable and represent the best effort under the current situation. The use of "professional judgment" is often subsumed in the selected distributional characteristics, e.g., analytical distributions truncated at the 99th percentile. Sensitivity analyses should be performed to assess the impact of such judgments.

The Panel also supports the Agency's approach in using the REJV data for empirical estimates of pesticide use patterns for residential exposure scenarios. The Panel provides the following suggestions and comments on the residential exposure analysis.

1) Co-occurring residential applications: The REJV data may lack sufficient information to address co-occurring application events. One plausible scenario is a series of NMC applications to trees, ornamentals, and the home garden, all by the same person in one extended event. This scenario might arise "for convenience," when extra tank mix remains after an originally intended single use on, for example, trees. Exposures in such a scenario may not be adequately characterized by a probabilistic approach based on the REJV. It may be reasonable to conduct a separate deterministic analysis to determine the plausible upper end of exposure.

2) Exposure following a professional application: This scenario is not specifically addressed. The residential exposure analysis using REJV data deals only with post- application exposures associated with homeowner applications. Residential exposure after professional application is a realistic possibility, however, even though, as the Agency has indicated, the home presents fewer occasions for professionals to apply NMCs than it does for the residents themselves. Adding the scenario of professional application in the home would modestly increase the probability and frequency of estimated residential exposure.

3) REJV and NHGPUS comparison: The REJV appears to be a superior database--being more recent than the NHGPUS data--and can better address certain exposure scenarios associated with residential use of pesticides. Nonetheless, the proprietary nature of this database is likely to limit its usefulness and presents some difficulties in achieving transparency for risk assessment. Also, as a one-time survey, the REJV will also be outdated in a few years. Moreover, while Calendex and CARES both use REJV, Lifeline uses NHGPUS. The Panel advises the Agency to compare empirical use patterns generated from REJV and from NHGPUS, especially in a manner similar to that which is used by Lifeline. The comparison may add to the support for either database and enhance the future utility of the REJV.

4) Potential new database: It is recognized that there is no CSFII- or PDP-like database for residential exposure as yet. However, two 5-years, multi-million dollar research projects funded by EPA NCER are now collecting longitudinal data on relevant activities including residential pesticide uses. Unless there are policies that specifically bar it from doing so, the Agency is encouraged to communicate with the grant recipients to ensure that the future data will be of a quality suitable for cumulative risk assessment

models.

5) Further use of REJV database: EPA has not explored any means of using the REJV database other than to estimate the percentage of households that use particular pesticides and the frequency of product-specific applications. Current use is limited to roughly 1200 complete records. All other responses in the REJV go unused. A more complex statistical analysis might enable EPA to utilize all the REJV data. For example, the methods of censored data analysis can be employed, treating the incomplete records as time-censored data. This step will require EPA to begin viewing the REJV data less like the CSFII records and more like its data on water residues. To fully utilize the REJV data a model of household usage and frequency of use will be needed. The data should be sufficient for a three part model that involves estimating i) the probability of a product-specific event like lawn pesticide treatment on a given day or week, ii) the distribution of the number of total such events in the household in a year (e.g., number of lawn pesticide treatments) and iii) the distribution of the times between events (e.g., time to next lawn pesticide treatment). All three of these components would need to be region-specific. Furthermore, certain types of residents might not perform certain types of activities (condo and apartment dwellers might not do shrub and lawn applications). Incorporating such considerations into the analysis should not be too difficult and would improve this aspect of the risk assessment by simplifying the process and by helping describe the components in terms of statistical distributions.

RESIDENTIAL QUESTION #2

Uncertainties Associated with the Hand-To-Mouth Assessment

To assess non-dietary ingestion (mg/day), the following four key factors are used in the models:

- Residue Concentration (turf residues, pet fur residues, and residues from hard indoor surfaces)
- Hand to mouth frequency (number of events per hour)
- Surface area of the inserted hand parts (cm2)
- Exposure time (hours/day)

Other factors include both saliva extraction efficiency and wet hand adjustment factor. This exposure estimate is then used along with the Relative Potency Factor (RPF) and Benchmark Dose to estimate risk. In the Preliminary N-methyl carbamate assessment, risk estimates for non-dietary oral exposure result in the lowest Margins of Exposure (MOEs), and would therefore be of greatest concern to the Agency; however, these low MOEs appear to be due in part to the incorporation of micro-activity data into our macro activity models. As a result, the non-dietary ingestion scenarios in the

Preliminary N-methyl carbamate cumulative risk assessment are the least refined.

The residue concentration values are derived from individual residue dissipation or deposition studies which are discussed in the Residential Chapter (Section E) of the Cumulative Risk Assessment document. The exposure durations are taken from the Agency's Exposure Factors Handbook. The hand to mouth frequencies and hand surface areas come from behavior studies relying either on observational data of young children using video tape analysis, trained observers, or parental observers. However, study data that evaluated hand-to-mouth frequency and surface area mouthed is difficult to interpret. Specifically, comparison of study results can be difficult due to differences in study practices and methodologies. For example, there are no standard definitions of mouthing (superficial contact, licking, biting, fraction of hand inserted) and thus the data for these behaviors likely differs among studies as a result of the investigators definitions. In addition, the degree to which ancillary data (such as surface area of hand contacted or inserted, the duration of contact, and the length of videotaping) are collected and reported differ among studies. This makes broad-based and generally-applicable interpretation difficult. Nevertheless, Drs. Zartarian and Xue allowed us to use their preliminary distributional analyses of these children's video data in this assessment. The studies used in the hand to mouth frequency analysis performed by Zartarian and Xue are briefly summarized in a table provided in a memorandum dated August 8, 2005 and provided to the Panel under separate cover.

The distributions of hand-to-mouth frequencies and surface area mouthed used in the Preliminary NMC CRA were based on the analysis performed by Zartarian and Xue (as detailed above). In the aggregate models used in the NMC cumulative assessment, each separate iteration selects a single value for the hand to mouth events variable from a distribution of hand to mouth frequency values. Also, each separate iteration of the model selects a single surface area from a distribution of the fraction of hand mouthed. These values are multiplied by the residues and exposure durations which are similarly selected from a distribution of residue and exposure durations as described above. This relatively simple selection process, however, ignores the numerous complexities and interrelationships involved in this critical behavior pattern. For example, the area of hand that is mouthed during a given event may correlate inversely with the frequency of mouthing events. Specifically, more frequent hand-to-mouth events may be associated with mouthing smaller fractions of the hand. The algorithms used in the NMC CRA however, (as established by the OPP Residential Standard Operating Procedures (SOPs) assume independence between these two parameters. This assumption likely leads to overestimates of exposures when upper percentiles of the hand-to-mouth frequency and area of hand mouthed distributions are combined. In addition, the macroactivity approach used in the NMC CRA aggregate models is based on the following assumptions:

- The mouthing frequency (events per hour), as recorded during the course of observational studies, continue at the same rate for the entire exposure duration selected; in reality, a high-end mouthing frequency recorded over a short time

interval (e.g., one hour) may not be likely to continue at the same intensity over a longer time period (e.g., 6 or 8 hours)

- ❑ The hand is fully replenished with residues from a contaminated surface (e.g., the lawn, pet or hard flooring) between each hand to mouth event

- ❑ The contact frequency and surface area data used in this assessment are taken from observational studies in which all hand contacts were recorded as hand-to-mouth events, regardless of the fraction of hand mouthed. Additionally, no adjustment was made for the duration of time the hand remained in the mouth.

R2a. The methodology used in the NMC CRA in which micro-activity data are used in macro-activity approach likely leads to systematic overestimates of exposure when upper percentiles of mouthing frequency and surface area of hand mouthed are combined. Does the Panel agree that this methodology does indeed overestimate exposure? Can the Panel suggest improvements to this methodology to further refine exposure estimates?

R2b. Does the Panel have suggestions for an alternative approach than the one used to estimate the non-dietary oral exposure pathway in the Preliminary NMC CRA? For example, would the use of a time weighted frequency value based on random hourly draws of hand frequency distributions more accurately estimate hand-to-mouth exposures?

Response

This Panel appreciates the Agency's effort in responding the comments made by the previous SAP meeting in February 2005 regarding the issue of incorporating the non-dietary oral ingestion in the cumulative residential exposure assessment. The Panel is pleased to see the Agency's effort in this development.

The Panel recognized that non-dietary oral ingestion is a highly variable element of the residential exposure assessment, and perhaps in the overall cumulative exposure assessment. It is therefore understandable that the non-dietary oral ingestion is the least refined component in the present version of the NMC cumulative risk assessment. As the Agency indicated in the PNMC documentation, an accurate assessment of non-dietary oral ingestion requires a substantial amount of information for four key factors. These are residue concentration, hand-to-mouth frequency, surface area of the mouthing part, and duration of exposure. At present, complete information on these four factors is lacking or conflicting.

Even more problematic is the analysis of hand-to-mouth activity data in a macro-activity approach using the default assumptions set out in the documentation. The Panel agreed that such an approach will overestimate exposure, as expressed by the MOE.

However, different Panel members have different suggestions to mitigate this problem.

As pointed out by one Panel member, problems can arise when simulation procedures use independent draws from two distributions for variables that are related in reality. This may result in overestimates or underestimates of exposure, depending on the nature of the association between the two variables. The variance of the product of two random variables, say $p=x \cdot y$ where x=mouthing frequency and y=surface area of mouthed hand, is $var(p) \approx y^2 var(x) + x^2 var(y) + 2xy \cdot cov(x,y)$. If the input values for x and y are drawn independently in the model run, the covariance term in this variance expression is treated as zero and the variance of the distribution for the resulting product values is simply equal to the first two terms of this expression. On the other hand, if x and y are positively correlated in reality, the variance of the true p (the target value) includes a positive covariance term. Therefore, the simulated values of p will have less variability than the target distribution and exposures will be underestimated (assuming that greater variability in the target distribution corresponds to greater extremes in the tails of the simulated distribution). The reverse will be true if x and y are negatively correlated (e.g., more frequent mouthing is associated with smaller areas of hand mouthed). In that case, modeling the product of two jointly distributed random variables as the product of independent draws from the x and y distributions will overestimate the variances of the target distribution for the product. In that way, draws from the simulated distribution would yield more extreme values than might be encountered in the real-world process being modeled, or, in other words, overestimate exposure at the high end.

The rules that govern variances of functions of random variables also inform us about the potential impact of employing a single value for an input variable over a protracted time as opposed to refreshing the value through independent draws throughout the exposure window. Without presenting the formal statistical argument here, using a single draw of an input variable for a protracted period of exposure will result in greater variability in exposures (hence greater extremes in values) than a modeling procedure that periodically returns to refresh the value of the input during the window of exposure. For example, fixing the residue on a child's hands for a two hour play period, while simpler to implement, will yield greater variability in the modeled distribution of exposures than a run that updates the residue concentration hourly during the exposure. Almost certainly, the composite simulated distribution will contain more extreme values than the target distribution, which arises from fluctuating exposures. The degree to which variability in the target distribution is overestimated will be inversely related to the autocorrelation of the input variable over time. A high autocorrelation will lead to a small overestimation. Note: fixing the value of an input variable for a time period is equivalent to assuming perfect auto-correlation of its values for subintervals of the larger time window. A sensitivity analysis is the best way to determine whether or not these effects have practical importance for the interpretation and use of the final exposure simulation.

Moving away from statistics, several Panel members suggested collecting more data and qualitative information from the videotapes and transcripts of studies on hand-

to-mouth exposure. A key question is whether the surface area of the hand (or fingers in mouth) co-varies with the frequency of mouthing. Panel members recognized that a major undertaking might be needed to answer this question. However, some preliminary work is justified to see how much additional information can be culled from the existing investigations and data.

One Panel member concluded that the Agency's idea of time weighting might lead to a more realistic assessment of exposure. This individual encouraged the Agency to pursue this approach by segmenting the time of contact into multiple short periods with separately determined frequency of behaviors. Using relatively large blocks of time would simplify an initial analysis along these lines. The impact of the time-weighted approach could then be evaluated before deciding whether to increase temporal resolution and use shorter time segments. The Agency followed a similar approach with its CCA exposure assessment. However, other Panel members noted that the paucity of available data would make it hard to tell which outcome is "better" if time weighting affects the analysis.

Another Panel member recommended excluding hand-to-mouth activity from the residential exposure assessment model until more data becomes available. Meanwhile the Agency should run simulations with a deterministic model to learn whether hand-to-mouth activity contributes appreciably to the overall residential exposure assessment. If the estimated exposure from hand-to-mouth activity does represent a substantial portion of the total residential exposure, the Agency is recommended to make the macro-activity assumptions of the model more realistic. In contrast, if hand-to-mouth exposure indeed represents only a small fraction of the total residential exposure, excluding that component from the assessment is justified for the following reasons:

1. The dermal exposure component takes into account the fraction of pesticide residue that would be ingested if hand-to-mouth activity does occur.
2. Oral exposures from dietary ingestion, and water consumption in certain regions, are a much more important component of the cumulative assessment and deserve more resources and attention as assessment methods are refined.

The Panel member who raised these points argued that current data for assessing non-dietary oral ingestion are insufficient both in quantity and in quality, and are unlikely to be sufficient in the foreseeable future. This individual concluded that, without good quality data to facilitate model development, inclusion of the hand-to-mouth component actually carries additional and unnecessary error and uncertainty forward to the cumulative risk.

RESIDENTIAL QUESTION #3

Distributional Analysis

Assessing residential exposure to pesticides is a complex process that must

consider exposure from a variety of sources via multiple routes. To account for exposure from different sources, the PNMC residential exposure assessment identifies scenarios where significant exposure may occur. Each of these scenarios is defined by a specific type of activity or set of activities that may result in exposure. Generally the relationships between these activities and the resulting exposures are well-defined in that algorithms, equations, and standard operating procedures exist for calculating exposure based on the activity being performed. However the supporting data sets used to estimate exposure for various residential scenarios range from robust (e.g., unit exposure values) to limited or sparse (e.g., lawn sizes, area treated, duration of exposure, and saliva extraction factors). Additionally, information characterizing the extent to which each activity contributes to exposure for a particular scenario does not always exist (e.g., the amount of time spent in home gardens performing activities such as hand weeding versus staking tomatoes or harvesting sweet corn).

In general, the Agency has attempted to fit distributions (as described in Appendix II.E.2 of the NMC CRA) to the exposure measurements for residential activities when supporting information exists to characterize the extent to which the activity contributes to exposure for the residential scenario of interest. However, the Agency has employed uniform distributions to the data sets for which such supporting information does not exist, (e.g. lawn sizes, area treated, duration of exposure, and saliva extraction factors). The Agency has elected to create such distributions when the available data are limited to such an extent that it is uncertain how well they represent national variability. The Agency believes use of uniform distributions to be conservative in estimating potential exposure since uniform distributions tend to overestimate exposure.

R3a. Please comment specifically on the Agency's use of lognormal distributions to estimate residential exposure and the statistical methods and procedures by which the Agency has selected particular distributions (e.g., probability plots and goodness-of-fit statistics).

Response

Probability plots and Shapiro-Wilk tests are as good a method as any for assessing a distribution. The lognormal has a moderately long tail and the results presented here, even though based on a few small samples, suggest that the lognormal is good enough. A lognormal distribution arises naturally when an observed variable is the product of many arbitrarily distributed variables, or a sum of variables on the log scale. The lognormal distribution is a common choice for environmental measurements and these results come as no surprise.

The question asks for a more rigorous evaluation of the goodness-of-fit methodology, however, and the following points need to be made.

- The power of any goodness-of-fit test will be too low in small samples and too high in large samples, so that a goodness-of-fit test never answers the relevant question: whether or not the distribution is good enough for the model to give reasonably accurate predictions.

- When the null hypothesis is true and many tests are done, the p-values for the tests should follow a uniform distribution (5% less than .05, 1% less than .01, etc.). One does not want all values to be "close to 1" as that would indicate that the data are closer to lognormal than they should be. There are a few too many very small p-values among the tests shown in the current document but, on closer inspection, the low p-values are mostly from inhalation data where several points are tied for minimum and lie on a horizontal line at the bottom of the plot: that is, points that are below the limit of detection and reported as 1/2 LOD. These can be ignored in a visual evaluation of the probability plot but will invalidate the Shapiro-Wilk test. The maximum likelihood estimation of the parameters, allowing for censoring, is correct, but the goodness-of-fit tests are not correct as shown.

- Censored data must be compared to a censored lognormal, with the tail below the detection level removed and replaced by a point mass at the LOD. A grouped-data chi-square test could be used but will be less powerful than Shapiro-Wilk. A reasonable and quick adaptation of Shapiro-Wilk for left censored data is accomplished by fitting a linear regression line to the QQ normal plot (with the normal quantiles on the X-axis and the log concentration quantiles on the Y-axis) and omitting censored values while performing the regression. This regression line provides estimates for the distribution mean and variance that are comparable to the censored-data MLE estimates and the R^2 value is the Shapiro-Francia statistic. Testing for normality (or log normality here, since the data have been log-transformed) is accomplished by determining whether the R^2 term is close enough to 1. The critical value depends on the number of censored values and the total number of observations but in general, if the R^2 is not greater than about 0.96, there is evidence that the data are not normal (or, in this case, log normal). Visual inspection of the straight line is often good enough and outlier values are usually very visible.

R3b. Does the Panel agree that the Agency's approach to creating and using of uniform distributions (i.e., ranges of values) for residential scenarios lacking adequate supporting information tends to overestimate exposure? Is the Panel aware of other data sources that may be better suited for assessing residential exposure scenarios of interest? Does the Panel have any suggestions regarding alternative distributions to use for scenarios where supporting exposure information is inadequate? To what extent should sensitivity analyses be used to assess the appropriateness of alternative distributions?

Response

Uniform distributions should never be used as they have no tails and will never generate extreme cases. In these applications, the true distributions are generally skewed to the right. In consequence, models using uniform distributions will understate the upper tails of the exposure distribution and lead to underestimates, not overestimates, of exposure. At a previous SAP meeting the Panel commented extensively on the use of uniform distributions in the context of the SHEDS analysis (Minutes of the meeting of August 30, 2002, No. 2002-06). Some pertinent quotations from that document are reproduced below:

> "The Panel felt that the extensive use of uniform distributions to represent either uncertainty or variability should be discouraged in favor of parametric distributions that do not have such strictly defined limits. Distributions with defined limits should generally be used only in cases where the limits can be firmly based on physical principles. The model should also allow use of Beta, Gamma and Weibull distributions, mixtures of any of the available distributions, and the ability to establish a distribution with a spike of probability at 0. The Beta distribution includes the Uniform as a special case and is more general as the distribution of a proportion. In the technical documentation the user should be cautioned to avoid the Normal distribution for values that are known to be non-negative and positively skewed, particularly where the standard deviation is over half of the mean.
>
> One Panel member expressed reservations about the use of a normal distribution for both the variability and the uncertainty about the mean of the surface-to-hand transfer coefficient; i.e. the surface-to-hand transfer coefficient among children is assumed to follow a Normal distribution and the uncertainty in the mean of that distribution is also described by a normal distribution. The Panelist expressed the belief that this Normal-Normal assumption for the surface-to-hand transfer coefficient could lead to substantial understatement of the uncertainty in this factor. In particular, the model as implemented had the variance of the mean surface-to-hand transfer coefficient less than the variance among children in surface-to-hand transfer factor. Given that the variance in surface-to-hand transfer coefficients is limited by the variability in hand surface area among children, this was considered highly implausible."

During previous SAP reviews of other probabilistic modeling efforts (e.g., CARES, Lifeline), Panel members have commented on the use of uniform distributions. Synopses from these past comments follow:

- Analysts often give the perceived simplicity of the uniform distribution as an important attraction for cases where there are limited empirical data. The uniform distribution, with its defined absolute upper and lower limits,

unfortunately provides an opportunity for analysis to fall into a trap that a particular parameter has zero chance of having values outside the range of a limited available data set. It is completely incorrect in general to assume that the largest and smallest values in a group of 9-30 data points or fewer represents the true minimum and maximum values that the variable can assume.

- Moreover there are few cases where the mechanisms that cause measurements or estimates of exposure-related parameters to vary among people create situations where there is no greater chance of producing a case near the center of a distribution than at its extreme end (as required for the uniform distribution to be correct). Factors that cause exposure to differ from one individual to another tend to interact multiplicatively—leading, when these factors are numerous, to expectations of a lognormal distribution. When one or more categorical factors are likely to have a strong influence on exposure (e.g., wearing short-sleeved vs. long-sleeved shirts) it is desirable to create mixtures of lognormal distributions, weighted by their expected frequency, to represent the influence of those different known cases.

- The uniform distribution is appropriate in cases where (1) it is physically impossible for the parameter to take on values outside the limits and (2) there really is no greater likelihood for values close to the center of the range rather than at either end. For example, there would be no problem in using a uniform distribution to represent the day of the week that a meteor might land. However, as many of the applications in the current model for both variability and uncertainty, the uniform distribution is often selected in cases where there can be no solid assurance that the parameter cannot take on values outside the stated range. In attempting to select a defined absolute range, the analyst is very vulnerable to the psychic trap of "overconfidence". "Overconfidence"— the general underestimation of uncertainty (assigning confidence limits that are too narrow) is one of the best documented phenomena in risk analysis. This applies to both subjective evaluations by experts and non-experts (Tversky and Kahneman, 1974; Alpert and Raiffa, 1982; Lichstenstein and Fischoff, 1977), and to supposedly "objective" numerical calculations by physicists (Shlyakhter and Kammen, 1992).

- Hattis and Burmaster (1994) gave a series of rules and examples of mechanisms that give rise to different distributional forms. Experience and the basic idea that variability is often the result of many factors acting multiplicatively indicates that the lognormal form is most often the best choice for exposure-related data where there is limited information. Both normal and lognormal distributions have just two parameters, and are thus no more "complex" statistically than a uniform distribution (and in that sense, less complex than the three-parameter triangular distribution). Derivation of the parameters of lognormal distributions can be done if a simple range is given together with the number of independent observations that gave rise to that range.

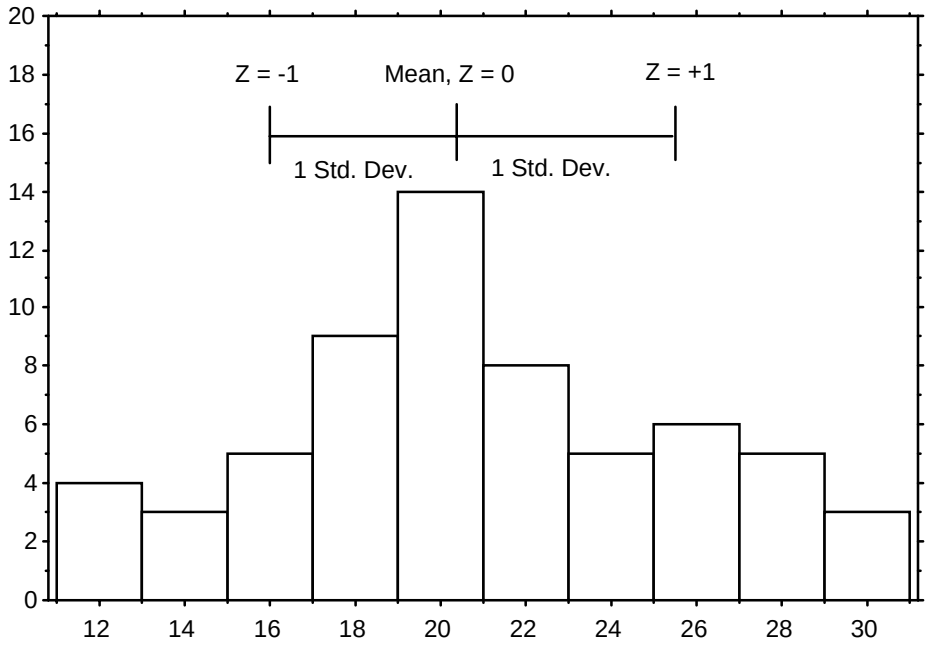
Means and other measures of dispersion, such as a standard deviation, can also be used to estimate the parameters of lognormal distributions.

One example of the use of uniform distributions in the NMC CRA is the breathing rate distribution. In describing the breathing rate data, the document says (p. 128)

> **"Breathing Rates:** The breathing rates used for this assessment are represented by a uniform distribution from 1 to 2 m^3/hour for light to moderate activity. This assumption is based on information from the EFH (USEPA, 1997). This distribution was used to assess exposure for all age groups."

In general, use of uniform distributions for describing inter-individual variability should be discouraged. Here, for comparison, are some breathing rate distributions collected in another context, direct breathing rate measurements from activity survey data of coal miners (Figures 1-3) and a general population of adults and children (Figure 4). It can be seen in the probability plots that these distributions are reasonably described using normal or lognormal distributions. A better distribution for the breathing rates in the NMC CRA might be to combine the assumed mean breathing rate with the dispersion from the observational studies depicted in these figures.

Fig. 1 Distribution of Measured Breathing Rates in Working Coal Miners



Data Source: Jones, C. O., Gauld, S., Hurley, J. F., and Rickmann, A. M. (1981). Personal differences in the breathing patterns and volumes and dust intakes of working

miners. Report to the Commission of the European Communities, Report No. TM/81/11, Environmental Branch, Institute of Occupational Medicine, Roxburgh Place, Edinburgh, Scotland.

Figure 2

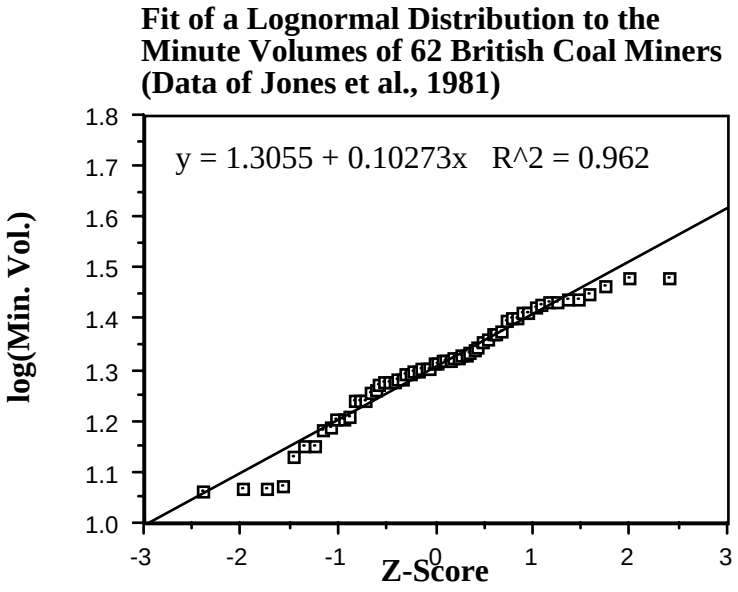
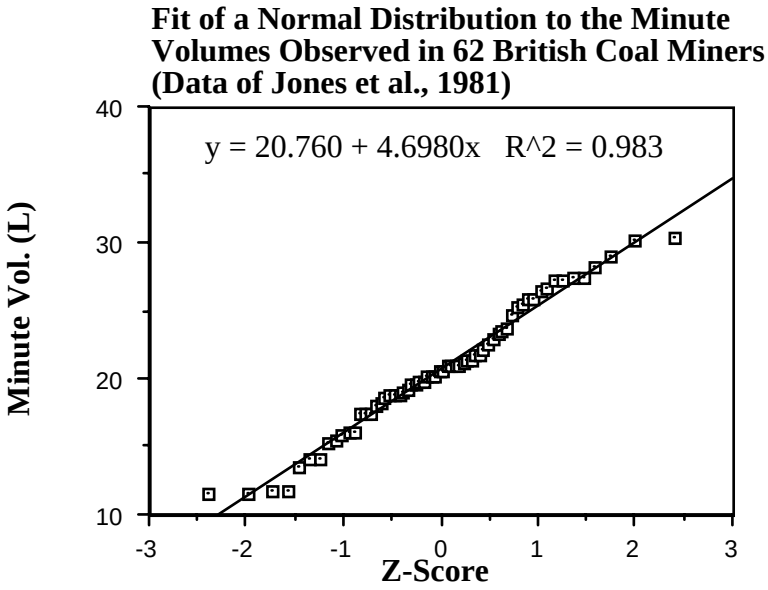
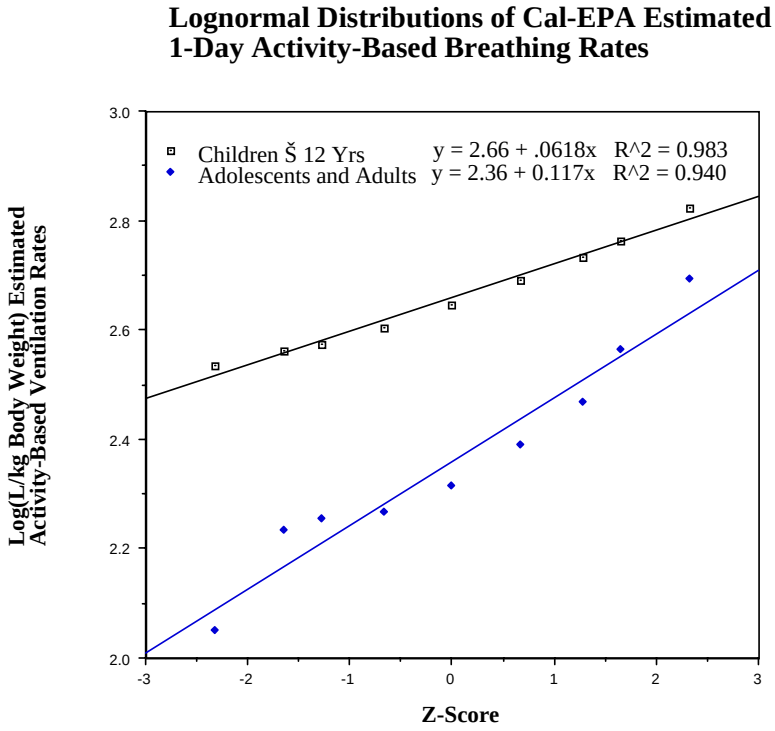


Figure 3



Source: Hattis, D., and Silver, K., "Human Interindividual Variability--a Major Source of Uncertainty in Assessing Risks for Non-Cancer Health Effects," *Risk Analysis*, Vol 14, pp. 421-431, 1994.

Figure 4



In summary, the Panel agreed that it is probably best to standardize on lognormal distributions; if you have enough information to pick the 2 parameters of a uniform, you should be able to pick the 2 parameters of a lognormal. If you restrict attention to 2-parameter positively skewed distributions on the positive axis, the best-known distributions (lognormal, gamma, logistic and Weibull) are quite similar to each other and the choice will not affect the model output significantly. A small sensitivity analysis comparing uniform with lognormal would be interesting.

Defaulting to the lognormal assumes that the value being modeled is positive-valued; if a distribution is symmetric and long-tailed, a shifted and scaled t on low degrees of freedom (3 parameters) could be tried, if you want something more general than the uniform (2 parameters) you could use the beta (4 parameters). If the uniform distribution is used, it will usually be a good idea to set the limits a bit wider than the observed range of values.

Use of uniform distributions as uncertainty distributions on unknown distributional parameters in 2-D Monte Carlo simulations is slightly more acceptable but still to be discouraged.

The Panel was not aware of other data sources better suited for assessing

residential exposure scenarios.

R3c. When the Agency fits distributions to various exposure values, the maximum value entered into the probabilistic models for a particular distribution is usually defined to be an upper percentile value such as the 99th percentile in order to ensure realistic input parameters. Recognizing that the Agency intends to perform sensitivity analyses to evaluate the effects of this truncation, please comment on the Agency's approach of truncating distributions that are input to the probabilistic models. Please comment on any other approaches that the Agency might use to evaluate uncertainties associated with choices about whether and where to truncate distributions.

Response

The Panel agreed that distributions should not be truncated unless there is a strong physical or biological reason to set an upper or lower limit. Truncation may eliminate only 1%, say, of the population, but it may be the most interesting 1%. Under truncation, the means and standard deviations of the distributions will be less than the nominal values and the assumptions of the simulation will not be quite the same as advertised. Truncation rules need to be set out in the documentation. A sensitivity analysis, comparing output with and without truncation, would help answer this question.

INTEGRATION

INTEGRATION QUESTION #1

The cumulative risk assessment guidance describes key principles for conducting these risk assessments. One such principle is the need to consider the time frame of both the exposure (e.g., When does exposure occur? What is the exposure duration?) and of the toxic effect (e.g., What are the time to peak effects and the time to recovery? How quickly is the effect reversed?). EPA's Preliminary Cumulative Risk Assessment for the N-methyl carbamates describes the current limitations in data and software to fully characterize the dynamic nature of exposure, effect, and recovery for this common mechanism group. In order to address these limitations, OPP performed an examination of the exposure patterns for records from the high end of exposure distribution and found that that a large fraction (~70%) of daily records contributing to the upper tail of the food exposure distribution represent single eating occasions. Regarding drinking water and residential/non-occupational exposure, EPA's preliminary assessment provided a characterization of the current availability regarding datasets and models and a description of the impact of these limitations on the risk estimates from specific exposure pathways (i.e., drinking water, residential).

I1a. Please comment on clarity and adequacy of the risk characterization provided in the preliminary cumulative risk assessment. Are there

important aspects with respect to the strengths and weaknesses of the risk characterization other than the ones we identified?

Response

The consensus of the Panel was that the cumulative risk analysis document, including the risk characterization, shows enormous progress since the last SAP review in February of this year. Even so, the next round would benefit from some additional background material to improve its accessibility to readers less familiar with the basic assumptions and features of the analysis. Some Panelists found the oral presentation of the risk characterization clearer and more helpful than the written version in the existing document. Finally, one Panelist suggested that, in light of the specific geographic focus of the high-end risk analysis involving significant ground water exposures, it might be more appropriate to present this material as a descriptive scenario rather than emphasizing the limited geographic derivation of the underlying data.

The next report also would benefit from further thought about the application of the relative potency factor paradigm to a group of AChE inhibitors that inhibit cholinesterase on time scales that are short but somewhat different. The relatively rapid recovery of the inhibition and the consequently short time unit for analysis pose an enormous challenge for risk characterization. With daily possibilities for exposure and inhibition there are 365 (or 366) opportunities per year for an adverse event to occur. Therefore a risk characterization document needs to discuss and explain why readers should focus on such alternatives as (A) the worst day experienced by any individual in a one-year period (B) the entire life stage (0-2 years? 0-20 years?) when there might be unusual developmental susceptibility or (C) through (Z) other plausibly relevant exposure and risk descriptors. These are to some extent risk-management judgments, but the risk characterization should be designed to frame and clarify the information that the risk manager and the public might reasonably consider, using the best technical insights we have into the likely dynamics of causation and the relevant dosimeters for adverse effects.

Other than AChE inhibition itself, the acute and readily apparent health effects from exposure to cholinesterase inhibitors go under the acronym SLUD. For these acute effects, experience indicates that a BMD_{10} for brain cholinesterase inhibition is a conservative (health protective) value. Moreover, because of the short time between cholinesterase inhibition and the manifestation of these signs, it is reasonable to consider that peak cholinesterase inhibition levels are the causally relevant measure of internal dose for modeling the risk of adverse responses in individuals.

However, readily observable high-dose effects are not the only, or even the most important responses of concern for population exposures to cholinesterase inhibitors. As recently reviewed by Slotkin (2004), cholinergic signaling plays a vital role in several phases of neurodevelopment including the migration of the cells that will become mature neurons to the locations in the brain (and, likely, elsewhere) where they are needed; the

formation of connections with other neurons, and the survival of the connections (synapses) through phases where unused connections are pared back and lost. It is possible that even subtle strengthening of the signaling via some cholinergic synapses will lead to survival of some connections in preference to other non-cholinergic pathways, and therefore have subtle long term consequences for function. This theory is based on current understanding of fundamental processes of neurodevelopment. For representative research articles and reviews on this large topic, see Lauder (1985), Whitaker-Azmitia (1991), Hohmann and Berger-Sweeney (1998), Lauder and Schambra (1999), Weiss et al. (1998). We now know that "neurodevelopment" of this type is not restricted to fetal life, but continues well after birth. In fact, there is evidence that synaptic rearrangement, as well as the proliferation and planned death of neurons, continues well into adolescence in rats (Bayer et al., 1982; Bayer, 1983) and also humans (Huttenlocher, 1990).

Long-term health effects may also result from the adaptation of cholinergic signaling systems to cholinesterase inhibition. There are several recent reports of unexpected long-term effects from military and agricultural anticholinesterase agents, in some cases when exposures were insufficient to induce acute cholinergic signs (Baker and Sedgwick, 1996; Kelly et al., 1997; Kassa et al. 2004; Tochigi et al. 2002; Abu-Qare and Abou-Donia, 2002; Yokoyama et al. 1998; Sanchez-Santed et al. 2004; Jamal et al. 2002). Long-term animal studies with military nerve agents have recently led to suggestions of a need to revise LOAELs determined on the basis of short-term experiments (VanHelden et al. 2003; 2004), even though the short-term experiments utilized quite a mild effect (changes in pupil size) as the measure of response. Neuroscientists on the Panel considered that longer-term adaptive responses are unlikely unless cholinesterase inhibition reaches levels that alter synaptic physiology by overriding the margin of safety for cholinergic transmission. The Panel finds, however that the concern for subtle developmental and adaptive effects does warrant further discussion in the risk characterization, to help decision-makers and the public put the Agency's choice of benchmark doses in perspective.

Suspected developmental and adaptive effects might be more directly dependent on a time-weighted integral of cholinesterase inhibition than on peak inhibition levels on specific days. There is a relatively straightforward way that EPA can use information it has already developed to perform an alternative set of exposure assessments based on this "Area Under the Curve" of cholinesterase inhibition dosimeter. That is, the Agency can do a parallel set of exposure analyses using an alternative set of Relative Potency Factors that incorporate both the potency of each carbamate for producing peak inhibition, and the rate at which that inhibition is reversed (at the lowest available doses in the experiments already analyzed by EPA).

In rats, the estimated reversal half-lives for inhibition by carbamates vary widely and the associated 95% confidence limits suggest statistical significance among some pairs (Table 1.B.6). In particular, the reversal half-life of the proposed index chemical, oxamyl (0.75h with tight confidence limits of 0.66h-0.88h) is much shorter than that of

formetanate (4.05h, with confidence limits of 3.02h-5.44h) and methiocarb (2.77h, with confidence limits of 1.91h-4.01h).

Table 1 below illustrates a simple approach to construct an alternative set of Relative Potency Factors (these might be designated "RPF*") using a preliminary index of time-integrated relative potencies. The index would simply consist of the product of the BMD-based RPFs and the individual chemical half-lives for cholinesterase inhibition reversal—renormalized so that, as before, the value for the index chemical is set at 1. Because oxamyl has the shortest half-life for inhibition reversal, this approach tends to increase the relative potency factors for all the other compounds— by over 7-fold in the case of formetanate.

Table 1
Preliminary Calculation of Time-Integrated Relative Potency Factors (RPF*)

	Peak BMD-Based RPF	Brain Inhib Reversal T1/2 (hr)	BMD*T1/2	Renormalized RPF* Relative to Oxamyl
Aldicarb	3.32	1.52	5.05	6.73
Carbaryl (0-10 dose)	0.12	1.83	0.22	0.29
Carbofuran	1.19	2.49	2.96	3.95
Formetanate	1.89	5.4	10.21	13.6
Methiocarb	0.14	2.77	0.39	0.52
Methomyl	0.38	0.8	0.30	0.41
Oxamyl	1	0.75	0.75	1.00
Primicarb	0.02	1.90[a]	0.038a	0.051
Propoxur	0.09	2.69	0.24	0.32
Thiodicarb	0.7	1.90[a]	1.33a	1.77

[a]For compounds with missing values for the inhibition reversal half-life, the geometric mean of the half lives for other compounds (1.90) has been substituted in the calculation.

Some mechanistic reasoning and modeling could lead to modest further modifications to the RPF* calculations when translated into human equivalents. All of the N-methyl carbamates are expected to leave behind the same chemical moiety on AChE. Purely spontaneous chemical reversal of the inhibition therefore is expected to occur at an identical rate across different carbamates in the common mechanism group, and should at least be similar across species from rats to humans. Data to confirm this expectation would not be difficult to obtain.

Given the similar behavior of the N-methyl carbamates at the level of the target enzyme, the differences in their inhibition reversal half-lives are likely to reflect differences in their rates of clearance from the body by excretion and metabolism. To the degree that a particular chemical's brain inhibition reversal does depend on this kind of slow clearance, then the ordinary scaling principles for metabolism-based pharmacokinetics (Travis et al. 1990; Boxenbaum, 1980 and 1982; Reese and Hattis, 1994) suggest that the clearance will be slower in people—according to convention,

roughly in proportion to Body Weight raised to the 0.25 power. Thus, the baseline expectation is that this kind of pharmacokinetic clearance should be about 4 times slower in a 70 kg person than in a 0.3 kg rat: $(70/0.3)^{0.25} = 4$. To translate the rat brain inhibition reversal half lives to human equivalents, therefore, a factor of 4 might be applied to the chemicals whose brain inhibition reversal rate is markedly slower than the in vitro reactivation half-lives that may be observed in experiments now under way. In cases where the estimated brain AChE inhibition reversal half-lives are similar to those measured at the appropriate temperature in vitro, however, no such interspecies projection factor should be applied. If modeling studies indicate that the processes limiting the rate of inhibition reversal can be apportioned between the spontaneous chemical processes and clearance from the body reservoir(s), then of course the interspecies slowing factor should only be applied to the active metabolic clearance process. This circumstance could therefore eventually lead to some differential adjustment of the RPF* across species [and, by extension, across age groups for the very young and the very old whose clearance rates tend to be less than those for young adults (Ginsberg et al. 2002 and 2005). Meanwhile, these considerations suggest that the in vivo inhibition half-lives for some N-methyl carbamates might be long enough to call into question a basic assumption of the proposed cumulative risk assessment for this common mechanism group. In particular, if one applies a 4.1-fold inter-species scaling factor to the 5.4 hr half-time for reversal of brain AChE inhibition in rats, one obtains a predicted half-time of 22 hr in the 70 kg human adult. Such a long half-time would force the risk assessment model to address carryover of inhibition from one day to the next. In considering this issue, the Agency should take into account cases where there is a dose dependency for inhibition reversal half-lives. In these cases, projections should utilize the estimates from the lowest feasible dose rate because it may be most relevant to the expectations for exposure at BMD_{10}.

I1b. Is the Panel aware of additional data which would aid the Agency in its cumulative risk characterization for the N-methyl carbamate pesticides? For example, is the Panel aware of any available data on the timing of water consumption events or can the Panel make any recommendations regarding reasonable assumptions that could be made to help characterize the estimated risk? Are there other sensitivity analyses and further investigations that would be equally or more important than the ones we identified?

Response

In response to this question, and because of the clear importance of the local drinking water pathway to the analysis, the Panel has undertaken some very simple pharmacokinetic modeling on this subject. There is good reason to suspect that the current method of analysis—lumping all daily exposure into a single event—introduces a systematic distortion in the expected effects on peak inhibition levels between dietary exposure (which, for upper percentiles, seem to be mostly traceable to single eating

events per day) vs. drinking water, which clearly occurs in several different events distributed throughout the day. The Panel has asked itself, how might the RPFs for peak brain AChE inhibition be adapted to accommodate a likely scenario for time-dependent water consumption?

The following analysis is based on what the Panel understands to be one plausible pattern of diurnal drinking water consumption that the EPA uses for pharmacokinetic analyses. This pattern consists of three water consumption events (each delivering 25% of daily consumption) at meal times separated by 5-hour intervals, and two between-meal drinking events, each delivering 12.5% of daily consumption. Figure 5 compares the results of this consumption scenario for expected peak inhibition levels in rats following water intake of BMD_{10} amounts of the NMC with the longest inhibition reversal half life (formetanate) and the NMC with the shortest half life (oxamyl). (The Panel has not attempted to incorporate further adjustments to translate results into half-lives for general and special human subpopulations).

Figure 5 shows that this hypothesized pattern of drinking water exposure does indeed lead to quite different expectations for peak daily cholinesterase inhibition for the daily drinking water doses of the two NMCs. The 5.4 hour half life for formetanate leads to appreciable buildup during the day to about 5.9% inhibition, whereas the predicted peak inhibition for oxamyl is only a little more than the 2.5% expected from each mealtime drinking event considered separately. If inhibition did not reverse between events, of course, or if the total daily dose were delivered in one bolus, the expected peak inhibition would be 10%. Table 2 illustrates how this finding can be translated into a simple numerical adjustment to the Relative Potency Factors.

Figure 5

Expected Time Pattern of Brain Acetylcholinesterase Inhibition for a Drinking Water Pattern of Exposure to Formetanate (Inhibition T1/2 = 5.4 Hours) vs Oxamyl (Inhibition T1/2 = 0.75 Hours)

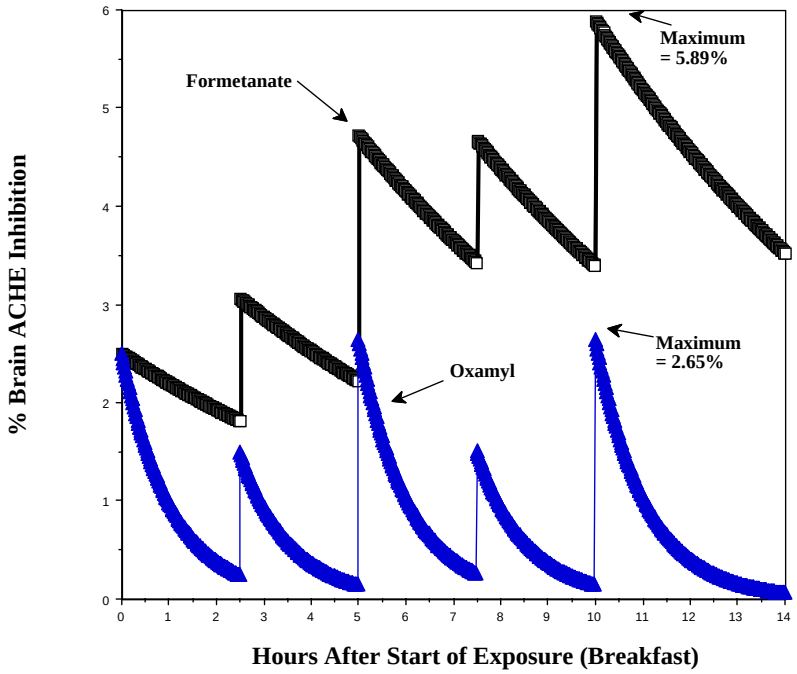


Table 2

Indicated Relative Potencies for Formetanate vs. Oxamyl for Maximal Daily Inhibition for a Drinking Water Pattern of Exposure

	Half Life (hr)	Maximum % Inhibition for BMD_{10} Exposure
Formetanate	5.4	5.89
Oxamyl	0.75	2.65
Ratio	7.2	2.23

In Table 2, the 2.2-fold upward adjustment of the formetanate RPF relative to oxamyl is significant, but less than the full ratio of the two half lives that would be indicated if AUCs were the desired causal dosimeter for a particular type of toxic response. Alternatively, the overall RPF for oxamyl itself in drinking water might be

adjusted downward to .265 of the RPF for oxamyl in the single oral doses expected for most high-percentile dietary exposures.

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

Other comments on the analysis:

The discontinuity that is apparent for some exposure routes between the end of the year vs. the beginning of the year needs to be resolved. It cannot be true that all pet collars are really applied on Jan 1; a more random day needs to be chosen for this and similar modes of exposure, with allowance for cross-year exposures as needed to fully represent realistic patterns for the start of an exposure event vs. the day of actually delivered exposures.

Institutional (e.g., school) and occupational exposures should be incorporated into the assessment.

It will become important to supplement whole- and half-brain data with measurements of causally relevant amounts and durations of AChE inhibition in brain regions that are mechanistically connected to specific developmental and other effects. Brain regions differ in basal levels of AChE, and may therefore differ somewhat in sensitivity to inhibition. Whole-brain and half-brain measurements are considered by some in the field to be the "wave of the past". On the other hand, data on region-specific inhibition of AChE are currently quite limited, and it is recognized that such measurements are associated with much higher variability than those from whole brain. For these reasons there was no Panel consensus in favor of incorporating regional studies at this time.

One Panelist considered that the non-dietary oral exposures were likely to be much more uncertain than other sources of exposure. If the EPA agrees, the next document might discuss the relative confidence of the analysis in the quantification of exposures by various routes, and consequent implications for research and risk management priorities.

Another topic for discussion in the next iteration of the document is the fact that areas of the country were chosen where carbamate residential exposure and use is likely to be higher than that in the rest of the country because of greater pest pressures, among other circumstances. This may suggest a scenario presentation rather than the implication of a full US-South regional analysis. Other suggestions by individual Panelists included subpopulation analyses by ethnic groups, groups with different dietary habits (e.g., vegetarians) and other categorizations of people that might be associated with differences in exposures. Such categorizations might ultimately be helpful in

formulating options for information programs and other risk management efforts.

REFERENCES

HAZARD

Hammond, P., T. Jelacic, S. Padilla, and S. Brimijoin. 1996. Quantitative, video-based histochemistry to measure regional effects of anticholinesterase pesticides in rat brain. Anal. Biochem. 241:82-92.

WATER

(In response to question W1)

Liu G., S. Dai, Y. Qian, and Q. Gan. 2003. Experimental Study on Effect of Anion Surfactant on Degradation Rate of Aldicarb in Soil. *Journal of Environmental Science and Health, Part B: Pesticides, Food Contaminants, and Agricultural Wastes* 38:405-416.

Ou, L.T., P.S.C. Rao, K.S.V. Edvardsson, P.E. Jessup, and A.G. Hornsby. 1988. Aldicarb degradation in sandy soils from different depths. *Pesticide Science* 23 (1):1-12.

Steenhuis, T.S., S. Pacenka, and K.S. Porter. 1987. MOUSE: A Management Model for Evaluating Groundwater Contamination from Diffuse Surface Sources Aided by Computer Graphics. *Appl. Agric. Res.* 2:277-289.

(In response to question W2)

Darnault, C.J.D., T.S. Steenhuis, P. Garnier, Y.-J. Kim, M.B. Jenkins, W.C. Chiorse, P.C. Baveye, and J.Y. Parlange. 2004. Preferential flow and transport of Cryptosporidium parvum oocysts through the vadose zone: Experiments and modeling. *Vadose Zone J.* 3:262-270.

Glass, R.J., G.H. Oosting, and T.S. Steenhuis. 1989. Preferential Solute Transport in Layered Homogeneous Sands as a Consequence of Wetting Front Instability. *J. Hydrol.* 110:87-105.

Jarvis, N.J., M. Stahli, L. Bergstrom, and H. Johnsson. 1994. Simulation of dichlorprop and bentazon leaching in soils of contrasting texture using the MACRO model. *J. Environ. Sci. Health, A* 29(6):1255-1277.

Kim, Y.-J., C.J.G. Darnault, N.O. Bailey, J.-Y. Parlange, and T.S. Steenhuis. 2005. Equation for Describing Solute Transport in Field Soils with Preferential Flow Paths. *Soil Science Society of America J.* 69:291-300.

Kung, K-J.S. 1990. Preferential flow in a sandy vadose zone: 1. field observation. *Geoderma* 46:51-58.

Sadeghi, A.M., A. R. Isensee, and A. Shirmohammadi. 2000. Influence of soil texture and tillage on herbicide transport. *Chemosphere* 41:1327-1332.

Selker, J.S., T.S. Steenhuis, and J.-Y. Parlange. 1996. An engineering approach to fingered vadose pollutant transport. *Geoderma* 70:197-206.

Shipitalo, M.J., W.M. Edwards, and W.A. Dick. 1990. Initial storm effects on macropore transport of surface-applied chemicals in no-till soil. *Soil Science Society of America Journal* 54(6):1530-1536.

Singh, P., R. Kanwar, K.E. Johnsen, and L.R. Ahuja. 1996. Calibration and evaluation of subsurface drainage component of RZWQM V.2.5. *Journal of Environmental Quality* 25(1):56-63.

Steenhuis, T.S., C. Jackson, K.-J.S. Kung, and W.H. Brutsaert. 1985. Measurement of Groundwater Recharge on Eastern Long Island. *J. Hydrol.* 79:145-169.

Steenhuis, T.S. and W.H. van der Molen. 1986. The Thornthwaite-Mather Procedure as a Simple Engineering Method to Predict Recharge. *J. Hydrol.* 84:221-229.

RESIDENTIAL

(In response to question R3b)

Alpert, M. and H. Raiffa. 1982. A progress report on the training of probability assessors. in *Judgment Under Uncertainty, Heuristics and Biases,* D. Kahneman, P. Slovic, and A. Tversky, eds., Cambridge University Press. N. Y. pp. 294-305.

Hattis, D. and D. E. Burmaster. 1994. Assessment of variability and uncertainty distributions for practical risk analyses. *Risk Analysis* 14:713-730.

Lichtenstein S. and B. Fischoff. 1977. Do those who know more also know more about how much they know? *Organizational Behavior and Human Performance* 20:159-183.

Shlyakhter, A. I. and D.M. Kammen. 1992. Sea-level rise or fall?" *Nature* 253:25.

Tversky A. and D. Kahneman. 1974. Judgment under uncertainty: Heuristics and biases. Science 185, 1124-1131, In: *Judgment Under Uncertainty: Heuristics and Biases,* Edited by: D. Kahneman, P. Slovic and A. Tversky. Cambridge University Press. N. Y. 1982 pp. 3-20.

INTEGRATION

Abou-Donia, M.B. 2003. Organophosphorus ester-induced chronic neurotoxicity. *Arch. Environ. Health*. 58(8):484-497.

Abu-Qare, A.W. and M.B. Abou-Donia. 2002. Sarin: health effects, metabolism, and methods of analysis. *Food Chem. Toxicol.* 40(10):1327-1333.

Baker, D.J., and E.M. Sedgwick. 1996. Single fibre electromyographic changes in man after organophosphate exposure. *Human & Experimental Toxicol.* 15:369-375.

Bayer, S.A., 1983. [^{3}H]Thymidine-radiographic studies of neurogenesis in the rat olfactory bulb. *Exp. Brain Res.* 50:329-340.

Bayer, S.A, J.W. Yackel, and P.S. Puri. 1982. Neurons in the rat dentate gyrus granular layer substantially increase during juvenile and adult life. *Science* 216:890-892.

Boxenbaum H. 1980. Interspecies variation in liver weight, hepatic blood flow, and antipyrine intrinsic clearance: extrapolation of data to benzodiazepines and phenytoin. *J. Pharmacokinet. Biopharm.* 8(2):165-176.

Boxenbaum, H. 1982. Interspecies scaling, allometry, physiological time, and the ground plan of pharmacokinetics. *J. Pharmacokinet. Biopharm.* 10: 201-227.

Ginsberg, G., D. Hattis, B. Sonawane, A. Russ, P. Banati, M. Kozlak, S. Smolenski, and R. Goble. 2002. Evaluation of Child/Adult Pharmacokinetic Differences from a Database derived from the Therapeutic Drug Literature. *Toxicological Sciences* 66:185-200.

Ginsberg, G., D. Hattis, A. Russ, and B. Sonawane. "Pharmacokinetic and Pharmacodynamic Factors that can Affect Sensitivity to Neurotoxic Sequelae in the Elderly." Environmental Health Perspectives doi:10.1289/ehp.7568 (available at http://dx.doi.org/) Online 26 May 2005, in press.

Hohmann, C.F. and J. Berger-Sweeney. 1998. Cholinergic regulation of cortical development and plasticity: new twists to an old story. *Perspect. Dev. Neurobiol.* 5:401-425.

Huttenlocher, P.R. 1900. Morphometric study of human cerebral cortex development. *Neuropsychologia* 28:517-527.

Jamal G.A., S. Hansen, and P.O. Julu. 2002. Low level exposures to organophosphorus esters may cause neurotoxicity. *Toxicology* 181-182:23-33.

Kassa J., G. Krejcová, F. Skopec, J. Herink, J. Bajgar, L. Sevelová, M. Tichý, and M.

Pecka. 2004. The influence of sarin on various physiological functions in rats following single or repeated low-level inhalation exposure. *Inhal. Toxicol.* 16(8):517-30.

Kelly, S.S., G.E. de Blaquière, F.M. Williams, and P.G. Blain. 1997. Effects of multiple doses of organophosphates on evoked potentials in mouse diaphragm. *Hum. Exp. Toxicol.* 16(2):72-78.

Lauder, J.M. and U.B. Schambra. 1999. Morphogenetic roles of acetylcholine. *Environ.Health Perspect. 107 Suppl.* 1:65-69.

Rees, D.C. and D. Hattis. "Developing Quantitative Strategies for Animal to Human Extrapolation" Chapter 8 in Principles and Methods of Toxicology, 3rd Edition, A. W. Hayes, ed., Raven Press, New York, 1994, pp. 275-315.

Sánchez-Santed F., F. Cañadas, P. Flores, M. López-Grancha, and D. Cardona. 2004. Long-term functional neurotoxicity of paraoxon and chlorpyrifos: behavioural and pharmacological evidence. *Neurotoxicol. Teratol.* 26(2):305-17.

Tochigi M., T. Umekage, T. Otani, T. Kato, A. Iwanami, N. Asukai, T. Sasaki, and N. Kato. 2002. Serum cholesterol, uric acid and cholinesterase in victims of the Tokyo subway sarin poisoning: a relation with post-traumatic stress disorder. *Neurosci. Res.* 44(3):267-272.

Travis, D.C., R.K. White, and R.C. Ward. 1990. Interspecies extrapolation of pharmacokinetics. *J. Theor. Biol.* 142:285-304.

Weiss, E.R., P. Maness, and J.M. Lauder. 1998. Why do neurotransmitters act like growth factors? *Perspect. Dev. Neuobiol.* 5:323-335.

Whitaker-Azmitia, P.M. 1991. Role of serotonin and other neurotransmitter receptors in brain development: basis for developmental pharmacology. *Pharmacol. Rev.* 43:553-561.

Yokoyama K., S. Araki, K. Murata, M. Nishikitani, T. Okumura, S. Ishimatsu, and N. Takasu. 1998. A preliminary study on delayed vestibulo-cerebellar effects of Tokyo Subway Sarin Poisoning in relation to gender difference: frequency analysis of postural sway. *J. Occup. Environ. Med.* 40(1):17-21.