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

OFFICE OF
PREVENTION, PESTICIDES, AND
TOXIC SUBSTANCES

April 15, 2005

MEMORANDUM

SUBJECT: Transmittal of Meeting Minutes of the FIFRA Scientific Advisory Panel Meeting Held February 15 - 18, 2005 on the N-methyl Carbamate Cumulative Risk Assessment: Pilot Cumulative Analysis

TO: James J. Jones, Director
Office of Pesticide Programs

FROM: Myrta R. Christian, Designated Federal Official
Joseph E. Bailey, Designated Federal Official
FIFRA Scientific Advisory Panel
Office of Science Coordination and Policy

THRU: Larry C. Dorsey, Executive Secretary
FIFRA Scientific Advisory Panel
Office of Science Coordination and Policy

Clifford J. Gabriel, Director
Office of Science Coordination and Policy

Attached, please find the meeting minutes of the FIFRA Scientific Advisory Panel open meeting held in Arlington, Virginia on February 15 - 18, 2005. This report addresses a set of scientific issues being considered by the Environmental Protection Agency pertaining to the N-methyl carbamate cumulative risk assessment: pilot cumulative analysis.

Attachment

cc:

Susan Hazen
Margaret Schneider
Anne Lindsay
Margie Fehrenbach
Janet Andersen
Debbie Edwards
Steven Bradbury
William Diamond
Arnold Layne
Tina Levine
Lois Rossi
Frank Sanders
Richard Keigwin
Randolph Perfetti
William Jordan
Douglas Parsons
Enesta Jones
Vanessa Vu (SAB)
Anna Lowit
David J. Miller
Nelson Thurman
Dirk Young
David Hrdy
Jeff Evans
Steve Nako
Stephanie Padilla
R. Woodrow Setzer
Ginger Moser
Miles Okino
Jerry Blancato
Fred Power
Curtis Dary
Tom Nolan, USGS
OPP Docket

FIFRA Scientific Advisory Panel Members

Stephen M. Roberts, Ph.D. (Chair of the FIFRA SAP)
Janice E. Chambers, Ph.D.
H. Christopher Frey, Ph.D.
Stuart Handwerger, M.D.
Steven G. Heeringa, Ph.D.
Gary Isom, Ph.D.
Kenneth M. Portier, Ph.D.

FQPA Science Review Board Members

John Adgate, Ph.D.
William Brimijoin, Ph.D.
George B. Corcoran, Ph.D.
Lutz Edler, Ph.D.
Bernard Engel, Ph.D.
Scott Ferson, Ph.D.
Lawrence Fischer, Ph.D.
Natalie Freeman, Ph.D.
Gaylia Jean Harry, Ph.D.
Dale Hattis, Ph.D.
James P. Kehrer, Ph.D.
Chensheng Lu, Ph.D.
Peter Macdonald, D. Phil.
David MacIntosh, Ph.D.
Robert Malone, Ph.D.
Christopher Portier, Ph.D.
Nu-May Ruby Reed, Ph.D., D.A.B.T.
Barry Ryan, Ph.D.
Michael Sohn, Ph.D.
Tammo Steenhuis, Ph.D.
Michael D. Wheeler, Ph.D.

SAP Minutes No. 2005-01

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

N-METHYL CARBAMATE PESTICIDE CUMULATIVE RISK ASSESSMENT: PILOT CUMULATIVE ANALYSIS

Session 1: Issues Related to Cumulative Hazard Assessment
Session 2: Physiologically Based Pharmacokinetic/ Pharmacodynamic (PBPK/PD) Modeling for Carbaryl
Session 3: Drinking Water Exposure Analysis
Session 4: N-methyl Carbamate Exposure Assessment: A Pilot Case Study.

FEBRUARY 15 - 18, 2005
FIFRA Scientific Advisory Panel Meeting, held at the Holiday Inn - National Airport, Arlington, Virginia

TABLE OF CONTENTS

N-Methyl Carbamate Cumulative Risk Assessment: Pilot Cumulative Analysis

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 contents of the meeting minutes do 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 or Joseph E. Bailey, SAP Designated Federal Officials, via e-mail at christian.myrta@.epa.gov or bailey.joseph@epa.gov.

In preparing the meeting minutes, the Panels carefully considered all information provided and presented by the Agency, 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.

INTRODUCTION

The FIFRA SAP has completed its review of the set of scientific issues being considered by the Agency pertaining to the N-methyl carbamate pesticide cumulative risk assessment: pilot cumulative analysis. Advance notice of the meeting was published in the *Federal Register* on January 6, 2005. The review was conducted in an open Panel meeting held in Arlington, Virginia, on February 15 - 18, 2005. The meeting included four separate sessions as follows.

- **Session 1 - Issues Related to Cumulative Hazard Assessment**

- **Session 2 - Physiologically Based Pharmacokinetic/ Pharmacodynamic (PBPK/PD) Modeling for Carbaryl**

- **Session 3 - Drinking Water Exposure Analysis**

- **Session 4 - N-methyl Carbamate Exposure Assessment: A Pilot Case Study**

Sessions 1 and 2 were chaired by Dr. Gary Isom and Myrta R. Christian served as the Designated Federal Official. Sessions 3 and 4 were chaired by Dr. Steven Heeringa and Joseph E. Bailey served as Designated Federal Official. A public comment period was held at the beginning of the 4-day meeting on February 15, 2005.

PUBLIC COMMENTERS

Oral statements were presented as follows:

On behalf of the Natural Resources Defense Council:
Jennifer Sass, Ph.D., Natural Resources Defense Council

On behalf of Bayer Crop Science: Abraham Tobias, Ph.D.

SAP Minutes No. 2005-01

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

N-METHYL CARBAMATE PESTICIDE CUMULATIVE RISK ASSESSMENT: PILOT CUMULATIVE ANALYSIS

SESSION 1: ISSUES RELATED TO CUMULATIVE HAZARD ASSESSMENT

FEBRUARY 15, 2005
FIFRA Scientific Advisory Panel Meeting, held at the Holiday Inn - National Airport, Arlington, Virginia

Myrta R. Christian, M.S.
Designated Federal Official
FIFRA Scientific Advisory Panel
Date: April 15, 2005

Gary Isom, Ph.D.
FIFRA SAP Session Chair
FIFRA Scientific Advisory Panel
Date: April 15, 2005

Federal Insecticide, Fungicide, and Rodenticide Act
Scientific Advisory Panel Meeting
February 15, 2005

N-METHYL CARBAMATE PESTICIDE CUMULATIVE RISK ASSESSMENT: PILOT CUMULATIVE ANALYSIS

SESSION 1: ISSUES RELATED TO CUMULATIVE HAZARD ASSESSMENT

PARTICIPANTS

FIFRA SAP Session Chair

Gary Isom, Ph.D., Professor of Toxicology, School of Pharmacy & Pharmacal Sciences, Purdue University, West Lafayette, IN

Designated Federal Official

Myrta R. Christian, M.S., FIFRA Scientific Advisory Panel Staff, Office of Science Coordination and Policy, EPA

FIFRA Scientific Advisory Panel Members

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

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

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

FQPA Science Review Board Members

William Brimijoin, Ph.D., Chair, Pharmacology, Mayo Clinic and Medical School, Rochester, MN

George B. Corcoran, Ph.D., Professor & Chairman, Department of Pharmaceutical Sciences, Eugene Applebaum College of Pharmacy & Health Sciences, Wayne State University, Detroit, MI

Lutz Edler, Ph.D., Head, Biostatistics Unit C060, German Cancer Research Center, Heidelberg, Germany

Scott Ferson, Ph.D., Senior Scientist, Applied Biomathematics, Setauket, NY

Lawrence Fischer, Ph.D., Center for Integrative Toxicology, Michigan State University, East Lansing, MI

Gaylia Jean Harry, Ph.D., Neurotoxicology Group Leader, National Institute of Environmental Health Sciences, Research Triangle Park, NC

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

James P. Kehrer, Ph.D., Director, Center for Molecular & Cellular Toxicology, College of Pharmacy, The University of Texas at Austin, Austin, TX

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

Christopher J. Portier, Ph.D., Director, Environmental Toxicology Program, National Institute of Environmental Health Sciences, Research Triangle Park, NC

Nu-May Ruby Reed, Ph.D., D.A.B.T., Staff Toxicologist, 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

Michael D. Sohn, Ph.D., Scientist, Environmental Energy Technologies Division, Lawrence Berkeley National Laboratory, University of California, Berkeley, CA

Michael D. Wheeler, Ph.D., Assistant Professor, Departments of Pharmacology & Medicine, University of North Carolina, Skipper Bowles Center for Alcohol Studies, Chapel Hill, NC

In Session 1 of this meeting, the FIFRA SAP met to consider and review N-methyl carbamate pesticide cumulative risk assessment: pilot cumulative analysis, issues related to cumulative hazard assessment. EPA acknowledges that there are toxicological characteristics unique to the N-methyl carbamates which need to be considered in a cumulative risk assessment for this group. Specifically, the mechanism of action for this group of pesticides is carbamylation of the acetylcholinesterase (AChE) active site. This chemical change is reversible, allowing for relatively rapid recovery from inhibition. OPP is collaborating with laboratory scientists and statisticians from EPA's National Health and Environmental Effects Research Laboratory (NHEERL) to evaluate biological and empirical aspects of recovery. EPA solicited comment from the SAP on specific issues related to dose-response modeling of AChE data, empirical estimation of time to recovery, and the impact of the laboratory method used to measure AChE inhibition on estimates of relative potency. The agenda for this SAP meeting session involved an introduction, background, and detailed presentations of the issues related to cumulative hazard assessment provided by Dr. Anna Lowit (Health Effects Division, Office of Pesticide Programs, EPA), Dr. Stephanie Padilla, Dr. R. Woodrow Setzer, and Dr. Ginger Moser, (Office of Research and Development, National Health and Environmental Effects Research Laboratory, EPA). Dr. Clifford Gabriel (Director, Office of Science Coordination and Policy, EPA), Mr. Jim Jones (Director, Office of Pesticides Programs, EPA), Mr. Joseph J. Merenda, Jr. (Office of Prevention, Pesticides and Toxic Substances, EPA), and Dr. Tina Levine (Acting Director, Health Effects Division, Office of Pesticide Programs, EPA) offered opening remarks at the meeting. Dr. William Brimijoin and Dr. Gaylia Jean Harry presented comments submitted by Dr. Janice Chambers.

SUMMARY OF PANEL DISCUSSION AND RECOMMENDATIONS

The overall issue posed to the panel for the Day 1 discussions was the suitability of the Agency's near-term plans for quantitative estimation of Relative Potency Factors (RPFs) for use in the carbamate cumulative risk assessment. An initial question was the reliability and usability of the basic cholinesterase inhibition data supplied to the EPA by pesticide registrants. The second generic question called on the Panel to evaluate three aspects of the Agency's plans for quantitative analysis of those data: (a) the basic biological/toxicological assumptions underlying the analyses, (b) the statistical methodology incorporating a simplified pharmacokinetic approach planned for estimating RPFs, and (c) the later potential for use of probabilistic methods to characterize uncertainty distributions related to standard point-estimate uncertainty factors that have traditionally been used in Agency risk analyses for non-cancer effects.

The Panel's principal response to the first issue was based on a qualitative graphical comparison of cholinesterase inhibition observations for several carbamates from an extensive set of Agency-conducted experiments using a modern radiometric method with results from (a) Agency-conducted experiments using an older version of a colorimetric method, and (b) results from Agency files of experiments reported by pesticide registrants using an updated modification of the colorimetric method. Briefly, these comparisons revealed considerable discrepancies between the results of the radiometric assay and the results from the older colorimetric method. However discrepancies were much less apparent, though not completely absent, between the radiometric results and the results of the modified colorimetric procedure that has evidently been used by registrants in recent years. The committee stressed that ultimate conclusions on this issue must be based on a full quantitative analysis of the dose-time-response relationships seen for the radiometric data in relation to the registrant data. If a full statistical analysis reveals anomalies in the comparisons of key parameters for particular compounds, then some adjustment of registrant data for individual compounds may be needed. In the meantime, on the assumption that proper care was taken by the registrants in the conduct of the assays, it appears that the results of the studies submitted for pesticide registration provide a fairly reliable database of AChE inhibition in relation to carbamate exposure. Thus, on balance, the Panel supports EPA's position that the data are suitable to support the cumulative risk assessment.

On the second question, in general, the Panel approves the way the Agency is approaching the empirical analysis of AChE data for N-methyl carbamates. The empirical model being used should be flexible and is likely to fit the data, although the Panel also offered several suggestions on details of the primary modeling approach and modest variants that should be considered by the Agency. Five simplifying assumptions articulated by the Agency were all generally considered reasonable, although there were reservations on some, and the Panel re-emphasizes the need for quantitative statistical analyses of the uncertainties resulting from departures from these assumptions where possible. Finally, the Panel welcomed the openness on the part of EPA staff to explore the potential for using distributional approaches in noncancer risk assessments to the extent possible. The Panel offers several suggestions in the spirit of encouraging the Agency to help develop and use existing and emerging probabilistic techniques. The

Panel cautions, however, that the process of replacing the uncertainty factors with distributions based on empirical data is far from straightforward. In addition to distributional techniques for analysis, this process will involve new articulation of risk management goals, which are, of course, beyond the scope of the Panel's charge to offer advice on technical issues. For example, although technical methods can in theory provide a description of how confident we should be that the population incidence of harm of a given severity level for a given exposure is less than a series of defined values, there has as yet been no articulation of how to interpret such a distributional description in the derivation of an RfD or an "Acceptable Daily Intake" (ADI). What fraction of the population is intended to be "covered" with what confidence by the current system of single-point "uncertainty" factors?

PANEL DELIBERATIONS AND RESPONSE TO QUESTIONS

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

Questions

Issue 1.1: Laboratory method for measuring cholinesterase inhibition

As discussed in the paper (Cumulative Hazard Assessment: Issues for the FIFRA SAP), in toxicology studies performed for pesticide registration, typically, acetylcholinesterase (AChE) inhibition is measured using modified Ellman, spectrophotometric methods (Ellman et al, 1961). Scientists at EPA's National Health and Environmental Effects Research Laboratory (NHEERL) have performed a series of time course and dose-response studies (Appendices 1 and 2) with seven *N*-methyl carbamate pesticides. These studies have compared radiometric and spectrophotometric methods following acute dosing in rats. EPA has provided plots of these studies along with plots of dose-response data from single dosing rat studies submitted for pesticide registration (Appendix 2). Prior to the completion of these studies, there was a concern that studies submitted to EPA for pesticide registration may underestimate relative potency. Specifically, using spectrophotometric methods, recovery of inhibition can occur prior to analysis if the proper precautions are not taken in the laboratory.

Statistical analyses evaluating results of the radiometric data generated by EPA and the spectrophotometric data reported by registration studies have not been performed. However, in general, based on visual observation of these plots (Appendix 2), there appears to be good concordance, particularly at doses at or near 10% inhibition, between the results of studies submitted for registration purposes where spectrophotometric methods were used and the results of studies performed by EPA where radiometric analyses were performed. Nostrandt *et al.* (1993) have previously shown that modified Ellman assay gave answers comparable to the radiometric assay if some special precautions were taken. EPA does not know the exact conditions used in various laboratories performing registration studies. However, it appears that the AChE data provided in the registration studies are of sufficient quality for evaluating relative potency.

Question 1.1

> **Please comment on EPA's observations regarding the results of radiometric studies conducted by NHEERL and the results of studies submitted for pesticide registration.**

Panel Response

The Panel debated whether the comparative studies of seven N-methyl carbamates at NHEERL give clear evidence that recent AChE data provided by the

registrants to EPA are of sufficient quality to support a cumulative risk assessment. The Panel commended the EPA scientists for their well-designed and valuable study and reached a consensus that the results demonstrated "reasonable reliability" in the registrant data. Several unresolved issues were noted, however, and qualifications and reservations were expressed as indicated below.

The Agency presented dose-response data generated by two different methods for a number of carbamates (coded A to H) of varying potency. One set of data was based on radiometric assays involving minimal sample dilution and minimal assay duration. These data were expected to approximate the true levels of AChE inhibition. A second set of data was based on automated spectrophotometric assays, with no attempt to limit sample dilution or assay time. The data from these "conventional Ellman assays" were expected to be more variable and to underestimate true levels of AChE inhibition by an unknown amount. The purpose of this comparison was to gain perspective on the quality of data previously supplied to EPA by the registrants of the same seven carbamates, who carried out their assays with a version of the spectrophotometric method.

The time-course and dose-response plots for the seven selected N-methyl carbamates in Appendix 2 of the EPA document differed between the radiometric and the spectrophotometric methods. The differences did not follow a consistent pattern and therefore did not lend themselves to easy correction (such as a simple multiplicative adjustment in the percentage of inhibition measured by one method to convert to percentage of inhibition measured by the other method.) It would be interesting to search systematically for the factors that could account for the differences observed. Beyond the factors discussed by the Agency (dilution, temperature) there may be others (e.g. experimental design, type of carbamate, other assay details) that could explain at least some of the discrepancies between assays. This information could help in judging the value and limitations of spectrophotometric methods for determining cholinesterase inhibition.

For objective comparison of the two assay approaches, one should probably parameterize the time-response and dose-response data from each method and use derived parameters to describe the differences. The time course of AChE inhibition in both data sets consisted of a sharp drop followed by a recovery to 100% of control activity. One Panelist suggested that the recovery phase could be fitted to a two-parameter Michaelis-Menten type of curve. A one-parameter exponential recovery model, as used in the current proposal, is also an option. Similarly, the dose-response data could be described by a slope parameter obtained with regression techniques. Because low-dose effects are important for risk assessment, one should explore how to calculate best the low-dose slopes of the different dose-inhibition curves. For compounds where shoulder effects appear, e.g. owing to the normalization against control values, data in the shoulder region could be dropped from the regressions or modeled with the same kind of Michaelis-Menten approach used for the organophosphates. That is, the Agency could model the dose-response curves and express potencies as reciprocals of the benchmark doses associated with a predetermined

percentage of inhibition (e.g., the reciprocal of the ED10, or BMD10, at which cholinesterase is inhibited by 10%). Differences between compounds could so be described - possibly quantitatively- more objectively than when viewing and interpreting the plots only.

Without prejudging the outcome of a rigorous methodological/statistical comparison, the Panel discussed the reliability of EPA's radiometric approach to assessing AChE inhibition by carbamates. A point not addressed in the submitted assay protocols was how long samples waited between homogenization and assay. The Panel heard in session that the delay was up to 90 minutes. Such a delay might have allowed substantial reactivation if samples had been highly diluted. However, the Panel was informed that tissues for these assays were homogenized in less than 5 volumes of buffer and were kept at ice temperature. Although the Panel would prefer direct evidence that reactivation did not occur before the assay, it is reasonable to conclude that the NHEERL radiometric data reflect the true levels of carbamate-induced AChE inhibition. This accuracy contrasts with the data from the same sample set obtained with a conventional Ellman procedure. In the NHEERL study, the conventional data were much more variable, especially with red cells, and in most cases they underestimated AChE inhibition in samples exposed to higher levels of carbamate.

Turning to the registrant data, all based on spectrophotometric methods, the Panel saw generally good agreement with the EPA/NHEERL results obtained using the radiometric method. In many cases, the levels of inhibition reported by the registrants were similar to those in NHEERL's radiometric data. This agreement would be surprising if the registrants had used standard automated Ellman assays. In public comment, however, an industry representative stated that the registrants as a group have made a concerted shift to a more rigorous modified method. The Panel recognizes that a modified Ellman procedure will reliably measure AChE inhibition if tissue dilution, assay time, and assay temperature are minimized as described by Nostrandt et al. (1993). One Panel member pointed out that running such assays in a real-time or "kinetic mode" offered the important advantage of being able to detect reactivation as it occurred, which is not possible with the end-point determinations in a radiometric assay. Taking the data at face value, one can conclude that the registrants' results do give a reliable index of cholinesterase inhibition with most of the tested carbamates. Assuming similar quality in the studies on the carbamates not tested by NHEERL, then most if not all of those registrant data should also be useable for risk assessment.

The agreement between the data sets from NHEERL and registrants should not be overstated, however. In the first place, as mentioned above, firm conclusions await a rigorous statistical analysis. Second, the graphs in Appendix 2 are somewhat misleading because, in standardizing for the control, the curves are forced together on the left (low dose region). This squeezing effect can account for the apparent "optical agreement" at low levels of AChE inhibition. Other options that could be explored by a statistician with access to the raw data would be expressing the comparative results in terms of ratios of the observed percentage of inhibition—either in aggregated form for all animals studied,

or on the basis of distributions of individual animal results in a form such as comparative box plots. Of more concern were specific instances in which the NHEERL and registrant data diverged systematically. Sometimes the divergence was conservative for risk assessment; i.e., the registrant data predicted greater inhibition than the NHEERL data, but in other cases the divergence was non-conservative. With compound D, for example, the registrant data underestimated relative potency in the brain by a factor of about two. With compound C the underestimation of potency in red blood cells (RBC) was equally great. The formal analysis should quantify the uncertainty in comparative potency estimates arising from such cases.

In light of the available information, the Panel was open to the possibility that these individual discrepancies merely reflect random effects in, for example, statistical mean control for AChE activity. Variations in the control for AChE activity are likely since individual animals can differ by 50% or more. As one Panel member pointed out, it would be advantageous in the future to collect a series of timed blood samples from the same animal, starting before the toxicant is administered. A modified study design with same-animal controls will reduce variability and increase statistical power. Unfortunately, but obviously, repeated sampling cannot be extended to tissues like brain tissues. While brain AChE remains a very important target site, future efforts should be devoted to develop models for RBC data as well, considering AChE activities in RBC are significantly correlated with activities in brain tissues as presented by the Agency's scientists. The concern about using brain AChE alone was raised by several SAP panel members in previous meetings, and for the purpose of validating the PBPK/PD models and cumulative risk assessments, an accessible specimen sample, such as red blood cells, is essential. In this case brain sampling from humans is not feasible. Additionally, issues of statistical power come to the fore in attempts to determine low levels of AChE inhibition. Previous SAP meetings have emphasized that precise determinations of BMD10, critical for estimates of relative potency, require accurate measures of inhibition at or below the level of 10%. Using present methods, this goal is unachievable unless group sizes are very large, even with a time-series study design. Novel approaches to address this problem are worth seeking.

Returning to the question of why some registrant data appeared to underestimate the levels of AChE inhibition, the Panel could not exclude the possibility that certain studies suffered from shortcomings in experimental procedure. In order to evaluate those concerns it is important to review the exact conditions used, and what Good Laboratory Practice (GLP) quality standards were implemented in the various laboratories. The standard operating procedures of the registrants should be made available to the Agency together with the original data if they are to be considered for the risk assessment. One Panel member suggested that a Round Robin test using known activities of AChE in human or animal tissue samples should be adapted so the variations between assays or between labs can be quantified within a pre-determined percentage. Similar activities are being undertaken in the State of California. The Panel also considered the possibility of compound-specific effects on enzyme activity under different assay conditions. Data assessing these possibilities would be of valuable. Ultimate conclusions should be based

on a full quantitative analysis of the absorption rate constant, recovery rate constant and AUC of the dose-response curves. Of these, the parameter most likely to vary widely across laboratories/animals is the absorption rate constant. However the recovery rates and the AUCs should agree. If these metrics show anomalies for particular chemicals after a full statistical analysis, then some adjustment of registrant data for individual compounds may be needed. In the meantime, on the assumption that proper care *was* taken by the registrants, it appears that the results of the studies submitted for pesticide registration provide a fairly reliable database of AChE inhibition. Thus, on balance, the Panel supports EPA's position that the data are suitable to support the cumulative risk assessment.

Issue 1.2: Empirical modeling of AChE Data

Part A. Benchmark dose calculations:

In the EPA's cumulative risk assessment of the organophosphorus pesticides (OPs), a decreasing exponential model was used to develop benchmark dose estimates. The FIFRA SAP previously endorsed this approach (FIFRA SAP, 2001 & 2002). EPA plans to use again a decreasing exponential model in its benchmark dose estimates for the N-methyl carbamate pesticides, with the addition of a component to model the time course of AChE inhibition. This model was provided in the Eqs. 1 – 4 and the associated text (See Cumulative Hazard Assessment: Issues for the FIFRA SAP and Appendices 3-4).

Question 1.2a

Please comment on the appropriateness of using the model provided in Equations 1 – 4 to calculate benchmark dose estimates based on cholinesterase inhibition for the N-methyl carbamate pesticides.

Panel Response

In general, the Panel approves of the way the EPA is approaching the empirical analysis of AChE data for N-methyl carbamates. The empirical model being used should be flexible and is likely to fit the data. The multiplicative model $f(t,d) = A[1-g(d)h(t)]$, is a straightforward empirical modeling approach that has the advantage of being both practical and transparent. However, when proceeding with this modeling, the Agency should be aware that the multiplication of the dose function with the time function could yield a dose-time-response surface that could, in some occasions, lack sufficient flexibility for fitting all of the carbamate data sets (see e.g. Figure 1 in Appendix 3). The inclusion of a term combining time t and dose d may need to be considered.

On the dose scale, the model is a bit over-parameterized and it will prove difficult to find an optimal solution. Constraints on the model parameters and/or Bayesian methods could reduce this problem. However, the SAP feels a more complicated model

is probably not justified.

There could be greater emphasis on statistical hypothesis testing (e.g. does the ED10 change from one strain to the next; across sexes). In addition, because this is an empirical model, goodness-of-fit needs to be carefully assessed when using this to make risk projections. The Agency is encouraged to continue exploration of the Aikake Information Criterion (AIC) for comparative assessment of the quality of model fits to the data.

The time component of the dose-response function is the classic solution to a simple PK process and adequate as a first look at the time course effects on dose-response. The dismissal of the human data for direct estimation and its projected use as a "validation" of an allometric scaling of the rat model seems premature (especially in light of the limited analysis used by the EPA to review the agreement between the two methods for evaluating AChE). The Panel encourages the Agency to either directly estimate parameters using the human data or try to fit the human and rat data simultaneously with the species extrapolation built into the modeling.

The use of the Benchmark Dose (BMD) concept and terminology for the cumulative risk assessment of carbamates needs clarification. Otherwise it could be confused with the BMD approach used in standard risk assessment (RA) of chemicals. There, the BMD is used for the definition of a starting point (point of departure) for the low dose extrapolation and the definition of risk specific doses. In contrast, in the present context of cumulative RA of carbamates, the BMD approach has been implemented to derive relative potency factors (RPFs). In addition, this BMD method does not incorporate a process to estimate a BMD lower confidence level (BMDL) which is usually calculated in standard RA. Possible later Monte Carlo uncertainty or sensitivity analyses might benefit from a series of upper and lower confidence limits on the estimates of mean RPFs.

If estimates of the relative potency factors are to be obtained for individual animals, the solution of averaged parameter values will not be the same as the average solution. In addition to the uncertainty distributions for the central estimate RPFs suggested in the previous paragraph, the Agency may need to generate solutions for random sets of animals to look at a variability distribution of relative potency factors. The RPFs will be calculated as ratios of two BMDs, say D_R / $D_{R\text{-index}}$, where the denominator will be the BMD of the index chemical, $D_{R\text{-index}}$ and D_R the BMD of the compound to be considered for an inclusion in the cumulative RA. From a statistical point of view, the RPF is then a ratio of two quantities where each was estimated from dose-response data and possesses therefore statistical variability. Consequently, the variability of the RPF can be calculated and is a function of the statistical variability of D_R and $D_{R\text{-index}}$. In order to calculate the standard error of the RPF =D_R / $D_{R\text{-index}}$ according to a well known statistical method, the appropriate standard error estimates--s.e.(D_R) and s.e.($D_{R\text{-index}}$)--are required.

A source of uncertainty in this BMD approach stems from the choice of R=10% inhibition as the critical effect size. When inspecting the figures of Appendix 2 one realizes that the slope of the dose-response curve is very shallow as one approaches 10% inhibition. The shallow slope translates into a large variability in estimates of the dose that causes this level of inhibition. Therefore, one may question whether the choice of R=10% is optimal or whether it should be replaced or compared to other choices of the critical effect size. Furthermore, since the BMD for the cumulative RA is derived using a dose-time-response function f(t, d) the D_R is also a function of time, better denoted $D_R(t)$. When using the simple PK model, the RPF is also a function of time t after exposure better denoted $RPF(t)= D_R (t)/ D_{R\text{-}index}(t)$. When representing the response as a surface over the dose-time coordinate system, the BMD is similar to an isobole, a parameter well known in combination experiments. RPFs are then simply ratios of isoboles from different experiments and should be presented as such. If there is stability over time, then that can be noted and used.

The multiplicative model f(t,d) uses the time of the peak effect, which is usually not known, as a model parameter to be estimated from the dose-time data. Since measurements near the time of the peak effect are the exception rather than the rule in experiments like the present one, it is often difficult to obtain precise estimates of that time. This difficulty reduces the interpretability of what is otherwise an important model parameter in the simple PK model. The data presented by the Agency came from experiments that measured AChE inhibition no earlier than 15 minutes after exposure. The Agency should attempt to obtain reliable measurements of AChE inhibition at earlier time points. If such data cannot be obtained, it may be possible to use a simpler model for the time component that does not require a biphasic curve but instead uses an instantaneous drop at exposure.

On the other hand, one Panelist expressed a preference for the Agency to adapt the simple pharmacokinetic model in a direction that is more in line with a mechanistic interpretation of the original equation for the time course of cholinesterase inhibition. The current bi-exponential equation implies that all of the inhibition that will ultimately occur is available immediately at the outset of the AChE regeneration process. This is not strictly correct because none of the AChE enzyme molecules can undergo regeneration before they have actually been inhibited. In the view of this Panelist, a more mechanistically faithful model/equation would be a solution to the following differential equation: $d[\text{Inhibition}]/dT = k_1[\text{unabsorbed}] – k_2[\text{Inhibition}]$, where [unabsorbed] is the maximum possible inhibition if all the carbamate were delivered to the site of action in the brain instantaneously.

This equation does not appear to have a closed-form solution, but it can be solved numerically using appropriate software. It will avoid the need in the present equation for the statistical constraint that k_1 is arbitrarily set to be less than k_2 (which will likely be violated for dermal exposure experiments).

A concerted effort is needed to find data on intrinsic clearance rate for any of these chemicals since there may be compound specific changes that could affect the regeneration rate. The Agency needs additional guidance and could possibly use a simple biochemical assay to get the theoretical recovery rates.

Finally, the SAP was pleased that the Agency chose to perform the calculations using the R software package. The equations were nicely coded and commented. Transparency in EPA's model development and application will continue to improve the public's confidence in the EPA's work.

Part B. Simple pharmacokinetic model

As discussed in the background document prepared for the FIFRA SAP, EPA is committed to improving methodologies and approaches for conducting cumulative risk assessments. To this end, EPA has begun development of a simple, pharmacokinetic (PK) based approach for incorporating recovery of cholinesterase inhibition in risk estimates. The simple PK model is more sophisticated than conventional relative potency approaches but less data-intensive than physiologically-based pharmacokinetic/ pharmacodynamic approaches and thus provides a pragmatic method for considering PK and/or mechanistic information in risk estimates. There are still, however, limitations to the application of this approach for the N-methyl carbamate cumulative risk assessment—namely, the capability for cumulative exposure models to output distributions of exposure (in mg/kg or similar units) to individuals. Given this limitation, EPA continues to pursue practical methods for improving risk assessment methods. It is unclear at this time the degree to which this simple PK approach may be used in the cumulative risk assessment for the N-methyl carbamate pesticides. However, as EPA continues to work towards improving its risk assessment methods, EPA is requesting comment from the FIFRA SAP regarding aspects of the development and application of this simple PK approach.

Question 1.2.b

Please comment on the simplifying assumptions used in the simple PK approach to predicting cholinesterase inhibition. Please include in your comments whether these assumptions tend to underestimate or overestimate potential risk. These assumptions are:

- **The inhibitor is cleared quickly from the target tissue, so that recovery time mostly depends upon the rate of decarbamylation of AChE.**
- **Competition among multiple inhibitors for AChE or clearance pathways is quantitatively insignificant.**
- **Inhibitors do not modify the affinity of AChE for other inhibitors (e.g., by binding to a site on the AChE molecule that has allosteric effects), or such effects are quantitatively insignificant.**
- **It is appropriate to ignore resynthesis of new AChE molecules in the time-frame of interest (1 – 6 hours).**
- **The model for effects in humans can be calibrated by scaling parameters of models fit to rodent data.**

Panel Response

The consensus of the Panel is that at low dose levels the assumptions listed below for the current model are probable valid. There are, however, limited actual data to address or support these assumptions for carbamate pesticides. As with any modeling exercise, the accuracy of the model is limited by the quality of the original data. Thus, the Agency is encouraged to examine all available data sets including the human data comparison mentioned in the document, EPA studies on mixtures, and data provided by industry both in and outside of the registration packages for individual pesticides. When evaluating these and other assumptions, the Agency is encouraged to consider how the model would need to be adjusted to account for any assumptions. The adaptations needed to alter or remove some assumptions might not appreciably change the fundamental model structure or implementing software, while accounting for other assumptions may require a new model or software.

Assumption 1. The inhibitor is cleared quickly from the target tissue, so that recovery time mostly depends upon the rate of decarbamylation of AChE.

The assumption that recovery of AChE activity mainly reflects the rate of de-carbamylation is sound, given the half-life values of current members of the cumulative assessment group and the rate of exchange between fat and blood that can be inferred

from simple pharmacokinetic information.[1] This assumption would be true at low concentrations. However, with one compound, inhibition lasted twice as long as with most of the others. The intrinsic rate of decarbamylation with this compound might be atypically slow, but that would be surprising since all of the tested carbamates should generate identical enzyme adducts. Therefore, assumption one may not be universally true. In other words, the rate of recovery of enzyme activity in rats treated with certain N-methyl carbamates might be driven by extrinsic factors, other than intrinsic rates of decarbamylation. Should new carbamate pesticides with much higher degrees of lipophilicity and longer half-life values become available, and appear to a significant extent in the environment, the validity of this assumption should be revisited. Some reevaluation also might be in order if elimination rates from human fat are much slower than the observed regeneration half lives for acetylcholinesterase activity in blood. That may be the case because, as indicated by the human model parameters in ERDEM, the ratio of blood flow to fat volume is much lower in humans than in rats.

Assumption 2. Competition among multiple inhibitors for AChE or clearance pathways is quantitatively insignificant.

The assumption that competition among multiple inhibitors for AChE or for clearance pathways is quantitatively insignificant should hold across much of the anticipated human exposure range. However, this simplifying assumption may not hold

[1] For example, the ERDEM pharmacokinetic model incorporates an estimated fat/blood partition coefficient of 17.1. For rats, the blood flow incorporated into this model is 0.51 liters/minute and the tissue volume is approximately .015 Liters for a 0.25 kg rat. Given this, the rate constant for elimination of carbaryl from fat is expected to be [blood flow /(tissue volume * fat//blood partition coefficient = .51/(.015*17.1) = 2.0/hour. This rate constant directly implies a half life in fat of [ln(2)/2.0 = 0.35 hours], or about 21 minutes. Thus, for the rat experiments, release from the fat seems unlikely to appreciably distort observations of the half life of about 1.7 hours attributed to enzyme regeneration. This would continue to be true even if the actual elimination half life for rat were about twice as long (41 minutes) as implied by an alternative rat tissue/blood partition coefficient estimation model suggested by one Panel member (Ginsberg, G. L., Pepelko, W. E., Goble, R. L., and Hattis, D. B. "Comparison of Contact Site Cancer Potency Across Dose Routes: Case Study with Epichlorohydrin," Risk Analysis Vol. 16, pp. 667-681, 1996; Walker, K., Hattis, D., Russ, A., and Ginsberg, G. "Physiologically-Based Toxicokinetic Modeling for Acrylamide—Risk Implications of Polymorphisms and Developmental Changes in Selected Metabolic Enzymes," Report from the George Perkins Marsh Institute, Clark University, and the Connecticut Department of Public Health to the U. S. Environmental Protection Agency under Cooperative Agreement #827195-0, December 2004; Ginsberg, G. L., Goble, R. L., and Hattis, D. B. "Slope Factor Comparison Across Dose Routes: Case Study with Epichlorohydrin," Report to the U.S. Environmental Protection Agency by TRC Environmental Corporation, April, 1994. Spreadsheets incorporating the alternative assumptions for estimating partition coefficients in rats and humans will be supplied to EPA on request.

under conditions of relatively high exposure to two or more carbamate pesticides, where competition for the AChE protein and for clearance pathways is more likely. Limited data with carbaryl and malathion indicate that malathion decreased the rate constant of absorption and beta-phase elimination of radiolabeled carbaryl when both were administered orally at 10 mg/kg to rat (Waldron and Abdel-Rahman, 1986). This study suggests that at higher exposures, competition for the target as well as for elimination pathways would be expected.

Assumption 3. Inhibitors do not modify the affinity of AChE for other inhibitors (e.g., by binding to a site on the AChE molecule that has allosteric effects), or such effects are quantitatively insignificant.

The assumption that inhibitors do not modify the affinity of AChE for other inhibitors is sound, based on existing knowledge. This is again especially true at low concentrations. The potential interaction of carbamates should be acknowledged, and accounted for when such interaction significantly influences risk analysis. Data to be obtained from the ongoing EPA studies of the combined 7 compounds should provide additional information on possible interactions. However, the current design of these mixture studies may not be sufficient to fully resolve critical questions of effect additivity. The inclusion of neural cell cultures as a complementary model system to assess specific interactions should be considered by the Agency. The validity of this assumption should remain under surveillance for new carbamates that may interact by allosteric means with AChE. If allosteric interactions are discovered, their effect on enzyme sensitivity will have to be defined before one can determine whether the operating assumption will tend to overestimate or underestimate risk.

Assumption 4. It is appropriate to ignore resynthesis of new AChE molecules in the time-frame of interest (1 – 6 hours).

Simplification of any model is desirable especially in reducing the number of parameters for which limited or no data are available. AChE protein is normally produced in excess with the majority being degraded. The half-life of AChE synthesis in rodents is approximate 7 days, providing a net turnover of approximately 10%/day or 0.5%/hr. This turnover leads to the expectation that the maximum impact of new enzyme synthesis on the overall rate of reversal of AChE inhibition should be approximately 3%. The relative impact of resynthesis of new AChE molecules is greater at higher levels of inhibition and becomes less important as activity returns to normal levels. Re-synthesis of new AChE molecules is therefore unlikely to contribute much to the process of recovery during the first 6 hours.

Assumption 5. The model for effects in humans can be calibrated by scaling parameters of models fit to rodent data.

Body weight$^{3/4}$ scaling is the usual assumption for model extrapolation of bulk metabolic capacity among species, in parallel with observations of the scaling of

metabolic rates. There are no obvious data to indicate that the model for effects in humans should not be initially calibrated by scaling parameters of the models fit to rodent data. If these parameters are identified as key contributing factors by sensitivity analysis the Agency should provide some guidance for targeted efforts to test the scaling assumptions for key parameters with limited human studies.

As indicated earlier in the response to Question 1.1, ultimate conclusions should be based on a full quantitative analysis of the absorption rate constant, recovery rate constant and AUCs/dose. Of these parameters, the most likely to vary substantially across laboratories/animals is the absorption rate constant. In contrast, the recovery rates and the AUCs per dose should match from experiment to experiment. If these parameters show discrepancies for particular chemicals, given a full statistical analysis, some adjustment of registrant data for individual compounds should be considered.

Question 1.2c

> **EPA historically has utilized (default) uncertainty factors for interspecies and intraspecies extrapolation. EPA's issue paper (and related appendices 3-4) suggests that application of the simple, PK approach to estimation of risk provides an opportunity to consider probabilistic methods in uncertainty analysis for cumulative hazard assessment. Please comment on biological and quantitative factors which may be important for consideration in the event probabilistic methods were to be used to perform uncertainty analysis in cumulative hazard assessment.**

Panel Response

This question reflects a welcome openness on the part of the EPA staff to explore the potential for using distributional approaches in noncancer risk assessments to the extent possible. The presenters exhibited a commendable dedication to achieve congressionally mandated public health protection goals with the aid of the best inputs of technical understanding that can be mustered in a reasonable time. The Panel offers the suggestions below in the spirit of encouraging the Agency to help develop and use emerging probabilistic techniques. However, the Panel is not able to offer a straightforward incremental path to utilize the insights gleaned from those techniques in the current system of single-point uncertainty factors that has developed over the decades since the original 1954 paper of Lehman and Fitzhugh (Lehman and Fitzhugh, 1954). In part, this is because there has only been a beginning to public discussion of how to separate the risk assessment and risk management elements embedded in the current RfD procedure by specifying quantitative objectives that should be met by RfDs in terms of variability and uncertainty (Hattis et al., 2002). That is, for exposure at an RfD level,

what degree of confidence should there be that the population incidence of some specified degree of biological effect is below a particular value?

The discussion below will first cover one Panelist's perspective on some issues in the potential use of probabilistic techniques to help inform choices of the values of specific uncertainty factors (those for interspecies projection and human inter-individual variability, as supplemented with the additional FQPA factor). Then there will be some more general comments on opportunities and pitfalls in the use of probabilistic techniques in the analysis of currently available data for the carbamate cholinesterase inhibition agents.

Interspecies Factor:

The goal of this factor is to convert the external doses associated with a sensitive but still "adverse" toxicological endpoint in a sensitive species of animals to doses that would be expected to cause the same "benchmark" incidence and severity of effects in humans. Distributional analyses of empirical data on interspecies projection are potentially useful in the following way: By assembling data bases of putatively analogous cases for relatively well studied chemicals, the "uncertainty" in interspecies projections for a chemical for which there are inadequate direct test information are estimated by observing the variability among a set of analogous chemicals with better information. There is, of course, some need for judgment in the choice of a particular set of chemicals that can be considered "analogous" cases for a specific untested chemical.

The most recent distributional analyses relevant to this factor (Hattis et al., 2002) for multi-dose exposures are based on data for putatively analogous indices of systemic subchronic toxicity for 61 anti-cancer agents in animals and people assembled by Price et al., 2002. [The endpoints used as benchmarks for toxicity in animals and people were not identical but were those historically used for prediction of human potencies from animal data by the National Cancer Institute—usually LD10 levels were used for rodents and compared with "maximum tolerated doses"[2] inferred from human studies of the anti-cancer agents. For dogs and monkeys, the benchmarks were typically TDL (Toxic Dose Low) values.[3]] The results of this analysis, in brief, are that where data from single species are used to project the potency of the same chemical in humans, the geometric mean of the estimates are close to what would be expected for a body weight$^{3/4}$

[2] "Maximum Tolerated Doses" are operationally defined as the dose level at which none of six or one of six patients experience dose limiting toxicity with the next higher dose level having two or more patients experiencing dose limiting toxicity [Storer B.E (1989), Design and analysis of phase I clinical trials. *Biometrics.* 45(3), 25-37; and Edler L (2001) Overview of Phase I Trials. In Handbook of Statistics in Clinical Oncology (ed. J Crowley) Marcel Dekker, 1-34.]

[3] This is defined as the lowest dose to produce pathological alterations in hematological, chemical, clinical, or morphological parameters, the doubling of which produces no lethality.

projection. There is, however, a considerable spread in the relationship between the human toxic potencies calculated from this scaling assumption and the potencies estimated from human data. For example, for the 18 chemicals for which there are rat and human potency estimates, the geometric mean human potency for systemic toxicity is about 90% of the value projected from the rat data; however the 95th percentile of the distribution corresponds to more than four times the value projected from the a body weight$^{3/4}$ scaling rule.

A similar distributional analysis is possible for acute toxicity using a database of thousands of LD50 observations in different species of animals assembled by Rhomberg and Wolff (1998). Unfortunately, there are only four chemicals in this data base where estimates of human LD50s are provided. However, analyses have been made of the distributions of interspecies differences for the diverse species represented in the collection. On this basis, the central tendency of interspecies projections using these acute toxicity data is close to dose/body weight^{1} rather than the body weight$^{3/4}$ scaling rule observed for the subchronic data discussed above. The preliminary explanation for the difference in the central tendency scaling rules for subchronic versus acute toxicity is that the key causally relevant internal dosimeter for systemic subchronic toxicity tends to be an AUC (integrated Area Under the Curve of concentration X time) for the toxicant, for which energy-dependent elimination rates are important. By contrast, it is thought that the key pharmacokinetic determinant of acute toxicity following a single dose of a typical toxicant is an internal maximum concentration of the toxicant, which is likely to be more strongly influenced by simple dilution of the administered dose in an appropriate distribution volume--which tends to scale across species in direct proportion to body weight. This comparison has not, however, been done within mode-of-action groups. It might be informative to select the anticholinesterase agents from the existing Rhomberg and Wolff (1998) data base and evaluate the interspecies scaling of that subset. The full Rhomberg and Wolff (1998) data base is available from the first author.

In the context of the analysis of carbamate toxicity, for the chemicals where doses causing comparable degrees of cholinesterase inhibition can be determined, it would be useful to examine the distribution of human/animal toxic potencies for whatever degree of inhibition is selected as the “benchmark” response.

Human Inter-individual Variability Factor

It is important to begin this discussion by mentioning the distinction between “dose-effect relationships” (where the dependent variable is a continuous response, as in a cholinesterase level) and “dose-response relationships” (where the dependent variable is the fraction of individuals who show a quantal, or binary, response such as death or a cleft palate.) In the case of a quantal response that is caused by threshold mechanism, a “benchmark dose” (BMD) is interpretable as the fraction of individual animals (or people, if human data are used) who have thresholds for response that are lower than the defined “benchmark” response frequency (often 10%). Therefore, for a quantal effect, the BMD contains some element of inter-individual variability, at least within the group

of animals (or humans) that was directly studied.

By contrast, where the biological response variable studied is a continuous parameter, as in the case of a cholinesterase level, the BMD is usually[4] defined in terms of a change (such as a 10% reduction) in the group mean of the parameter under study. The chosen extent of change in the continuous variable is sometimes referred to as the "benchmark response" (BMR). In this case, there is no component of individual variability in the BMD—essentially the BMD is an estimate of the dose that causes the designated percentage change in the continuous variable (BMR) in the typical or average member of the experimental animal group tested. Therefore, for a continuous variable like a particular degree of acetylcholinesterase inhibition, the purpose of the human inter-individual variability factor should be seen as the multiple of dose needed to convert the external doses associated with a particular percentage of inhibition (or other associated biological effects) in a typical or average person, to the dose that can be expected with high confidence to produce a corresponding cholinesterase inhibition (or other associated effects) in an acceptably small proportion of exposed humans. Stated in this way, questions are raised as to how small is an "acceptably small proportion of exposed humans" and how high the confidence level should be that the incidence of cholinesterase inhibition at the BMR level is less than the incidence deemed "acceptable." These are policy determinations that are beyond the scope of the technical advice requested from the Panel, although it can be noted that there are preliminary proposals for operational answers to these questions in the existing literature (Hattis et al., 2002).

With this background it also should be noted that inter-individual variability in the dose producing a given amount of cholinesterase inhibition in people does not cover all of the pharmacodynamic variability. In the taxonomy of human variability proposed in the past work of one Panelist (Hattis et al., 1999), pharmacodynamic inter-individual variability has two components:

- Differences among people in the amount of the active agent at the active site needed to produce a particular change in a physiological parameter (e.g. cholinesterase inhibition), and

[4] Alternatives to this have been suggested in some cases. For example some have advocated defining an "effect" for a continuous parameter in terms of the change in the number of individuals who are some number of standard deviations (or log standard deviations for a lognormal distribution) away from the group mean based on the distributional statistics of an unexposed control population. However this "quantalization" approach discards some of the inherent statistical strength of a continuous biological response parameter, and, unless there is some basis for choosing the number of standard deviations used to define abnormality, appears no less arbitrary than picking some percentage change in the group mean level for defining a benchmark response.

- Differences among people in the amount of change in the physiological parameter needed to cause an adverse response (e.g., an acute change in cholinesterase function needed to cause short term symptoms, or developmental effects during possible "windows of vulnerability", or a longer term adaptation needed to produce possible adaptive synaptic changes that could mediate longer term adverse alterations in neural or neuromuscular signaling as has recently been reported for some organophosphate agents) (van Helden et al., 2004; Delgado et al., 2004; Sánchez-Santed et al., 2004; Kassa et al., 2001a; Kassa et al., 2001b; Kassa et al., 2001c). Subtle developmental effects are theoretically possible, for example, an early life marginal strengthening of signaling for some cholinergic pathways relative to other non-cholinergic pathways. This is because, during some sensitive times, initially formed synaptic connections may be lost if there is few signaling traffic along them. One Panelist noted that identifying sensitive window(s) for this kind of effect may be complex because different portions of the human nervous system develop (and perhaps undergo this synaptic "pruning") at different times/ages (Johnston, 1995; Lichtman et al., 2000; Lowel and Singer, 1992; Walsh and Lichtman, 2003).

Because the types of variability mentioned in the second bullet would not be captured even if there were perfect measurements of the human inter-individual variability in changes in cholinesterase activities as a function of dose, the Agency should be cautious in its consideration for possible changes in the normal 10-fold human inter-individual variability factor for anticholinesterase agents in the light of human clinical studies of cholinesterase inhibition from carbamates that have been conducted by or for the pesticide registrants. The variability indicated by those human studies should be analyzed and compared with variability information for other general toxicants, but possible changes in the uncertainty factor for inter-individual variability should be limited by the fact that the full pharmacodynamic pathway to functional response is not captured in cholinesterase inhibition measurements standing by themselves. The FQPA 10-fold factor can be thought of as a policy-based recognition of the possibility of developmental effects mentioned in the second bullet above.

Other General Comments

Panelists expressed general support for exploration of probabilistic approaches that would depart from the routine application of multiple factors of 10 for the various uncertainties in interspecies projection and the extent of human inter-individual variability. However, there was skepticism that the results of probabilistic uncertainty analysis could, in the near term, be used to completely replace the traditional interspecies uncertainty factors without both PK and PD data for validation. Furthermore, while there was optimism that pharmacokinetic scaling for interspecies projection is likely to be approachable, the Panel had reservations on the current capability to treat the PD portion of the pathway, especially with the relatively sparse data specific to carbamate pharmacodynamics. For the inter-individual variability factor, the Panel expressed a

need for explicit consideration of the possible effects of human genetic differences leading to differences in absorption or metabolic elimination. One Panelist expressed a strong preference that most of the analytical resources available for this effort be devoted to statistical analysis of the cholinesterase inhibition dose-response data, with no more than 10% of the available analytical resources devoted to a possible probabilistic analysis.

The Panel expressed its support of the suggestion made by the public presenter (Dr. Sass) to fully document and expose to independent verification the various models used for analysis of the cholinesterase inhibition information, as called for in the EPA's Council for Regulatory Environmental Modeling (CREM) guidelines (http://epa.gov/osp/crem.htm).

Another Panelist, with some explicit support from other members, made four distinct methodological points for EPA to consider in framing its probabilistic analyses:

- First, uncertainties other than those arising from simple sampling errors should be included. Conventional uncertainty statistics such as standard errors are based solely on observable fluctuations in the available data. However empirical experience has shown that even in such mathematically sophisticated fields as physics, as improved experimental techniques allow better measurements, there is a strong tendency for the updated measurements for such parameters as fundamental particle constants to be found to fall outside of the bounds of previously stated statistical confidence limits more often than would be expected by chance. The likely generic explanation for this is that there are generally unsuspected sources of systematic error (miscalibration of instruments; unrepresentativeness of statistical samples in surveys, etc.) that mean that existing standard errors should generally be expanded beyond those conventionally calculated from internal differences within data sets. Techniques have been suggested to do this (Shlyakhter, 1994; Hattis and Burmaster, 1994), but these have not yet been widely adopted.

- Second, the issue of dependence should be thought through for any probabilistic analysis. Dependency issues are of two kinds. First, one parameter may depend on another because of real causal mechanisms that connect them. For example, when fat is absorbed into the blood following a fatty meal, this will be likely to both increase the fat/blood partition coefficient for a volatile lipophilic chemical in a physiologically based pharmacokinetic (PBPK) model, leading to somewhat greater short term absorption from an inhalation exposure, and, in parallel, decrease the tissue/blood partition coefficients, and hence the rate of elimination of the chemical from a variety of lipid-containing model compartments (fat, brain) with venous blood exiting the organs. A second distinct source of dependency is induced correlations in estimation errors for different parameters in the same regression or other statistical modeling analysis. For example, in the simple pharmacokinetic model presented it is quite likely that the simultaneous

estimation of cholinesterase inhibition half life and "absorption" half life" induces cross-correlations in estimates of these parameters. Appropriate use of the calculated estimation errors for these parameters in a probabilistic uncertainty analysis requires direct or indirect consideration of the dependencies of the error distributions for one of these parameters on the others.

- Third, although transformations can be convenient ways to enforce some constraints on parameter values in an estimation procedure (as was mentioned in the discussion of the simple pharmacokinetic model), such transformations may not always be ideal for deriving uncertainty distributions.

- Finally, although it is prudent to use the guidance mentioned in the simple pharmacokinetic model justification to keep the model simple, there may be some merit in exploring even simpler models for sensitivity analysis (for example, eliminating the "absorption" half life).

A further concern that was mentioned by one Panelist (but was shared by other Panelist) was for more analysis of errors caused by model uncertainty or misspecification. What uncertainty arises from possible errors in the mathematical forms chosen for representing resynthesis or reactivation of inhibited acetylcholinesterase, or specifying the relationships between cholinesterase inhibition kinetics in rats versus people? Although the quantitative information for carbaryl and other carbamates are likely to be among the better data sets, they should not be expected to yield very precise estimates of the values of model parameters. This reinforces the need for careful thought, sensitivity and uncertainty analyses.

REFERENCES

Delgado E., McConnell R., Miranda J., Keifer M., Lundberg I., Partanen T., Wesseling C. (2004). Central nervous system effects of acute organophosphate poisoning in a two-year follow-up. Scand J Work Environ Health. 30(5):362-70.

Edler L (2001). Overview of Phase I Trials. In Handbook of Statistics in Clinical Oncology (ed. J Crowley) Marcel Dekker, 1-34.

Ginsberg, G. L., Goble, R. L., and Hattis, D. B. (1994). Slope Factor Comparison Across Dose Routes: Case Study with Epichlorohydrin. Report to the U.S. Environmental Protection Agency by TRC Environmental Corporation.

Ginsberg, G. L., Pepelko, W. E., Goble, R. L., and Hattis, D. B. (1996). Comparison of Contact Site Cancer Potency Across Dose Routes: Case Study with Epichlorohydrin. Risk Analysis Vol. 16:667-681.

Hattis, D. and Burmaster, D. E. (1994). Assessment of Variability and Uncertainty Distributions for Practical Risk Analyses. Risk Analysis 14:713-730.

Hattis, D., Banati, P., Goble, R., and Burmaster, D. (1999). Human Inter-individual Variability in Parameters Related to Health Risks. Risk Analysis 19:705-720.

Hattis, D., Baird, S., and Goble, R. (2002). A “straw man” proposal for a quantitative definition of the RfD. Drug and Chemical Toxicology 25:403-436.

Hattis, D., Walker, K., Russ, A., and Goble, R. (2004). Role of Dosimetric Scaling and Species Extrapolation in Evaluating Risks Across Life Stages. I. Dosimetric Considerations in Childhood. Draft Report to the U.S. Environmental Protection Agency Under RFQ No. DC-03-00009, October, 2004; also see website http://www2.clarku.edu/faculty/dhattis

Jager, T., Vermeire T. G., Rikken M. G., van der Poel P. (2001). Opportunities for a probabilistic risk assessment of chemicals in the European Union. Chemosphere 43(2):257-64.

Johnston, M. V. (1995). Neurotransmitters and vulnerability of the developing brain. Brain Dev 17(5):301-6.

Kassa, J., Koupilova, M., Vachek, J. 2001a. Long-term effects of low-level sarin inhalation exposure on the spatial memory of rats in a T-maze. Acta Medica (Hradec Kralove). 44(3):93-96.

Kassa, J., Koupilova, M., Vachek, J. (2001b). The influence of low-level sarin inhalation

exposure on spatial memory in rats. Pharmacol Biochem Behav. 70(1):175-179.

Kassa, J., Koupilova, M., Herink, J., Vachek, J. (2001c). The long-term influence of low-level sarin exposure on behavioral and neurophysiological functions in rats. Acta Medica (Hradec Kralove). 44(1):21-27.

Lehman, A. J. and O. G. Fitzhugh (1954). 100-fold margin of safety. Assoc. Food Drug Off. U.S. Q. Bull. 18:33-35.

Lichtman, J. W. and Colman, H. (2000). Neuron. Synapse elimination and indelible memory 25(2):269-78.

Lowel, S. and Singer, W., (1992). Selection of intrinsic horizontal connections in the visual cortex by correlated neuronal activity. Science 255:209-212.

Nostrandt, A.C., Duncan, J.A., Padilla, S. (1993). A modified spectrophotometric method appropriate for measuring cholinesterase activity in tissues from carbaryl-treated animals. Fundam. Appl. Toxicol. 21:196-203.

Pieters, M. N., Kramer, H. J., Slob, W. (1998). Evaluation of the uncertainty factor for subchronic-to-chronic extrapolation: statistical analysis of toxicity data. Regul Toxicol Pharmacol. 27(2):108-11.

Pieters, M. N., Freijer, J., Baars, B. J., Fiolet, D. C., van Klaveren, J., Slob, W. (2002). Risk assessment of deoxynivalenol in food: concentration limits, exposure and effects. Adv Exp Med Biol. 504:235-48

Price, P. S., Swartout, J. C., Schmidt, C., and Keenan, R. E. (2002). Characterizing inter-species uncertainty using data from studies of anti-neoplastic agents in animals and humans. Human and Ecological Risk Assessment in preparation, --paper and data available on the web at http://www2.clarku.edu/faculty/dhattis.

Rhomberg, L. R. and Wolff, S. K. (1998). Empirical scaling of single oral lethal doses across mammalian species based on a large database. Risk Anal. 18(6):741-753.

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

Shlyakhter, A. I. (1994). An improved framework for uncertainty analysis: Accounting for unsuspected errors. Risk Analysis 14:441-447.

Slob, W., Pieters, M. N. (1998). A probabilistic approach for deriving acceptable human intake limits and human health risks from toxicological studies: general framework. Risk

Anal. 18(6):787-98.

Storer, B. E. (1989). Design and analysis of phase I clinical trials. Biometrics 45(3):25-37.

van Helden, H. P., Vanwersch, R. A., Kuijpers, W.C., Trap, H. C., Philippens, I. H., Benschop, H. P. (2004). Low levels of sarin affect the EEG in marmoset monkeys: a pilot study. J Appl Toxicol. 24(6):475-83.

Vermeire, T., Stevenson, H., Peiters, M. N., Rennen, M., Slob, W., Hakkert, B. C. (1999). Assessment factors for human health risk assessment: a discussion paper. Crit Rev Toxicol. 29(5):439-90. Review.

Waldron Lechner, D., Abdel-Rahman, M. S. (1986). Kinetics of carbaryl and malathion in combination in the rat. J Toxicol Environ Health. 18(2):241-56.

Walker, K., Hattis, D., Russ, A., and Ginsberg, G. (2004). Physiologically-Based Toxicokinetic Modeling for Acrylamide—Risk Implications of Polymorphisms and Developmental Changes in Selected Metabolic Enzymes. Report from the George Perkins Marsh Institute, Clark University, and the Connecticut Department of Public Health to the U. S. Environmental Protection Agency under Cooperative Agreement #827195-0, December 2004

Walsh, M. K., Lichtman, J. W. (2003). In vivo time-lapse imaging of synaptic takeover associated with naturally occurring synapse elimination. Neuron. 37(1):67-73.

SAP Minutes No. 2005-01

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

N-METHYL CARBAMATE PESTICIDE CUMULATIVE RISK ASSESSMENT: PILOT CUMULATIVE ANALYSIS

SESSION 2: PHYSIOLOGICALLY BASED PHARMACOKINETIC/ PHARMACODYNAMIC (PBPK/PD) MODELING FOR CARBARYL

FEBRUARY 16, 2005
FIFRA Scientific Advisory Panel Meeting, held at the Holiday Inn - National Airport, Arlington, Virginia

Myrta R. Christian, M.S.
Designated Federal Official
FIFRA Scientific Advisory Panel
Date: April 15, 2005

Gary Isom, Ph.D.
FIFRA SAP Session Chair
FIFRA Scientific Advisory Panel
Date: April 15, 2005

Federal Insecticide, Fungicide, and Rodenticide Act
Scientific Advisory Panel Meeting
February 16, 2005

N-METHYL CARBAMATE PESTICIDE CUMULATIVE RISK ASSESSMENT: PILOT CUMULATIVE ANALYSIS

SESSION 2: PHYSIOLOGICALLY BASED PHARMACOKINETIC/ PHARMACODYNAMIC (PBPK/PD) MODELING FOR CARBARYL

PARTICIPANTS

FIFRA SAP Session Chair

Gary Isom, Ph.D., Professor of Toxicology, School of Pharmacy & Pharmacal Sciences, Purdue University, West Lafayette, IN

Designated Federal Official

Myrta R. Christian, M.S., FIFRA Scientific Advisory Panel Staff, Office of Science Coordination and Policy, EPA

FIFRA Scientific Advisory Panel Members

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

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

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

FQPA Science Review Board Members

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

William Brimijoin, Ph.D., Chair, Pharmacology, Mayo Clinic and Medical School, Rochester, MN

George B. Corcoran, Ph.D., Professor & Chairman, Department of Pharmaceutical Sciences, Eugene Applebaum College of Pharmacy & Health Sciences, Wayne State University, Detroit, MI

Lutz Edler, Ph.D., Head, Biostatistics Unit C060, German Cancer Research Center, Heidelberg, Germany

Scott Ferson, Ph.D., Senior Scientist, Applied Biomathematics, Setauket, NY

Lawrence Fischer, Ph.D., Center for Integrative Toxicology, Michigan State University, East Lansing, MI

Gaylia Jean Harry, Ph.D., Neurotoxicology Group Leader, National Institute of Environmental Health Sciences, Research Triangle Park, NC

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

James P. Kehrer, Ph.D., Director, Center for Molecular & Cellular Toxicology, College of Pharmacy, The University of Texas at Austin, Austin, TX

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

David MacIntosh, Sc.D., Senior Associate, Environmental Health & Engineering, Inc., Newton, MA

Christopher J. Portier, Ph.D., Director, Environmental Toxicology Program, National Institute of Environmental Health Sciences, Research Triangle Park, NC

Nu-May Ruby Reed, Ph.D., D.A.B.T., Staff Toxicologist, 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

Michael D. Sohn, Ph.D., Scientist, Environmental Energy Technologies Division, Lawrence Berkeley National Laboratory, University of California, Berkeley, CA

Michael D. Wheeler, Ph.D., Assistant Professor, Departments of Pharmacology &

Medicine, University of North Carolina, Skipper Bowles Center for Alcohol Studies, Chapel Hill, NC

INTRODUCTION

In Session 2 of this meeting, the FIFRA SAP met to consider and review N-methyl carbamate pesticide cumulative risk assessment: pilot cumulative analysis, physiologically based pharmacokinetic/pharmacodynamic (PBPK/PD) modeling for carbaryl. OPP is collaborating with scientists from EPA's National Exposure Research Laboratory (NERL) to develop a PBPK/PD model for carbaryl within the Exposure Related Dose Estimating Model (ERDEM) Platform (Blancato et al., 2002; Okino et al. 2004). The carbaryl model will form the basic structure of a generalized model for the N-methyl carbamates. A Quantitative Structure Activity Relationship (QSAR) database of physicochemical descriptors and provisional PK and PD parameter values has been assembled for selected N-methyl carbamates. The completeness and representativeness of the QSAR database will influence the application of the PBPK/PD model for use in the cumulative risk assessment of the N-methyl carbamates. EPA solicited comment on specific aspects of the appropriate use of ERDEM for this Risk Assessment. The agenda for this SAP meeting involved an introduction, background, and detailed presentations of the issues related to PBPK/PD modeling for carbaryl provided by Dr. Anna Lowit (Health Effects Division, Office of Pesticide Programs, EPA), Dr. Miles Okino, Dr. Jerry Blancato, Mr. Fred Power, and Dr. Curtis Dary (Office of Research and Development, National Exposure Research Laboratory, EPA).

SUMMARY OF PANEL DISCUSSION AND RECOMMENDATIONS

The Panel commends and supports the Agency's effort to develop a PBPK/PD model for carbaryl within the Exposure Related Dose Estimating model (ERDEM) platform and encourages its continuing improvement as the basic structure applicable for N-methyl carbamates (NMC).

Comments were made regarding the coverage in the Agency background document. Specifically, the given information is insufficient for a comprehensive review regarding the completeness and availability of database for model development. Greater transparency and details also are needed for tracking model building decisions and calibration.

Comments and suggestions were made on several areas of model construction, parameter estimation, and model fitting. The Panel expressed concerns about the complexity of the model that appears to go beyond what the available data can support. Several suggestions for its simplification are given. On the other hand, there are areas that may be inadequately modeled. These include considering the inactivation of carbaryl when it is released from binding to cholinesterase, and the bioavailability of carbaryl through dermal uptake. If necessary, it may be possible to place greater emphasis on the lower dose range pertinent to human exposures when calibrating model fit through evaluating the discrepancies between model predictions and experimental data. Thoughtful considerations should be given to the decision to focus the model fit on the blood:brain partition coefficient. A set of fat:blood partition coefficients are provided for initiating a sensitivity analysis. With respect to the PD portion of the model, the general lack of PD data is noted.

The Agency's initial effort is focused on group average parameters. Future efforts can explore and assess parameter range and variability. Modeling data from individual animals may reveal a unique set of coefficients for model parameters. When aggregated data are applied to the model, the pertinence of prediction at the 99.9th percentile warrants careful consideration. It is important to consider the model's utility when inputting biomonitoring data collected in humans. Instead of using only the portion of data collected in Missouri, biomonitoring data collected in California also should be used to characterize human exposures. Finally, the Panel provides detailed suggestions on statistical evaluation, specifically for determining model complexity and parameter estimation, assessing model's goodness-of-fit, and for sensitivity and uncertainty analyses.

PBPK/PD modeling opens up a variety of possible endpoints for risk assessment. The Panel provides perspectives for each of the five endpoints proposed by the Agency in the context of duration and level of exposure, toxicity implication, and the expression of risk. The use of more than one endpoint is encouraged because it can enhance the understanding of how risk is defined and facilitates risk communication.

It is too early in the model development to determine its application in cumulative risk assessment for the NMC. The Agency is encouraged to continue its development. General considerations for the future development of the ERDEM model include ensuring that the sizable model output is organized with options for ease in data and file management and export, and broadening the model application by adding the capability to model endpoints at target fetal tissues.

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

Questions

Issue 2.1: Data Requirements for PBPK/PD Models

The document "Assessment of carbaryl exposure following turf application using a physiologically based pharmacokinetic model" describes the application of a carbaryl specific physiologically based pharmacokinetic/pharmacodynamic (PBPK/PD) model to a case study of exposure for selected exposure scenarios involving broadcast applications of a liquid formulation of carbaryl to turf. A PBPK/PD model was developed based on available laboratory studies, and applied to post-application exposure scenarios.

The PBPK/PD model provided a structure to evaluate the pharmacokinetic (PK) and pharmacodynamic (PD) data available for carbaryl. This case study was instructive in discovering data gaps, such as blood:brain partition coefficient values, isolated metabolism rates, and identification of specific metabolites. Data published in the open literature were generally incomplete for the purposes of PBPK/PD modeling, where simultaneous tissue concentration, excretion and effect data are ideal for model evaluation. Registrant data provided some constraints for PBPK/PD parameter values, but uncertainty remains in those values due to the dependence on fitting *in vivo* data and structure-activity methods (Poulin and Theil 2000; J Pharm Sci 89:16-35).

Question 2.1a

Please comment on the completeness of the data used to develop the PBPK/PD model.

Panel Response

The documents provided to the Panel do not contain a comprehensive review of the data available for potential use in model calibration, and so it is difficult to respond to this question fairly. Future documentation of the model and its derivation should include such potentially usable information. Table 1 in the background document is far from sufficient for this purpose. This is because (1) it evidently omits some data sets that the authors decided not to use, (in some cases for good reasons--e.g. because they were based on experiments with non-ring labeled carbaryl) and (2) because the observations themselves are not presented in tabular form. Presentation of the useable data in tabular form, with available information on observed variation and measurement errors in the data in an appendix, facilitates independent assessment of the correspondence between observations and model predictions.

Question 2.1b

Please comment on the way the data sets were used to estimate and constrain parameter values.

Panel Response

In general, the Panel is supportive of the ERDEM effort but considers the system to be in a relatively early stage of development. This is particularly evident in the application to a complex modeling problem such as that presented by the many metabolic/inactivation pathways and multiple physiological compartments of the current PBPK/PD models for carbaryl in rats, adult humans, and children of various ages. The potential power of the model is that it can project across dose levels, species, and age groups. However, there need to be caveats alerting the reader to the limitations and uncertainties in the model as currently developed and calibrated. The Agency indicated that the model will continue to be improved as more data become available, and this is encouraged by the Panel. Such improvements increase the likelihood that this model can be successfully extended to other carbamates.

This question elicited a wide variety of comments on different aspects of model development, calibration, and testing. The responses below are organized into four major headings:

- Model documentation and disclosure of the history of model development and calibration with observational data

- Fundamental model structure

- General conception and execution of the "fitting" problem. This includes
  - The problem of "overparameterization" and the balance of considerations in basing estimates of various model parameters on fitting results within the model versus exogenous information (such as organ volumes and blood flows)
  - The need for more formal statistical criteria for model fitting, instead of the current relatively informal procedure
  - The issue of model "validation"
  - The issue of group average versus individual values of model parameters

- Evaluation of the current ERDEM model fits and implications of these early results with respect to
  - Which observations should be emphasized in model calibration?
  - Anomalies in the current fitting, and implications for adapting model inputs and refitting

Model Documentation and Development History

The existing report provides only the most basic outline of the processes for using the data to calibrate the model. There is no presentation of what fitting techniques were applied, which parameters were fixed or fitted, or what data were used to evaluate the fits. Several Panelists also indicated that additional documentation of the modelers' decisions in response to inadequacies in first-generation model fits is desirable. The oral presentation indicated a relatively informal process of adjusting the brain/blood partition coefficient. Some Panelists suggested that the EPA should undertake a more detailed and statistically formal effort to resolve the uncertainties in the calibration for the rat model before the projection is seriously used for estimating human delivered doses and inferring brain cholinesterase inhibition from the few urinary metabolite data available for human children.

Fundamental Model Structure

One Panel member made an observation that has significant implications for the structure of the carbaryl model. This was that carbaryl binding to the active site of acetylcholinesterase (AChE) is, essentially, a metabolic destruction step for the carbaryl molecule. Even if the cholinesterase enzyme molecule itself is regenerated by hydrolysis of the methylcarbamoyl moiety bound to the active-site serine, the original carbaryl is not regenerated. Simple calculations by this Panelist suggested that this might not be the primary mode of metabolism of carbaryl, but it could make a significant contribution to overall metabolism. If the contribution is appreciable then the model will need to be adapted. Instead of locating all metabolism of carbaryl in the liver (following a convention in PBPK modeling that may often be an imperfect reflection of reality) the model should provide for catalyzed hydrolysis of carbaryl in all compartments where AChE and butyrylcholinesterase enzyme activities are known to be present--particularly the brain, blood, and liver.

General Conception and Execution of the "Fitting" Problem

a) The issue of "overparameterization"

There is a well-known adage that represents the deceptively high quality of fits where there are few data and many parameters: "with four parameters I can fit an elephant; with five I can make it wave." Actually, a formal paper on this subject referred to in the Panel discussion suggests that many more than four parameters are required to create a mathematical function that will represent a drawing of an elephant, but the point is clear. If a relatively limited data set is used to estimate the values of many parameters, it will usually be possible to arrive at a "fit" that appears excellent, but which in reality has no predictive power because the curve has simply been made to arbitrarily follow the data points, including all the experimental errors in the data, regardless of the accuracy of the implied mechanistic structure of the mathematical relationships in the model.

PBPK/PD models generally have many more parameters than can be usefully estimated with available data sets on the time course of the concentrations of parent chemical and metabolites in different places in the body and excreta. Therefore, the tradition is to divide parameters into two subsets:

- The vast majority of the parameters are estimated using data that are exogenous to the model and are thought to be relatively robust whatever conclusions are drawn about the values of other model parameters. These exogenous inputs to the model typically include organ sizes, blood flow rates, alveolar ventilation, and tissue/blood partition coefficients based on direct measurements. The inputs may also include such relatively distinctive elements of the ERDEM model as transit rates of food through portions of the gastrointestinal system, bile and urinary flow rates.

- By contrast, a few types of parameters are usually estimated by "fitting" or "calibrating" those model parameters so that model outputs correspond to some set of time course concentration data collected following a defined in vivo experimental administration of the compound to animals or people. This second set of parameters typically includes Michaelis-Menten enzyme kinetic parameters (V_{max} and K_m—the asymptotic maximum velocity of the reaction and the concentration of substrate that elicits half the maximum reaction rate) and dermal absorption rates.

The ERDEM carbaryl model contains more parameters of the fitted category than usual, in part because the modelers tried to follow the whole ensemble of carbaryl metabolites in order to (1) use calibrating information in the form of total ^{14}C levels in some compartments, and urinary outputs of specific metabolites for rats, and (2) draw inferences about internal exposure and, indirectly, brain cholinesterase inhibition from urinary metabolite data in exposed children. Several Panelists were critical of the ERDEM modelers' choice to adjust the brain/blood partition coefficient (and only that one partition coefficient) to achieve better agreement between observed and predicted brain concentrations, in part because this converts a parameter that was estimated on the basis of exogenous physical chemical information into one that is fitted via the model to specific data (see below for a more extended discussion).

b) The need for more formal statistical criteria for model fitting

Many of the Panelists strongly favored a more formal statistical approach to model fitting/calibration using the existing data. This does not necessarily imply a completely equivalent (unweighted) treatment of all data—some types of data were judged by different Panelists to be more salient for risk-related predictions than others (see below), but development of some objective and statistically defensible procedure was widely considered to be an important aspect that should be considered for the next iteration of the ERDEM system.

c) Model “validation”

“Validation” is the term used to describe a process whereby some predictions from a model are compared to observed data that were not used in the original calibration of model parameters. This process is most useful when there are relatively large amounts of data available to estimate relatively few parameters, unlike the present case where the available data are relatively sparse in relation to the number of parameters.

Some Panelists with experience in PBPK/PD modeling were critical of “validation” as a term. They believe that it implicitly suggests that a model that survives a comparison with independent data is therefore “valid” or “true”. In fact, while such a successful test can increase confidence in model predictions, in general there may be many other but untested models that would have equal mechanistic plausibility and not be incompatible with the same data.

At least one of these Panelists suggested that instead of “validation”, modelers describe the process as “juxtaposition of model predictions and data.” This treats the data and the model predictions on a more even footing. In fact, when model predictions and data are incompatible, the data can be wrong, the model can be wrong, or both can be wrong in varying degrees. Discrepancies between model predictions and observations present the modeler with a need/opportunity to reflect, and create different hypotheses about the likely source(s) of the discrepancies. In this process it is helpful to look for patterns in the observed/predicted differences that can suggest the mechanistic origins of systematic errors in model structure or data. Are the discrepancies more prominent with one strain or gender of animals than another? At early versus late time points? At high versus low doses? In one tissue more than another? Each of these types of patterns can suggest hypotheses that can be the starting points for further examination with other information. Such hypotheses also can be used in sensitivity analyses to assess how much the ultimate risk-related result might be changed if the model was restructured in various ways, or data sources were weighted in ways that emphasized the least problematic or most likely relevant information for environmental exposures.

d) Group average versus individual values of model parameters

One Panelist pointed out that the ERDEM carbaryl model parameter values result from fitting one set of coefficients to an average response from replicate animals or people. However, if each individual has a unique set of coefficients for model parameters, the solution of averaged parameter values will not generally be the same as the average solution. To assess this, the modeling team should obtain, present, and analyze the raw data for individual animals if possible. The modeling team may eventually need to generate solutions at random sets of correlated coefficient values and look at the distribution of the results to obtain better insights into the effects of individual variability on both averages and variability distributions for risk related outcomes.

Evaluation of the Current ERDEM Model Fits and Implications of the Current Results

a) Which observations should be emphasized in model calibration?

The apparent widespread use of total ^{14}C levels in various organs for the purpose of calibration drew critical comments from several, but not all, Panelists because this represents an uncertain and varying proportion of unchanged carbaryl and metabolites at different times and places in the body. Virtually all who commented preferred to emphasize the data on unchanged carbaryl for model calibration as being more causally relevant for the endpoint of cholinesterase inhibition. However, one Panelist indicated that measurements of total ^{14}C were likely to have been more accurately measured.

One Panelist also emphasized the need to use some urinary metabolite information to help calibrate the rat model, because urinary metabolite data for human children represents essentially all the data that are available for calibration of the human model. In response, EPA staff members indicated that information on the total metabolite distribution cumulated at long times after exposure were an important part of the calibration of the relative values for the metabolism components for specific metabolites in the rat model.

Another preference that was widely shared among the Panelists was for calibration of the model to make predictions that most closely correspond with observations at relatively low doses that are closest to those likely to be experienced in people. The EPA staff presenter for the ERDEM model expressed a similar preference but the system he showed for summarizing the model/data discrepancies with dose as a linear regression fit may tend to implicitly emphasize discrepancies at relatively high exposure levels. Expression of discrepancies as the weighted sum of squares of log deviations of observations from predictions would avoid this inadvertent emphasis on the fits at high dose levels. More sophisticated likelihood methods (again, most likely based on a log transformation of observations and predictions) could similarly be used to specifically evaluate discrepancies at lower dose levels.

A final issue was the choice of the registrant's Missouri rather than the California data for drawing modeling inferences about human exposures. The reason given for this was that the Missouri data reflected adherence to the label directions for application of the carbaryl-containing products. This may be a reasonable policy-related consideration in the evaluation of the data for purposes of re-registration decisions. However for purposes of creating a model descriptive of actual human exposures the Panel suggested including the California data. If use at greater than recommended rates occurred within the registrant's own controlled experiment, it is at least plausible that similar departures from ideal application occur in the practical use of carbaryl-containing products in the much less controlled settings of residential use by consumers and perhaps also commercial applicators. Moreover, the California data could provide quantitative information on the responses of urinary biomarkers to exposures over a much broader

range of exposures—allowing better calibration and testing of model predictions and inputs such as assumed human metabolism "K_m"s that would otherwise be assumed to be identical to estimates from the animal data.

b) Anomalies in the current fitting, and implications for adapting model inputs and refitting

The Panel agreed that the model developed has promise, as outlined in a summary slide, to

- Provide structure to evaluate the effects of diverse data on inferred risk
- Efficiently evaluate different exposure scenarios
- Identify factors affecting dose

However, the Panel had difficulty evaluating the quality of such outputs without a better understanding of the source of all the parameters put into the model, and improvements to reconcile many of the model/data departures that are apparent in the Appendix.

Discussion in an earlier subsection emphasized the need to reflect on patterns in model/data discrepancies to form mechanistic hypotheses about the source(s) of possible systematic errors. In doing this, the modeler also asks: (1) which parameter estimates are more vulnerable to larger changes because of experimental variability or alternative interpretations (e.g. alternative metabolite compositions of total ^{14}C) and (2) which parameters are key to the determination of the ultimate assessment parameters that affect risk (e.g. unchanged carbaryl versus metabolite concentrations). The conclusions from the first reflection process help identify which parameters should be allowed to vary, if they are influential in affecting the model fit to observed information. The decision to radically adjust the brain partition coefficient, and only that one, is one of the more questionable choices in model development in the opinion of several Panelists. The results of the current data contain interesting indications that perhaps errors in estimating other partition coefficients may be contributing to the excessive model predictions of brain carbaryl concentrations that caused the modelers to drastically reduce the predicted brain/blood partition coefficient from what was predicted on physiochemical grounds.

Many of the data/model fits are not as good as they should be for a final model. In particular, the model seriously overestimates brain ^{14}C concentrations and somewhat overestimates brain and blood ^{14}C (Figures E2 and E3, p.130 in EPA draft document "Assessment of Carbaryl Exposure Following Turf Application Using a Physiologically Based Pharmacokinetic Model"). On the other hand, the model (after the suspect adjustment of the brain/blood partition coefficient) does better at estimating the brain carbaryl concentration but if anything underestimates brain carbaryl at short times after exposure. There are huge discrepancies between data and model predictions for blood naphthol and naphthol sulfate concentrations (Figures E10 and E11, p.134). Brain and fat carbaryl concentration kinetics also show appreciable differences between data and

model predictions (Figures E18 and E19, p.138). Peak brain carbaryl (Figure E18) is under predicted by two-fold. In the case of fat, the time of occurrence of peak levels (Figure E19) indicates some fundamental problem with the partitioning or effective flow rates or blood brain transfer kinetics. Similarly, in Figure E22 (p.140) it can be seen that the predicted brain napthol concentrations underestimate the observed data by about seven-fold. Liver and blood napthol comparisons are off by something like ten-fold (E24, p.141). Therefore, improvements in the model for rats are needed before extending it to make inferences about cholinesterase inhibition from urinary metabolite data in humans.

Of these anomalies, the one that is particularly likely to be salient for alternative model development, in the view of one Panel member, is the apparent pattern of delay in the observed time to peak levels of carbaryl in the fat relative to that predicted by the current model (Figure E19, p.138). This suggests one or two possible underlying differences between reality and the model inputs: (a) The estimated fat/blood partition coefficient is too low and/or (b) the rate of transfer from the blood to the fat is slower than expected from perfusion-limited transport due to a relatively slowly-reversible binding of carbaryl in the blood. Binding in the blood would tend to lower the effective brain/blood partition coefficient at longer times after oral exposure and also might contribute to an explanation of the failure (commented on by several Panelists) to measure detectable levels of carbaryl in the blood.

To begin to explore this issue, one Panelist utilized his own data base of rat and human tissue/blood partition coefficients and regression model formulae (previously used for a few different published PBPK models (Ginsberg et al., 1996; Ginsberg et al., 2004; Walker et al., 2004)) to estimate the fat/blood partition coefficients and the half-life for perfusion based removal of carbaryl from fat in rats and humans based on the same octanol/water partition coefficient for carbaryl used by the ERDEM authors. (Spreadsheets containing these data bases and the formulae for tissue/blood partition coefficient estimation are available in the EPA OPP Docket (Docket Number: OPP 2005-0405). The results are presented in Table 1. It can be seen that the alternative estimates of the fat/blood partition coefficients are about twice as large as the estimates currently used in the rat model in ERDEM, and about three times as large in the case of humans. This leads to expectations of considerably longer elimination half lives for carbaryl in fat—in the direction needed to at least partially account for the model/data difference for the fat compartment seen in the rat model used in ERDEM. It is suggested that the set of alternative partition coefficients could be used as an alternative starting point for sensitivity analysis for the ERDEM carbaryl modeling exercise.

Table 1
Comparative Estimates of Fat/Blood Partition Coefficients and Elimination Half Lives for Carbaryl in Fat in the Current ERDEM Model versus an Alternative Regression-Based Method

Species	Body Weight (kg)	Fat Volume (L)	Fat Perfusion (L/Hour)	Fat/Blood Partition Coefficient	Source of Fat/Blood Partition Coefficient	Elim. Rate Constant from Fat (1/hr) by Perfusion	Half Life of Carbaryl in Fat (hr)	Half life of Carbaryl in Fat (min)
Rats	0.25	0.015	0.51	17.1	ERDEM	2.0	0.35	21
Rats	0.25	0.015	0.51	33.9	Hattis rat regression model	1.0	0.69	41
Humans--9 yr	20.4	2.35	16.9	17.1	ERDEM	0.42	1.7	99
Humans--9 yr	20.4	2.35	16.9	63.6	Hattis rat regression model	0.11	6.1	368

Issue 2.2: PBPK/PD Model Fidelity

The PBPK/PD model was developed in the Exposure Related Dose Estimating Model (ERDEM) platform. The ERDEM platform is, by design, highly structured and flexible for adaptation to new or emerging exposure and risk assessment needs. In PBPK/PD modeling, there is a need to balance completeness regarding anatomical/physiological pathways/routes with the desire for model simplicity. The model is required to simulate the relevant dose metrics and provide the capability to extrapolate from the laboratory setting to exposure scenarios of interest. Modeling runs with the ERDEM platform are typically short; thus computational time is not an issue. The important consideration is the *in silico* representation of the species and the connection with pathways and routes of exposure.

Question 2.2a

Please comment on the carbaryl PBPK/PD model structure for evaluating diverse exposure scenarios, including the exposure to children on the turf described in the report. Please include in your comments a consideration of the degree to which the compartments included in this model reasonably describe the PK and PD characteristics of carbaryl and provide the ability to extrapolate the model across species and scenarios, balanced against model simplicity.

Question 2.2b

As more PK and PD data become available, the model structure from this application may be applied to other N-methyl carbamates, including mixtures. Please comment on the suitability of the carbaryl specific PBPK/PD model structure as developed in the ERDEM platform for expansion to include other N-methyl carbamates.

Panel Response (to Question 2.2 a & b)

The Panel commends the Agency and its scientists who are building this comprehensive PBPK/PD model for their work which has lead to an important milestone in the utility of mechanism-based models. This model is relevant in the present context for cumulative risk assessment and will be important for doing biologically-based risk assessments for single compounds. The model has achieved a level of maturation where all of the relevant parts and modules are publicly available and ready for open scientific scrutiny. This model can be applied to the estimation of biomarkers of exposure and, to a lesser extent, biomarkers of effect.

At this stage of model development, it is important to consider the applicability and acceptability of the model. The model is extremely complex and the Panel expressed concern that this level of complexity may not have sufficient support in the available data. Some modules of the model are structurally identical (e.g. those used as simple storage compartments) and could be considered as sufficiently similar to be treated with one generic differential equation thereby reducing the apparent complexity. A second means of reducing apparent complexity would be to group modules of similar influence on the AChE levels (e.g. in response to the results of a sensitivity analysis). Finally, the Agency may want to consider grouping some of the different metabolic rates to reduce the need for 2 parameters for each metabolite.

However, some changes are needed that might increase complexity. Dermal absorption seems to be inadequately modeled. A portion of carbaryl absorbed by the skin and stored in the dermis may not be readily available for distribution in the circulation. This is evident by the results in Figure 12 on page 36 of the draft document "Assessment of carbaryl exposure following turf application using a physiologically based pharmacokinetic model" provided by the Agency to the Panel, in which the ERDEM model overestimates the peak total amount of C^{14} in blood by a factor of 2.5. An additional differential term describing the release of carbaryl from dermis to blood may need to be added to this model. Other results from the simulation for simultaneous oral and dermal exposure to carbaryl all show some degree of differences between model prediction and the observations. The Dermal Thin Film technology, which is being developed in ORD, also could be used to investigate the extent of carbaryl or N-methyl carbamate retention in dermis and the rate of release to blood.

The other missing component in the ERDEM model is a description of how the baseline AChE level in brain tissue is incorporated. The current approach of using animal baseline AChE activities as the input levels is inadequate and may be wrong. The

fundamental limitation is that baseline AChE brain level in humans of all ages is hard to obtain. This is why the SAP encourages EPA and the registrants to expand the focus on building PBPK/PD model to include both brain tissue and red blood cells (RBC). As long as the RBC AChE inhibition can be modeled appropriately and adequately in the PBPK/PD model, the brain tissue AChE inhibition can then be predicted with a much better accuracy.

Finally, on model changes, the model appears to treat AChE inhibition as a reversible binding mechanism when in fact, it is more of a metabolism mechanism; this needs to be corrected in later versions of the model.

The runtime version of ERDEM was made available by the Agency and it was possible to execute this version without problems. It produces a logfile where all relevant information on the set up, the performance and the results of the program can be inspected. This wealth of information requires answers to several questions:

- How will the model output be organized?
- Which tables and which graphs of the numerical results will be stored in which format in which directory of the user's PC? Will these be files that can be easily edited?
- How will the model characteristics be made available so that modeling can be replicated?

The model seems to be general enough to cover the PBPK/PD modeling of single carbamates. However, it is difficult for the Panel to discuss additions or subtractions to this model at this point because it has not been applied rigorously and to a broader class of agents. As long as other NMCs behave like carbaryl, this carbaryl PBPK/PD model should be able to be used for other NMCs and mixtures. Extra steps would need to be added for NMCs which requires metabolic activation and for those that have more than one active metabolite (e.g., aldicarb). Caution should be used when the components of the PD model are used for other NMCs without additional data. The major concern is that the reactivation time, or rate of AChE inhibition, will vary by NMC. The other concern is that AChE-inhibition by NMCs, rather than by carbaryl alone, is assumed to be dose-additive. While additional animal studies are being conducted to determine whether this assumption is valid, caution should again be taken. The Panel encourages the Agency to evaluate how the model performs when all seven carbamates and their metabolites are modeled simultaneously using different exposure routes.

The ERDEM model excludes a large part of the pharmacodynamics (PD) of carbamate toxicity. Documents provided by the Agency indicate that other PD endpoints are currently under investigation (e.g. motor activity) which may represent more advanced health endpoints than AChE inhibition. As the Agency develops the ERDEM model, they are encouraged to pursue these other PD endpoints.

The model uses various allometric scaling factors for adjusting model parameters including the determination of respective parameters for young children. Price et al. (2003) have collected a series of equations for estimating organ and tissue volumes using National Health and Nutrition Examination Survey (NHANES) anthropometric data. It was not clear from the documents and the presentation whether those equations have been used for the model, and if not, why they have been excluded.

The terms and techniques being used by the Agency for predictions and data comparisons at the 99.9 percentile (e.g. Figure 20 of the aforementioned Agency draft) need clarification. The model predictions do not appear to be 99.9% predictions of cumulative mass but are instead, based on a highly unlikely exposure event. The analysis does not appear to include other uncertainties and variabilities in the model that might impact the predictions. Moreover, comparing model predictions to data at the tails of the distribution is questionable at this time. The research going into developing the PBPK/PD model is based largely on general, aggregated observations. Until the model has been exercised in greater detail, the Panel suggests focusing the discussion on predictions at the median and within the interquartile range.

An additional comment was made regarding the priority for developing the structure ERDEM. One Panel member suggested that an additional component be added to ERDEM for modeling fetal *in utero* exposures, with the ability to model endpoints at target tissues, such as the fetal brain.

Issue 2.3: Statistical Model Evaluation Considerations

Development of PBPK/PD models is an iterative process such that the model is improved and revised as more data and information become available. In a regulatory setting, it's not unusual that model development begins before all data sets have been collected. Currently when using ERDEM, an initial model structure is developed based on the species physiology and known chemistry of the chemical and metabolites. The initial model structure consists of the differential equations and variables that correspond to the relevant compartments and metabolic transformations. The initial parameter values are estimates made by the researcher, often based on models of related chemicals. The model is considered provisional until the available data are evaluated. The parameter values and model structure are then updated to reflect the available data. As new data are made available, they are evaluated concurrently with the existing data against the model simulations, and the model is revised accordingly. This iterative process has been followed for the current carbaryl PBPK/PD model, where model evaluation was based on visual inspection and linear regression between the model results and data points (not included in the report).

Question 2.3

Please comment on statistical or mathematical analyses which could inform the need for model revisions as new data are made available.

Panel Response

The Panel discussion on this question centered on the following three issues:

- the nature and use of data for determining model complexity and parameter estimation,
- approaches to assessing goodness-of-fit of the model, and
- the need for assessing model complexity and parameterization sensitivity.

Model complexity and parameter estimation

The example model presented to the Panel is complex, with over 80 parameters: 25 physiological parameters, 24 physiochemical parameters, 35 biochemical parameters and 10 inhibition and recovery rate parameters. Norm settings for these parameters are provided in Tables 3 to 9 of the draft report "Assessment of carbaryl exposure following turf application using a physiologically based pharmacokinetic model." The response data, dose metrics, available to assess the goodness of model predictions are a small dataset provided in Table 1. Figures 7 to 19 demonstrate that the model when evaluated at the norm settings of the 80 parameters tends to follow the observed pattern of recovery, sometimes overestimating the amount of ^{14}C radiation or the amount of carbaryl, etc and other times underestimating. From a statistical point of view, the model is used to describe a complex set of multivariate data and hence it is not surprising that some over- and underestimation occurs.

The model is not parameterized to a specific individual, nor does it use a global parameter estimation method, such as non-linear least squares, maximum likelihood or other statistical approach. Some Panel members were not surprised by this, suggesting that the typical model building approach is to determine parameter values for individual compartments using available data for these compartments. Other Panel members felt that more formal model-building and parameter estimation should have been displayed.

The Panel discussed the model construction and parameter estimation process in some detail. Two approaches are available for parameter estimation. In the global approach, the whole problem is formulated as a large, multidimensional search for the values of the parameters that minimize the residual sums of squared deviations between the model predictions and any available data. This can be done with or without constraints on the parameter values. Unfortunately, with highly parameterized models and little data, model fit can be amazingly good, but the fitted model may be relatively useless for prediction purposes. This is not unlike the phenomenon of fitting higher order polynomials to limited data; at some point the flexibility of the model allows the fitted line to perfectly intersect all available data points.

As an alternative to a global approach, least squares methods may be employed in a piece-wise manner. Each parameter or related set of parameters is treated as a separate analysis. This approach should make it easier to incorporate constraint information and utilize or test intermediate variable values. The documents provided to the Panel discuss model fits for various compartments for the parent compound as well as the metabolites with various endpoints. It was pointed out that when a model is constructed or parameterized by successively incorporating information from different data sets with adjustments made to model parameter values with each addition, the order in which the data are incorporated and the changes made will influence the final model form and associated parameter values. This is not unlike what happens using stepwise regression in the presence of mild co-linearity. The strategy of model building, model refinement and parameter estimation are all interrelated; changes in any one affects the other and together influence the final model and its predictability. Therefore, transparent model building should include keeping track of what data are used to determine settings for each parameter. A chronology of model adaptations also should be available to reviewers.

An intermediate to the piece-wise approach and the global approach is to fit the model by using the data to inform us on only a couple of parameters on any particular run. This reduces the dimensionality of the fitting process by making the fit of the parameters conditional on the other parameters being held fixed at their current "confirmed" values. This is a useful approach when using data from many different experiments. In a sense one calibrates the model for each set of data and examines the predictions from the fitted model from each set. This approach can illuminate when additional information will be useful in improving model predictions.

The Panel also discussed the more difficult issues of the dimensionality and complexity needed in the model. It seems that the goal of a PB/PK model is to mimic reality and integrate current understanding of carbaryl-related processes. Compartments and parameters in the model are essentially there, based on theoretical justifications and not necessarily because of their ability to be directly observed and/or calibrated using achievable observations. Some compartments and/or parameters may have only indirect relationships to measurable outcomes and as such are only slightly correlated with these outcomes. This model allows us to ask "what if" questions whose answers cannot be directly derived from observations. A number of Panel members felt that the PB/PK model should be used as a check against simpler models that can be, and are, calibrated with existing data to important outcomes. It was felt that there is still the need to collate the relevant empirical evidence, conduct appropriate parameterizations that pass statistical muster, and express the results in a way that allows us to characterize uncertainty in the model predictions. Reference was made to comparing the PBPK/PD model results with the results from the simpler and empirically-based PK model discussed in a preceding session (Session 1: Issues related to cumulative hazard assessment). Comparisons such as these offer insights into the amount of model complexity that is necessary to describe relevant outcomes.

Characterization of prediction uncertainty for the full PBPK/PD model is difficult because there are both model form uncertainties as well as parameter uncertainties. For example, there was discussion relating to the situation where the compartment-specific data suggest one value for a parameter (e.g., dermal permeability coefficient), but the parameter value had to be dramatically changed from this data-derived value in order to get a better global model fit. At least one Panel member felt that this is exactly the kind of situation that would justify the need to reexamine not just the particular assigned parameter value but also the form of the compartment model. Contradictions such as these lead to increased scientific understanding and better final models. This also brought up the issue of how new information is used to modify the model and/or model parameters. There seems to be a need for protocols to help decide when new data improve on existing data. At least one Panel member suggested that human data are preferred over animal data, *in vivo* data preferred over *in vitro* data and that direct observations be preferred over derived values. Incoming new data also could be investigated with a series of models of increasing complexity (again, in a manner similar to what is done in model building in multiple regressions).

Some Panel members belief that a combination of model simplification (judicial use of Occam's Razor) and formal parameter estimation via traditional least squares or Bayesian methods (e.g. formal true Bayesian or hierarchical Bayesian updating techniques, Gelman et al 1996; Bois, 2000) will result in a model whose predictions are easier to evaluate. Many would appreciate seeing how various model simplification approaches would affect the final model fit, e.g. stripping off parts of a model, combining compartments or eliminating compartments.

For a given model formulation and where possible, parameter estimates should be accompanied by estimated standard errors and model predictions presented with associated prediction intervals. It was noted that the confidence intervals presented in Table C2 (p.103) and illustrated in Figure C1 (p.102) (draft document "Assessment of carbaryl exposure following turf application using a physiologically based pharmacokinetic model") are confusing; reporting lower limits that are negative and plotting only the upper limits. It was not clear if one- or two-sided intervals are used, nor was it clear which are desired.

Also, it was noticed that much of the data used for parameter estimation is censored in some way. Traditional least squares methodology is not directly applicable in this case, and how censored values are handled can have a dramatic effect on estimates and overall model fit. The more formal statistical estimation approaches may require censored data likelihoods or weighted likelihoods which are quite difficult to handle properly. The problem is multiplied when one is attempting to handle multiple datasets in the process.

Goodness-of-fit

Goodness-of-fit of the model is highly related to the complexity of the model and

parameter estimation. For this reason, the discussion on goodness-of-fit was interspersed throughout the Panel deliberations.

The applicability of statistical goodness-of-fit methods to assess the adequacy of model predictions is limited by the complexity of the PBPK/PD model, the type of data, the limited amount of data available to fit the model, and the aims of the modeling. Several views were expressed among the Panel members on the degree to which statistical approaches can and should be used.

Some Panel members preferred a more qualitative assessment of model goodness-of-fit, offering that adequacy might be judged by answers to the following questions.

- Does the model capture the main points of the biological processes?
- Does the model adequately describe available outcomes data for important dose levels and time ranges?
- What can be learned by fitting the model to each of the available data sets?
- Which changes to model parameters are needed to improve the fit in a specific key compartment and do these changes result in deterioration of the fit in other compartments?

A number of suggestions were made for more quantitative ways to assess model goodness-of-fit.

- If least squares methodology is used to estimate model parameters, the residual mean square can be used to assess model fit. Modifications to this approach include the Akaike Information Criterion (AIC) which joins the estimated mean square error to the number of parameters actually estimated in the model.
- Evaluate model complexity by building a set of nested sub-models by systematically removing or simplifying compartments. Again, residual mean squares, AIC or other criterion can then be compared, either informally or via formal statistical tests where available, to assess the significance of differences between full and reduced models.
- Implement dynamic graphics which would allow visualization of the changes in model fit produced when one or a small number of parameters change their values continuously (movie-type visualization of model sensitivity).
- Utilize some form of the Press statistic ($g_{\text{Press}} = \sum_i (\hat{y}_i - y_i)^2$) to evaluate the goodness of model predictions to data not used to calibrate the model.
- Utilize other criteria, such as the minimum absolute distance between observed and predicted values (MAD, $g_{MAD} = \sum_i |\hat{y}_i - y_i|$), or curve comparison statistics like an Anderson-Darling distance to measure goodness-of-fit.

- Consider weighting some times more than others in the goodness-of-fit criteria and evaluate how this changes model fits (e.g., a weighted Press statistic, $g_{WP} = \sum_i w_i \left(\hat{y}_i - y_i \right)^2$).
- If a likelihood function can be created, then likelihood goodness-of-fit statistics and associated tests become available.

Establishing a quantitative criterion for assessing level of model fit is important in determining at what point one achieves acceptable model fit. What is needed to set this value is the sampling distribution of the goodness-of-fit statistic. It is unlikely that this will be analytically available and hence some form of Monte Carlo simulation study will be needed to develop this distribution. Goodness-of-fit statistics are available to assess not just the overall goodness-of-fit of the model to observed outcomes, but can be used to identify where lack of fit occurs e.g. does the model underestimate brain concentrations, overestimate blood levels, etc.

The entire discussion by the Panel emphasized the need in this study to identify the critical end points for the model. It was not clear whether it was more important for the model to correctly predict peak concentrations, times at which peak concentrations were reached, peak or times for specific compartments (brain or blood) or some other aggregate measure of prediction fit (as mentioned above). Some Panel members were not certain that fits on the temporal scale were best. Other Panel members were concerned with whether model fits should be made on untransformed concentration values or log transformed concentrations. Model fits to untransformed concentrations tends to allow large deviations, typically at higher concentrations to dominate the goodness-of-fit statistic. Assuming multiplicative errors and natural log scaling allows better fits for low doses and helps avoid situations where parameter estimates are inappropriately estimated as negative.

There was some discussion about choice of the settings for the human adult model. Typically such models are designed for 20 year old males assuming 6% body fat. The Agency may wish to utilize NHANES data or the findings of the Price et al (2003) study to define a more realistic typical individual.

Sensitivity analysis

The ERDEM PBPK/PD model is a complex system that involves many interdependent functional relationships. Estimates or assigned values for many parameters of this system have varying degrees of current empirical support and all are subject to a degree of natural variability as well as "reducible" uncertainty for specific applications. There was general consensus among the Panel of the need to perform a sensitivity analysis for the PBPK/PD model. A limited sensitivity analysis was done but a more extensive analysis needs to be performed. In general, it was felt that there may be only a few parameters that correlate with model predictions. A common theme of the discussion was that, in general, researchers work very diligently to build the best model

they are capable of doing and then maintain skepticism in the model outputs. A sensitivity analysis is one way this is done.

Sensitivity analyses typically include an assessment of the effect of parameter uncertainty on model predictions. This is typically assessed by applying prior distributions to parameters and assessing the degree to which changes in parameter values correlate with changes in model predictions. With highly parameterized models this is a very difficult task. But it was felt that a proper uncertainty analysis allows users of the model, and eventually regulatory assessors, the ability to place model predictions in their proper context.

In terms of propagating uncertainties to endpoint predictions, the EPA should put more attention to errors caused by systematic model misspecification in the PBPK/PD model. This includes concerns with over parameterization. The uncertainty (and variability) in model predictions depends heavily on the selected model. If the model is incorrect, depending on the degree of error, the uncertainty in the final model predictions also will be incorrect.

In assessing uncertainty, it seems reasonable to include terms and pathways in the model that may have small contributions to the overall uncertainty, simply because their inclusion usually doesn't make much difference and leaving them out begs the question of what effect they might have had. A process to evaluate reductions in the dimensionality and complexity of the model is needed to support calibration of the models. Some Panel members felt that the dimensionality of the model should be solely the responsibility of the model builder, but others would like to see an analysis of how model form changes affect predictions and changes sensitivity.

The Panel also discussed the idea of letting the model form itself be uncertain. For example, the model developer could generate a suite of models, all of which are acceptable but which illustrate the range of understanding of the form of the underlying processes and expert opinion on how compartments might interact. Some models might be highly defined, consisting of many compartments and parameters, whereas others could be very simple, based on purely empirical relations. It also may be possible to assign a weight to each model proportional to the likelihood that the particular model properly describes the PBPK/PD system. As in a Monte Carlo sampling procedure, one could then generate model predictions of endpoints from each of the models, and from multiple samples of the parameter uncertainties for the respective model. This would result in a collection of model predictions based on uncertainty in both parameters and model specification and the resulting prediction distributions or prediction bands would better convey overall uncertainty. This information complements the measures of goodness-of-fit resulting from formal parameter estimation.

Not all Panel members wanted to go quite as far as the full parameter and model uncertainty analysis discussed in the previous paragraph. One member preferred a well-structured statistical analysis to a post-hoc uncertainty analysis; allocating 90% effort to

parameter estimation and only 10% to assessing uncertainty. Others felt that it was crucially important that the underlying uncertainty be assessed in explicit probabilistic simulations. One Panel member mentioned that a Bayesian Network approach (Borsuk et al., 2001, 2004) might be useful but another Panel member stated that this approach has been examined and found unacceptable for use with dynamic system models such as these PBPK/PD models.

There was discussion that advances in computational algorithms and computer technologies have eliminated many of the computational barriers to full uncertainty analysis. What is required is careful scientific specification of the model structure and its implied conditional, probabilistic relationships. Specification of prior probabilities for parameter values requires good empirical data or carefully elicited expert judgment. It does not remove the need to test how individual model components and changes in model specification will influence model outputs. Nor does it eliminate the need for sensitivity analyses to the model outputs.

Other related comments made during the Panel discussions include:

- The models discussed do not seem to explicitly incorporate inter-individual variability.
- The animal data typically used in calibration are aggregated and assumed homogeneous.
- Model fits might be more useful if they are targeted at the midpoint or center of the response distribution rather than at the 99.9^{th} percentile as seemed to be done in the report to the Panel.
- The judgment of those model fits would be substantially improved if individual response values were included on plots and not only summary data (means).
- Related to the quality of the mathematical and statistical analysis is the capability of the interface between the model application and stand-alone statistical analysis packages. The amount of data that can be created during a simulation run needs a robust statistical analysis package with modern tools of statistical diagnostics and regression methods.
- After creating Figures 7-19, the report talks about typical or expected ranges for a subset of the rate parameters as summarized in Table 13 (p.48). It would have been informative to overlay the most extreme predicted curves onto one or more of Figures 7-19. This would provide a "bound" on the prediction, allowing us to see if the observations at least fit within high-low predicted patterns. This is not exactly a confidence band but something more than the "point estimate" presented in original graphs.

Issue 2.4: Risk Metric

Historically, EPA has calculated margins of exposure (MOE) in its risk assessments for the N-methyl carbamate pesticides. These MOEs are calculated by

dividing environmental exposure concentrations by a point of departure identified from toxicity studies. These points of departure are typically no-observed-adverse-effect- levels (NOAELs) or benchmark dose estimates (BMDs). For the N-methyl carbamates, these NOAELs or BMD estimates are generally based on peak cholinesterase inhibition. The use of PPBK/PD models provides the opportunity to consider toxicological endpoints other than peak cholinesterase inhibition. Some potential toxicological endpoints include 1) peak concentration of the pesticide (or key metabolite) at the site of action; 2) total pesticide (or key metabolite) at the site of action over a period of time (e.g. area under curve); 3) peak cholinesterase inhibition; 4) inhibition at or above a pre-defined level of inhibition (e.g. BMD10); 4) duration of time for inhibition at or above a pre-defined level of inhibition. The current report explicitly provides the peak concentration of carbaryl and peak cholinesterase inhibition; although the other metrics are easily accessible from a model developed in the ERDEM platform.

Question 2.4

> **Given the toxicological characteristics of carbaryl and other N-methyl carbamate pesticides, please comment on the degree to which these toxicological endpoints are appropriate for purposes of developing a risk assessment.**

Panel Response

PBPK/PD modeling has the advantage over the conventional NOEL and MOE approach in incorporating the current understanding of mechanism of toxicity in describing the relationship between components of risk assessment, e.g., toxicity, exposure, and risk. The unique challenge of modeling the relationship between toxicity and exposure for carbaryl and other carbamate pesticides is that the pesticide and its active metabolite(s) are rapidly released after binding to cholinesterase.

The Panel generally agreed that, as far as enzyme binding is concerned, the inhibition of cholinesterase activity is the ultimate metric for expressing toxicity and risk. This could include both considerations of the peak inhibition level, the duration of sustained inhibition as expressed in the area under the curve (AUC), and the duration above a pre-defined inhibition level. On the other hand, the profiles of carbamate and its active metabolites at the site of action are pertinent metrics of exposure. These metrics are important for improving the understanding of both short- and long-term exposures (i.e., prolonged duration or repeated exposures) and can provide a linkage to the biological effects and the biomonitoring data in humans.

Thus, all of the five metrics suggested by the Agency are of interest and each of them can potentially be a valid metric for risk assessment, although some expressions may require further clarification. At this early stage of developing the PBPK/PD model

for carbaryl and possibly extending its application to other N-methyl carbamates and their cumulative exposure, the Agency is encouraged to explore the merit of each of these endpoints. In fact, multiple expressions of some of these metrics can enhance a logical and clear presentation of risk assessment and facilitate the understanding and communication between the exposure and risk assessors and between risk assessors and risk managers.

Specific comments for each metric are provided below with respect to their pertinence in risk assessment and consideration of time factor in exposure and toxicity.

1) peak concentration of the pesticide (or key metabolite) at the site of action.

The peak concentrations of pesticide or its key metabolites at the site of action are among the primary variables presented by the ERDEM model. The concentration of the parent compound is the key expression of exposure for evaluating the toxicity of carbaryl. This and the next metric (i.e., AUC) are the ultimate parameters of pharmacokinetic events after exposure, interfacing with the pharmacodynamic and toxicity metrics. When properly modeled, the peak concentration informs pathway saturation and the various degrees of probability of states of the carbaryl, e.g., free-circulating, bound to cholinesterase, metabolized as released from the enzyme complex. The peak level metric also supplies information on the speed at which the pesticide is taken up, the time-scale over which effects are likely to occur, and the maximum level of effect likely to occur.

The peak level also could serve as a pointer to compound-specific toxicities that lie outside the common mechanism and is especially relevant in an acute setting. One reason is that short-term adaptations reduce or eliminate physiological disturbances from minor to moderate levels of cholinesterase inhibition. When moderate levels of inhibition are sustained, several compensatory mechanisms come into play almost immediately, including reduction of acetylcholine release as well as receptor-desensitization and, later on, down regulation. Synaptic homeostasis is quickly restored. Hence, the AUC (see discussions under the next metric), is unlikely to be a good measure of acute toxicity.

2) total pesticide (or key metabolite) at the site of action over a period of time (e.g. area under curve)

The AUC metric includes the magnitude of effect implicitly, but also duration of effect. The concentration of the pesticide or its active metabolite in and of itself is irrelevant to the characterization of risk in isolation from the extent of enzyme inhibition. Thus, comments regarding this metric are not limited to the AUC of pesticide or its active metabolites but also the metric of AUC of cholinesterase inhibition.

An important consideration is that the compensatory responses that might be taken by the organism in response to perturbation in acetylcholine signal strength are not

necessarily benign. If sustained adjustments are made in particular synapses in the postsynaptic sensitivity or the presynaptic extent of acetylcholine release, these may well have subtle long term consequences for the transmission of signals that are important parts of neuron/neuron and neuron/muscle communication. Thus, the toxicological response to long sustained exposures could be very different from the acute exposure scenarios. Total AUC of inhibition or total time above a pre-defined level of inhibition (metric 5) can be important. In these cases, compensatory mechanisms may fail, or the consequences of the compensations may themselves be adverse. Therefore, models developed for the ERDEM platform should calculate and report these metrics.

The different toxicological implications between the peak cholinesterase inhibition (next metric) versus the AUC of inhibition are simply illustrated: in a poisoning event, 90% inhibition over 20 minutes is a more serious effect than 10% inhibition over 180 minutes. While the AUC metric would give the same results for the above two scenarios, the health outcome is likely to be substantially different. On the other hand, for non-acute poisoning events, e.g., 10% inhibition, the integrated AUC approach would be appropriate. It may be, for example, that long-term, low-level cholinesterase inhibition gives rise to specific health outcomes of interest. Further, repeated assaults may offer additional effects. The scenarios outlined in the document presented to the Panel - adult applicator and child playing on treated turf offer examples that could give rise to similar integrated effects with entirely different profiles.

3) Peak cholinesterase inhibition

As described above, peak inhibition of cholinesterase activity is the most important endpoint for evaluating acute toxicity with carbamate anticholinesterases. It is the most direct measure of adverse effect mediated through the common mechanism of toxicity for carbamates. Thus, for short term effects, peak inhibition is a good first judgment for the dosimeter that is most likely to be predictive of toxicity.

The expression of peak cholinesterase inhibition at a given dose is comparable to the current toxicity data that defines the threshold as the no-observed-effect level (NOEL) based on the peak inhibition measured at the assumed peak time of effects. Thus, the PBPK/PD model can be used to estimate the peak inhibition after human exposure and compared to the level defined as adverse when exceeded, i.e., threshold cholinesterase inhibition. Depending on the extent that a PBPK/PD model may address the inter-species and inter-individual variation in sensitivity (e.g., two individuals exposed to identical level of carbaryl may show different levels of inhibition due to differential susceptibility), uncertainty factor(s) may be modeled or applied in this comparison.

The similarity of this metric to the conventional NOEL or BMD has the added advantage of facilitating risk communication while transitioning into the mechanistic approach. Modeling peak inhibition also allows the possibility of correlating this metric to other related neurological endpoints (e.g., neurobehavioral and clinical signs).

An apparent unknown is what dosimeter is closest to a causal determinant of the subsequent pharmacodynamic processes—from overt symptoms in adults to putative marginal strengthening of signaling along some pathways rather than others during development to more subtle adaptations of the synaptic responsiveness to later cholinergic stimuli. This is where we really need progress in basic science to show how cells set and reset their set points for responsiveness in relation to the degree and duration of cholinesterase inhibition. Thus, for current assessments, both peak inhibition and the AUC (the amount of inhibition times time) should be used as dose metrics for assessing MOE and risks of response.

4) Inhibition at or above a pre-defined level of inhibition (e.g. BMD_{10})

One Panel member commented that this metric requires some clarification. Simplistically, a PBPK/PD model can define the dose that results in a pre-defined peak level of cholinesterase inhibition, equivalent to the BMD (e.g., 10% inhibition in the brain). In a conventional approach, this metric may then be used to generate the relative potency factor for acute toxicity of multiple carbamates; if it is determined that peak inhibition is the valid metric for expressing acute toxicity. The main drawback is that this expression does not implicitly account for the time factor as the next metric (i.e., duration of time for inhibition at or above a pre-defined level of inhibition). Instead, this metric suggests a dichotomous variable - exposure resulting in inhibition above a fixed level. Thus, this metric may be less useful than knowing how long such an inhibition occurred. Clearly there is a difference in inhibition above 10% that lasted for five seconds versus a similar inhibition, due to chronic exposure, that lasts for many days.

5) Duration of time for inhibition at or above a pre-defined level of inhibition.

Being able to take into account the duration/time factor is one satisfying feature of this metric. Assuming that the pre-defined level is selected with some intrinsic meaning (e.g., a health outcome of interest), the duration of the insult experienced by the body represents a significant measure. Short-duration insults are likely to produce less effect than long-term sustained insult of the same magnitude. The selection of this metric pre-supposes that a pre-defined level is below some frank effect level. Otherwise, the effect can be defined by other means. Furthermore, the duration should be relative to the expected clearance rate. For example, it can be asked: is this elevated level maintained for longer than the lifetime of carbaryl in the body and thus may be indicative of an ongoing, albeit low-level, exposure?

REFERENCES

Bois, F. (2000) Statistical Analysis of Fisher et al. PBPK Model of Trichloroethylene Kinetics. Env. Health Perspectives 8, Supplement 2: 275-282.

Borsuk, M. E., C. A. Stow, and K. H. Reckhow (2004). A Bayesian network of eutrophication models for synthesis, prediction, and uncertainty analysis. Ecological Modelling 173:219-239.

Borsuk, M. E., C. A. Stow, D. Higdon, and K. H. Reckhow (2001). A Bayesian hierarchical model to predict benthic oxygen demand from organic matter loading in estuaries and coastal zones. Ecological Modelling 143:165-181.

Gelman, A., F. Bois, and J. Jiang (1996). Physiological Pharmacokinetic Analysis Using Population Modeling and Informative Prior Distributions. J. of Am. Stat. Association 91(436): 1400-1412.

Ginsberg, G. L., Pepelko, W. E., Goble, R. L., and Hattis, D. B. (1996). Comparison of Contact Site Cancer Potency Across Dose Routes: Case Study with Epichlorohydrin. Risk Analysis 16:667-681.

Ginsberg, G., Hattis, D., Russ, A., and Sonawane, B. (2004). Physiologically-based pharmacokinetic (PBPK) modeling of caffeine and theophylline in neonates and adults: Implications for assessing children's risks from environmental agents. Journal of Toxicology and Environmental Health 67:297-329.

Price, P. S., Connolly, R. B., Chaisson, C. F., Gross, E. A., Young, J. S., Mathis, E. T., Tedder, D. R. (2003). Modeling interindividual variation in physiological factors used in PBPK models of humans. Critical Reviews in Toxicology 33:469-503.

Walker, K., Hattis, D., Russ, A., and Ginsberg, G. (2004). Physiologically-Based Toxicokinetic Modeling for Acrylamide—Risk Implications of Polymorphisms and Developmental Changes in Selected Metabolic Enzymes. Report from the George Perkins Marsh Institute, Clark University, and the Connecticut Department of Public Health to the U. S. Environmental Protection Agency under Cooperative Agreement #827195-0, December 2004.

SAP Minutes No. 2005-01

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

N-METHYL CARBAMATE CUMULATIVE RISK ASSESSMENT: PILOT CUMULATIVE ANALYSIS

SESSION 3: DRINKING WATER EXPOSURE ANALYSIS

February 17, 2005
FIFRA Scientific Advisory Panel Meeting held at the Holiday Inn National Airport Arlington, VA

Mr. Joseph E. Bailey
Designated Federal Official
FIFRA Scientific Advisory Panel
Date: April 15, 2005

Steven G. Heeringa, Ph.D.
FIFRA SAP Session Chair
FIFRA Scientific Advisory Panel
Date: April 15, 2005

Federal Insecticide, Fungicide and Rodenticide Act
Scientific Advisory Panel Meeting

February 17, 2005

N-methyl Carbamate Cumulative Risk Assessment: Pilot Cumulative Analysis

Session 3: Drinking Water Exposure Analysis

PARTICIPANTS

FIFRA SAP Session Chair

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

Designated Federal Official

Joseph E. Bailey, FIFRA Scientific Advisory Panel Staff, Office of Science Coordination and Policy, EPA

FIFRA Scientific Advisory Panel Members

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

Gary E. Isom, Ph.D., Professor of Toxicology, School of Pharmacy & Pharmacal Sciences, Purdue University, West Lafayette, IN

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

FQPA Science Review Board Members:

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

George B. Corcoran, Ph.D., Professor & Chairman, Department of Pharmaceutical Sciences, Eugene Applebaum College of Pharmacy & Health Sciences, Wayne State University, Detroit, MI

Lutz Edler, Ph.D., Head, Biostatistics Unit C060, German Cancer Research Center, Heidelberg, Germany
Bernard Engel, Ph.D., Professor, Agricultural & Biological Engineering, Purdue

University, West Lafayette, IN

Scott Ferson, Ph.D., Senior Scientist, Applied Biomathematics, Setauket, NY

Lawrence J. Fischer, Ph.D., Director, Center for Integrative Toxicology, Michigan State University, East Lansing, MI

Natalie Freeman, Ph.D., Associate Professor, Department of Physiological Sciences, College of Veterinary Medicine, University of Florida, Gainesville, FL

James P. Kehrer, Ph.D., Director, Center for Molecular & Cellular Toxicology, College of Pharmacy, The University of Texas at Austin, Austin, TX

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

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

David MacIntosh, Sc.D., Senior Associate, Environmental Health & Engineering, Inc., Newton, MA

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

Christopher J. Portier, Ph.D., Director, Environmental Toxicology Program, National Institute of Environmental Health Sciences, Research Triangle Park, NC

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

Michael D. Sohn, Ph.D., Scientist, Environmental Energy Technologies Division, Lawrence Berkeley National Laboratory, University of California, Berkeley, CA

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

Michael D. Wheeler, Ph.D., Assistant Professor, Departments of Pharmacology & Medicine, University of North Carolina, Skipper Bowles Center for Alcohol Studies, Chapel Hill, NC

INTRODUCTION

In Session 3 of this meeting, the FIFRA SAP met to consider and review the N-methyl carbamate pesticide cumulative risk assessment: pilot cumulative analysis, specifically issues related to drinking water exposure assessment. OPP solicited comment from the SAP on the use of existing ground-water models to provide a pilot ground-water exposure assessment for the N-methyl carbamate pesticides. Session 3 included presentations by Mr. Nelson Thurman and Dr. Dirk Young (Environmental Fate and Effects Division, Office of Pesticide Programs) and Dr. Tom Nolan (U.S. Geological Survey) pertaining to the use of ground water exposure models and transport models to help predict drinking water exposure estimates for the N-methyl carbamate pesticides. Dr. Steven Bradbury (Director, Environmental Fate and Effects Division, Office of Pesticide Programs) gave opening remarks for Session 3.

SUMMARY OF PANEL DISCUSSION AND RECOMMENDATIONS

The Agency solicited comments from the Panel on assessing regional groundwater vulnerability and predicting pesticide concentrations in groundwater for estimating exposures to N-methyl carbamate pesticides.

Data that can be used to assess nationwide groundwater vulnerability are quite limited. Therefore an obvious approach for assessing groundwater vulnerability to pesticides does not exist. EPA proposes to use a shallow groundwater vulnerability assessment that was developed for nitrates. However, the nitrate assessment may not be appropriate for estimating shallow groundwater vulnerability to pesticides. If the proposed approach is to be used, it should be compared with observed groundwater contamination data. Alternatively, an approach utilizing the nationwide State Soil Geographic (STATSGO) Data Base and the pesticide leaching models should be considered.

To predict concentrations of N-methyl carbamate pesticides in groundwater, the proposed approach appears to be conservative and health protective for most reasonable scenarios. In all but the most unusual cases, screens in the wells will draw from a large cross section of the aquifer, perhaps some depth below the groundwater table. This water will be more dilute than the concentration of leachate in the vadose zone, and additional degradation may occur before people drink the groundwater. Also, the level of detail incorporated in the model and the modeling decisions is consistent with or a bit more conservative than the assessments for other routes of cumulative exposure to people.

The Office of Pesticide Programs (OPP) presented illustrative applications of three groundwater models --- Root Zone Water Quality Model (RZWQM), Pesticide Root Zone Model (PRZM), and Leaching Estimation and CHemistry-Pesticides model (LEACHP) --- that it proposes to use. The three groundwater models are very similar. However, in response to specific questions brought to the Panel, important considerations include the following:

- ***Management practices and pesticide application***. The models must be able to address the management practices and pesticide application practices (e.g., pesticide incorporation, management practices that impact hydrology) that are important to N-methyl carbamate and other pesticide movement to shallow groundwater. Urban uses of N-methyl carbamate pesticides could be important, but it is unclear whether the models proposed could address such uses.
- ***Degradation product formation.*** The model(s) should simulate formation and movement of degradation products if they are found to significantly contribute to risk.

Preferential flow. Macropore or preferential flow may be essential to accurately predict transport of N-methyl carbamate pesticides and other pesticides to shallow groundwater.

- ***Hydrology and rainfall characteristics.*** Most models predict the water balance accurately. Correctly characterizing rainfall amount and intensity may be especially important when applying RZWQM, because it will not simulate macropore flow unless the rainfall is sufficiently intense.
- ***Pesticide fate.*** Some of the more sophisticated pesticide fate routines in models should possibly be avoided. Both the pesticide degradation rate and the adsorption strength are highly variable between fields and may explain the poor performance of the models during evaluation efforts.
- ***Tile drainage.*** If N-methyl carbamate pesticides are applied to areas that are tile drained, the selected model should simulate tile drainage for accurate loading estimates to surface water. However, tile drainage consideration may not be necessary in the model for the purposes of the cumulative risk assessment, especially if this is a screening level tool.
- ***Model calibration, sensitivity, testing, and ease of use.*** Model calibration may be necessary. If so, the ability to easily calibrate the model will be important. Other important considerations include the transparency in the model's calculations, the data requirements, post-analysis of the model output, internal documentation of code, ability to modify code, ability to interface the code with third-party graphical user interfaces (GUI), and ability to link to statistical analysis packages and geographical information system (GIS) packages.

Surface water and groundwater behave quite differently. Generally, groundwater travels slowly and thus spatial averaging is a poor assumption. The pesticide concentration in each well depends on its immediate surroundings and cannot be represented by an average concentration. Because wells have screens allowing water into them, temporal averaging should be considered.

To account for persistence of pesticides in groundwater, the background levels could be estimated with models, and decline in residues over time (estimated based on long-term trends in monitoring) should be considered. The background document clearly identified that carbamate transformation can be pH specific. Also, the transformations of some metabolites may be affected by concentration and anaerobic conditions.

PANEL DELIBERATIONS AND RESPONSE TO THE CHARGE

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

Question 1

Selection of regional vulnerable ground water sites:

In both the organophosphate (OP) cumulative assessment and this carbamate cumulative assessment, OPP identified regional drinking water exposure sites for surface water sources of drinking water that would represent one of the more vulnerable surface watersheds in each region. The process, which was deemed a valid approach by the 2002 SAP (USEPA, 2002), identified those areas where areas of high combined cumulative pesticide use coincided with drinking water sources which were particularly vulnerable to runoff. This served as a regional screening assessment in that if the regional cumulative risk assessment finds that exposure in water is not a significant contributor to the overall exposure in that area, it will not be a significant contributor in other areas in the region.

For ground water sources of drinking water, OPP is proposing to use a similar approach (described in Section C.2. of the Drinking Water section of the case study).

National coverages of the vulnerability of aquifers to pesticide contamination are limited in availability, and generally refer to vulnerability of the overlying soils and surficial geology to leaching. Does the SAP believe that the sources on relative ground water vulnerability identified in the case study provide an adequate screening-level assessment of the potential for contamination of shallow aquifers to pesticide contamination? If these vulnerability assumptions are inadequate, what can be done to improve the approach?

Panel Response

In the oral presentation, EPA staff indicated that the vulnerability assessment conducted for groundwater to date was not as systematic as that done for surface water. The Panel agreed that a more systematic assessment would be desirable.

As indicated by the EPA, the national coverages of the vulnerability of aquifers to pesticide contamination and related data are limited. The vulnerability assessments that are available are limited in the types of data considered in their creation and in the scales of the data used (e.g., the nationwide DRASTIC assessments). There are some state and regional groundwater vulnerability assessments, such as those conducted in Indiana

(Cooper et al., 1997). The results of these regional assessments point out the importance of the scale of underlying data used in their creation (Cooper et al., 1997). The DRASTIC groundwater vulnerability estimates for Indiana created with 1:250,000 scale GIS data resulted in significantly different estimates of vulnerability than those created using courser nationally-derived data. Further, Cooper et al. (1997) demonstrated the value of testing the vulnerability maps with observed groundwater quality data. Therefore, there are challenges in identification of groundwater vulnerability nationally.

To assess groundwater vulnerability, EPA used the method of Nolan et al. (2002) and three criteria to identify the most vulnerable regional drinking water exposure sites. The first two criteria were areas where carbamate pesticides are applied and areas where groundwater is used for drinking. The third criterion was to identify the areas most vulnerable to groundwater contamination.

The vulnerability assessment approach does not address off-label and improper uses of pesticides. Agricultural pesticide applications might generally show fairly high adherence to prescribed uses and application rates, but residential applications might show considerably less fidelity to label uses. One of the pesticides being considered has residential applications. It is conceivable that EPA's presumption that usage follows label guidance might seriously underestimate vulnerabilities in some suburban areas.

The Nolan et al. (2002) method to assess groundwater vulnerability was developed for nitrate, which may transport through the vadose zone or an aquifer differently than pesticides. For example, vulnerable areas for pesticides may largely be a function of soil organic carbon and macropore flow, which may not be as large a factor for nitrate transport to shallow groundwater. The method used by Nolan et al. (2002) also considers human population which may be unique to groundwater vulnerability to nitrate. The model coefficient associated with the depth to groundwater also may have the wrong sign on it in this model if it is used for N-methyl carbamate pesticides. Therefore, areas may be vulnerable to pesticide leaching to shallow groundwater that are not identified by the Nolan et al. (2002) approach; the vulnerability GIS coverage that has been used may not be the most appropriate approach for identifying vulnerability of groundwater to carbamates and other pesticides.

If the Nolan et al. (2002) vulnerability approach is to be used for identifying areas vulnerable to pesticides reaching shallow groundwater, it should be compared to existing data on known contaminated aquifers, both current and past, to assess its quality. The vulnerability model should agree well with field observations. Where there is disagreement, the EPA should investigate to understand why they differ.

If one examines the next step proposed by EPA in modeling groundwater vulnerability, there is possibly a disconnect between the proposed groundwater vulnerability approach and the modeling approach. For example, the groundwater vulnerability screening GIS map was created considering confined as well as unconfined aquifers. However, the modeling approach only considers very shallow unconfined

aquifers. The use of STATSGO (a nationwide soil GIS dataset) may be a more appropriate mechanism for vulnerability screening. The use of STATSGO combined with the models would allow the most vulnerable settings to be identified. Also, using STATSGO would be consistent with the proposed approach for modeling pesticide exposure in shallow groundwater, since soil properties play an important role in parameterizing the models, while aquifer properties are not considered.

Question 2

Use of leaching models for ground-water exposure assessments:

The Agency is basing its ground-water exposure assessment on private rural wells drawing its drinking water from an unconfined aquifer. The estimated exposure in drinking water from these wells is based on the concentration estimated at the top of this aquifer.

The three models that the Agency is considering for use in the cumulative assessment and in refined ground water exposure in aggregate (individual chemical) assessments – LEACHP, PRZM, and RZWQM – are leaching models that predict pesticide concentrations in water at some depth below the surface. As a result, the Agency is using estimated pesticide concentrations in the vadose-zone to represent concentrations in shallow ground water. As such, estimates would represent potential drinking water exposure from wells drawing from shallow, unconfined aquifers.

Is this approach a reasonable, health-protective approach for use in both cumulative and aggregate drinking water exposure assessments? If this approach produces an exceedance of essentially safe exposure levels, in what manner could a better estimate of exposure to pesticides in water be derived from existing data and modeling approaches?

Panel Response

For most reasonable scenarios, the proposed approach appears to be conservative and health protective. In most cases, the screens of wells are some depth below the groundwater table, allowing time for additional degradation before people drink the groundwater. The level of detail in the assessment and the assumptions are consistent or a bit more conservative than the assumptions made for exposure from other pathways. However, there are several cases where caution is needed and the approach may not be health protective.

If the pesticide is not entirely in the dissolved phase (as assumed in these models) and greatly in excess of normal agricultural use rate for example due to either a large accidental spill, an unusual surface runoff event or cleaning of tanks close to the well, the pesticide might reach a shallow unconfined aquifer in concentrations far in excess of model predictions.

The use of an aquifer model combined with the estimates of pesticides reaching shallow groundwater without calibration for local conditions is cautioned. Calibrations with experimental data would likely provide better estimates of actual pesticide levels reaching drinking water. One of the challenges will be adequately representing the aquifers because data will be limited and site specific conditions could be essential for accurately estimating pesticide concentration.

Data used for model validation might not represent the "real" concentration leaching into the aquifer. Caution should be used when estimating pesticide transport to shallow groundwater based on suction lysimeter and soil sample data. If macropores are present in the soil, pesticides can move from the soil surface to shallow groundwater with little evidence of this movement in the unsaturated zone (Malone et al., 2000). The best locations for pesticide samplers are in the capillary fringe of the groundwater or above textural interface layers. Wick pan samplers are preferred above gravity pan samplers. Tile lines give the overall best integrated sample (Boll et al., 1992; Shalit et al., 1995; Boll et al., 1997).

The Agency's tiered approach is supported. Tiered approach is commonly used in risk assessment as a good use of limited resources by first looking into the worst case scenario using conservative assumptions. This initial tier of analysis is less time consuming and data intensive. A refining tier of analysis for a more realistic scenario is performed if the worst case scenario is deemed unrealistic and shows exceedance of safe exposure level.

The Panel raised several questions during the Agency presentation concerning how the models will be evaluated for the closeness of their prediction to the realistic situations. The Panel recommended a more rigorous review of the chosen models for N-methyl carbamate pesticides, specifically for their use in cumulative risk assessment. When the worst case scenario exposure prediction exceeds the Agency pre-determined safe level, and that prediction is deemed a gross over-estimation, the input parameters, including the default concordance parameters (e.g., in RZWQM), should be reviewed specific for the cumulative assessment group of the N-methyl carbamate chemicals. Attempts should be made to compare model results to available monitoring data.

Question 3

Addressing the cumulative risk assessment needs:

The Agency has considered analyses of the capabilities of three ground water models for use in the carbamate cumulative exposure assessment and in individual chemical (aggregate) assessments. The major areas of evaluation – hydrology, management practices, pesticide transport processes, and ease of use of the model – are described in the background document *Drinking Water Exposure Assessment: Ground Water Model Evaluation*. For the cumulative assessment, OPP will compare results of all three models with each other and with available monitoring. The ultimate evaluation goal for the models is how well each meets the requirement of providing reasonable, health-protective estimates of pesticide residues for use in aggregate and cumulative drinking water exposure assessments.

In the SAP's estimation, how well do the three ground water models OPP proposes to use – RZWQM, PRZM, and LEACHP – compare in addressing hydrology? Macropore flow? Rainfall characteristics? Management practices? Pesticide fate and transport? Formation and movement of transformation products? Can the panel recommend other criteria that should be considered in evaluating the effectiveness of these models for estimating drinking water exposure for regulatory purposes?

Panel Response

The Panel's deliberations on each of the parameters are discussed in the following sections, along with recommendations for additional parameters that should be considered in the models.

Management Practices and Pesticide Application

It is necessary to identify which processes are most important for accurately modeling N-methyl carbamate pesticides. For example, incorporation of applied pesticides into the soil should probably be considered because this is a recommended agricultural practice in conjunction with the application of some N-methyl carbamate pesticides. Another important factor to consider is the amount of pesticide applied. The quantity used can usually be estimated based on sales records and on label instructions. The management practices (such as Integrated Pest Management) can greatly affect the amount of pesticide applied and is, therefore, an important consideration in modeling for this purpose.

Management practices usually have little effect on the water balance with the exception of those that include a winter cover crop (Walter et al., 1979). In these cases,

there is less percolation during the early spring. Therefore, management practices that include cover crops should be considered in the simulation of the hydrology.

Another concern is that agricultural uses may not be the only source of N-methyl carbamate pesticides, especially in suburban settings, yet these settings may still be drawing water from surficial aquifer wells. Non-agricultural use may be small in its overall contribution, but such uses may be important and could even dominate in selected situations. It may be important for the models to allow for consideration of non-agricultural sources .

Degradation Product Formation

N-methyl carbamate pesticides in the environment are relatively labile and readily undergo hydrolysis giving rise to the amino acid and the alcohol form of the leaving group, e.g, 1-naphthol. The leaving group component may itself be somewhat toxic. There are likely to be other products as well. These should be traceable as part of the modeling. Therefore, the model should simulate formation and movement of degradation products. For example, metabolites of aldicarb such as aldicarb sulfoxide and aldicarb sulfone may be more important than the parent compound (Kraft and Helmke, 1991; Smelt et al., 1995). This component must be thoroughly tested because there have been mass balance problems with transformation products in model testing.

Preferential Flow

Macropore flow is likely essential to accurately model carbamate transport to shallow groundwater. Even on loamy sand (Ritter et al., 1996), macropore flow can be an important process. Specifically, N-methyl carbamate pesticides such as carbofuran, can be transported to shallow groundwater via preferential flow (Isensee et al., 1990; Kladivko et al, 1999). Model codes including a description of preferential flow processes required less calibration efforts to meet the FOCUS model performance criteria on a sandy loam soil than those without such description (Thorsen et al., 1998).

The methods to simulate preferential flow are vastly different between the models. This should not be a surprise since the theory for simulating preferential flow is still under development and not understood well. Of interest here, most aquifers are overlain by sandy soils. Also, most N-methyl carbamate pesticides are applied on these sandy soils. On these types of soils, unstable fingered flow and fingered flow are important. Although RZWQM is the most sophisticated model in simulating preferential flow, it does not consider fingered flow (see Figure 2). To model fingered flow, the maximum intensity of a storm might determine the number of fingered flow paths (Selker et al., 1996). RZWQM clearly stands out in its sophistication to model macropore flow and is likely superior above the other models, but none of the models address the unstable fingered flow phenomena in sandy soils. At the present time, preferential flow including funnel flow in sandy soils could be included by model developers, since the initial theory has been developed and could be tested in field situations (Steenhuis et al., 2001; Kim et

al., 2005). EPA should be looking for these types of models for inclusion in the risk assessment.

Model users should be aware that it is simplistic to divide up the medium into one preferential flow region and a matrix component (Steenhuis et al., 1990). In reality, there is a continuum between the fastest flow path and the slowest flow paths. The experimental and theoretical work of Kung et al. (2000, 2005) and Gish et al. (2004) confirmed this concept.

In structured soil (e.g., silty, loamy, and clayey soil), macropore characteristics for RZWQM are easily parameterized. Little is known for sandy soils though and work is needed to parameterize RZWQM for sandy soils.

Hydrology and Rainfall Characteristics

Most models predict the water balance quite accurately, including the ones proposed for use by the Agency. Accurate determination of daily rainfall amount is essential and may be one of the most crucial issues concerning accurate simulation of water balance (Wagenet and Hutson, 1996). The rainfall rate is also important, especially if macropore flow is identified as an important process to include in the model. Skopp et al. (1981) varied how water was added to a column, which affected the rate of preferential flow, and consequently the breakthrough curves were completely different (also see Skopp and Gardner, 1992). Steenhuis et al. (1990) developed a preferential flow model whereby flow was directly related to the rate of application.

Correctly characterizing rainfall intensity may be especially important when applying RZWQM, because it will not simulate macropore flow unless the rainfall is sufficiently intense. That is, 24 hour or hourly rainfall may not be sufficiently intense to result in macropore flow with RZWQM. Pesticides may move to shallow groundwater by macropore flow in the sandy soils if heavy rainfall occurs shortly after pesticide application (Ritter et al., 1996). Subsequent rainfall events may be less important because the first storm after application can move solutes into the soil matrix, thereby reducing the potential for transport in macropores (Shipitalo et al., 1990). Therefore, efforts should be applied to develop breakpoint rainfall input files for the first few storms after pesticide application.

Pesticide Fate

Some of the more sophisticated pesticide fate routines in models should possibly be avoided. For example, the RZWQM irreversible binding routine did not provide accurate simulations of metribuzin in percolate, suggesting this concept may need to be explored further (Malone et al., 2004). Also, estimating parameters for kinetic sorption may be difficult. However, if most pesticide is transported through percolate shortly after application, irreversible binding and kinetic sorption may be less important than short-term half-life and equilibrium sorption.

Both the degradation rate and the adsorption strength are highly variable between fields. The variability of degradation rates between sites is one of the main problems for simulating the pesticide concentration in the aquifers with confidence, and this might be the main reason that the pesticide concentration in the USGS model evaluations (included as background documents) are so poorly simulated. Pivetz and Steenhuis (1995) and Pivetz et al. (1996) concluded that the degradation rate may vary significantly in time. Reasonable estimates can only be obtained if prior experiments have been carried out at the site. Because simple parameters such as the linear adsorption model and the first order degradation rate model are difficult to obtain, it will likely be too difficult to apply more complicated formulations of these relationships nationally.

Enhanced or accelerated degradation has been reported in N-methyl carbamate pesticides such as carbofuran (Cogger et al., 1998; Getzin and Shanks, 1990) and aldicarb (Suett and Jukes, 1988; Smelt et al., 1987). Enhanced degradation, defined as an increase in pesticide degradation with each application, may lead to higher or more frequent applications because insecticidal efficacy will be reduced.

Tile Drainage

If N-methyl carbamate pesticides are applied to areas that are tile drained, then the selected model should simulate tile drainage for accurate loading estimates to surface water. Recent research suggests that the pesticide isoxaflutole and its metabolite RPA 202248 are transported to the edge of the field through subsurface drains at about 1 m at rates similar to runoff (personal communications, Garey A. Fox, Assistant Professor, University of Mississippi). However, tile drainage may not be necessary in the model for the purposes of the cumulative risk assessment, especially if this is a screening level tool.

Model Calibration, Sensitivity, Testing, and Ease of Use

The use of any model of pesticide contamination from agricultural use is likely to give only an estimation of the contamination found in the surficial aquifer. The use of the PRZM system is well established with respect to EPA usage. This suggests that its use in this context may be supported above and beyond other ostensibly better, but perhaps more complicated and difficult to use models. Further, parameterization of the other models may be difficult since data may not be available for scenarios of interest.

Experience with LeachN, the nitrogen movement implementation of the Leach model, demonstrates the difficulty of getting these models to predict observed levels of materials, even with intensive measurement of soil characteristics, initial conditions and control of soil nutrient additions. EPA may wish to create their own tool for this purpose that they understand and can distribute and control.

The sensitivity of the models to key parameters that are likely to be changed in application of the models should be well understood. Also, the sensitivity of key

parameters relative to the group of models being considered would provide insight into the most applicable model or models for various situations. Perhaps probabilistic or nonlinear analysis of the model output might be the best approach to analyzing the parameter interdependencies.

The ease of model calibration and the ability to calibrate each model should be considered. Models often must be calibrated to perform adequately. Therefore, calibrations will be important considerations when using the groundwater models to estimate pesticides reaching shallow groundwater.

The ability to automate the use of each model so that it can readily utilize data from databases and GIS to facilitate automated model runs will likely be important for future applications of the models. Models that require the use of a model-specific interface may not work very well or at all for such purposes.

Other criteria to take into consideration are: transparency in the model's calculations, the data requirements, post-analysis of the model output, internal documentation of code, ability to modify code, ability to interface the code with third-party GUIs, and ability to link to statistical analysis packages and GIS packages. In analyzing a full cumulative risk assessment, users and reviewers must have models that are simple to evaluate during all portions of the risk assessment.

For model testing, as complete a dataset as possible should be used. Some of the discrepancies between model and data may be resolved by averaging samples taken over a broader area. The models have much in common, so discrepancies between models must be explicable in each case in terms of known differences in model assumptions. The dataset should include soil samples, runoff, and percolate, because neglecting one of these dissipation pathways may provide an incomplete picture of model performance (Malone et al., 1999). When preferential flow occurs, soil samples and suction lysimeters may not adequately characterize chemical movement to shallow groundwater, and percolate samples may be necessary (Malone et al., 2000).

The effects of water purification processes on carbamate levels also should potentially be considered. If such processes routinely eliminate most N-methyl carbamate pesticides from drinking water, this modeling would seem to be superfluous except for those drinking untreated water. Estimation of these purification effects would thus seem critical, particularly new filtration technologies and the effects of chlorination and chloramines disinfectant technology. At the current time, the potential for significant overestimation of carbamate levels from these models seems to exist. These factors, together with the low level of exposure to N-methyl carbamate pesticides from drinking

water relative to other exposure sources, suggest this modeling may be of value for only a subset of the public.

Question 4

Estimating cumulative carbamate exposures in ground water:

Available monitoring data, primarily from the USGS NAWQA program, indicate that more than one carbamate in the cumulative action group may occur together in ground water (see the drinking water exposure section of the case study). Co-occurrence in ground water results when more than one carbamate is used at different times on the same crop, on different crops in rotation on the same fields, or on different crops grown on adjacent fields.

For surface water sources of drinking water, OPP adjusted estimated pesticide concentrations in the modeled reservoir by percent crop area and percent crop treated in the watershed to reflect the dilution of untreated areas on the total pesticide load reaching the reservoir (see the description of the surface water exposure assessment in the analysis methods of the Drinking Water section of the case study).

Given that less mixing of water from different fields is expected in shallow aquifers, should OPP use similar adjustments to ground water concentrations estimated from the leaching models? If not, what recommendations does the Panel have to account for the potential contributions from different fields treated with carbamates?

Panel Response

Surface water and groundwater behave quite differently. Surface water travels relatively rapidly from any location in the watershed to the outlet, and the Agency's adjustment for surface water edge of field concentration by percent crop area and percent crop treated in the watershed is reasonable. Thus, the assumption is made that pesticide concentration in surface drinking water can be represented by the average spatial concentration of the runoff.

The assumption of taking spatial averages in the watershed is not valid for groundwater. Generally, groundwater travels slowly, usually less than 100 m/year. The pesticide concentration in each well depends on its surroundings and cannot be represented by an average concentration. Thus, the people drinking water from the wells are exposed to different concentrations. This is important for the risk analysis where highest concentrations are often of interest.

The different concentrations for different wells are schematically depicted in Figure 1. Pesticides are applied to a field as indicated. The pesticides travel in a distinct path to the river through the ground water. Figure 1 illustrates that only the shallow well downstream from the field contains pesticide, while the other wells are pesticide free. The drinking water well in Figure 1 samples water from a larger volume and will be affected by more than one field.

The screens of shallow wells are typically a minimum of 4 feet long and sample pesticides over these depths. This means that the daily input at the top of the aquifer is averaged over a certain time period. To help describe this better, Figure 2 depicts a close up view of the situation near the pesticide field from Figure 1. Only the water that fell on top of the field is shown in the groundwater depicted. It can be seen that the oldest water is the deepest, while the water which just arrived is near the surface of the aquifer. While the water is moving downward, it is, at the same time, moving sideways toward the river. The line in the groundwater in Figure 2 is a streamline. In this case water that falls at the upper edge of the field follows this path.

Although spatial averages cannot be taken for wells, temporal averaging can be done. It is not difficult to calculate a reasonable time period over which the pesticide concentrations from the vadose pesticide model should be averaged. The depth of "recharge layer" per year is simply the recharge per year, r, divided by the saturated moisture content, θ_s. The depth of water that represents water reaching the aquifer each year, D, can be calculated taking into account the adsorption partition coefficient, k, and the density of the soil, ρ, as:

$$D = \frac{r}{\theta_s + \rho k}$$

The time, T, over which the pesticides arriving to the groundwater should be averaged can then be simply written as:

$$T = \frac{S}{D}$$

where S is the screen length.

The location of stream lines can be found with the method presented by Gelhar and Williams (1974). This method was used in a model called MOUSE (Steenhuis et al., 1987). The model had many similarities to the models currently used by EPA (but without preferential flow paths). The details of simulating streamlines are given in these publications. The method consists of two simple equations; one that determines the displacement in the horizontal direction and the other in the vertical direction.

The two surface water modifiers, percentage of crop area (PCA) and percentage of crop treated (PCT), can reasonably be used in ground water prediction also. However, the persistence of N-methyl carbamate pesticides in ground water would require considerations of multiple years of pesticide use. This is different from focusing the surface water scenarios on spikes associated with runoff and its relatively fast decline to low or nondetectable levels after runoff events. Considerations also can be given for time of pesticide use and ground water movement. In terms of PCA, the pattern of change in land use over time should be accounted for (e.g., to consider pesticide use at sites that were once agricultural). When available, multiple years of data on PCT also may be useful to modify the default assumption of 100% PCT.

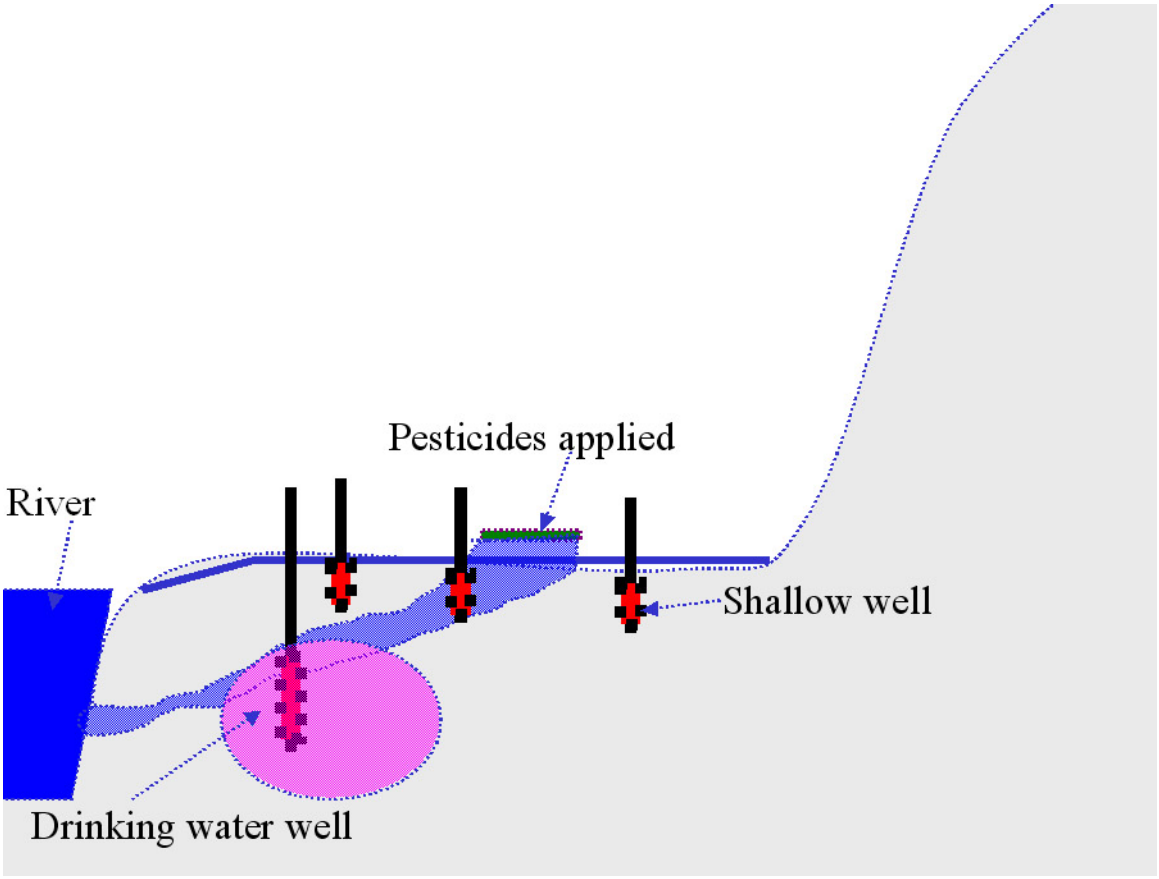

Figure 1. Schematic of flow paths in a valley with a river. Only the shallow well downstream of the field is affected by the field in which the pesticides are applied.

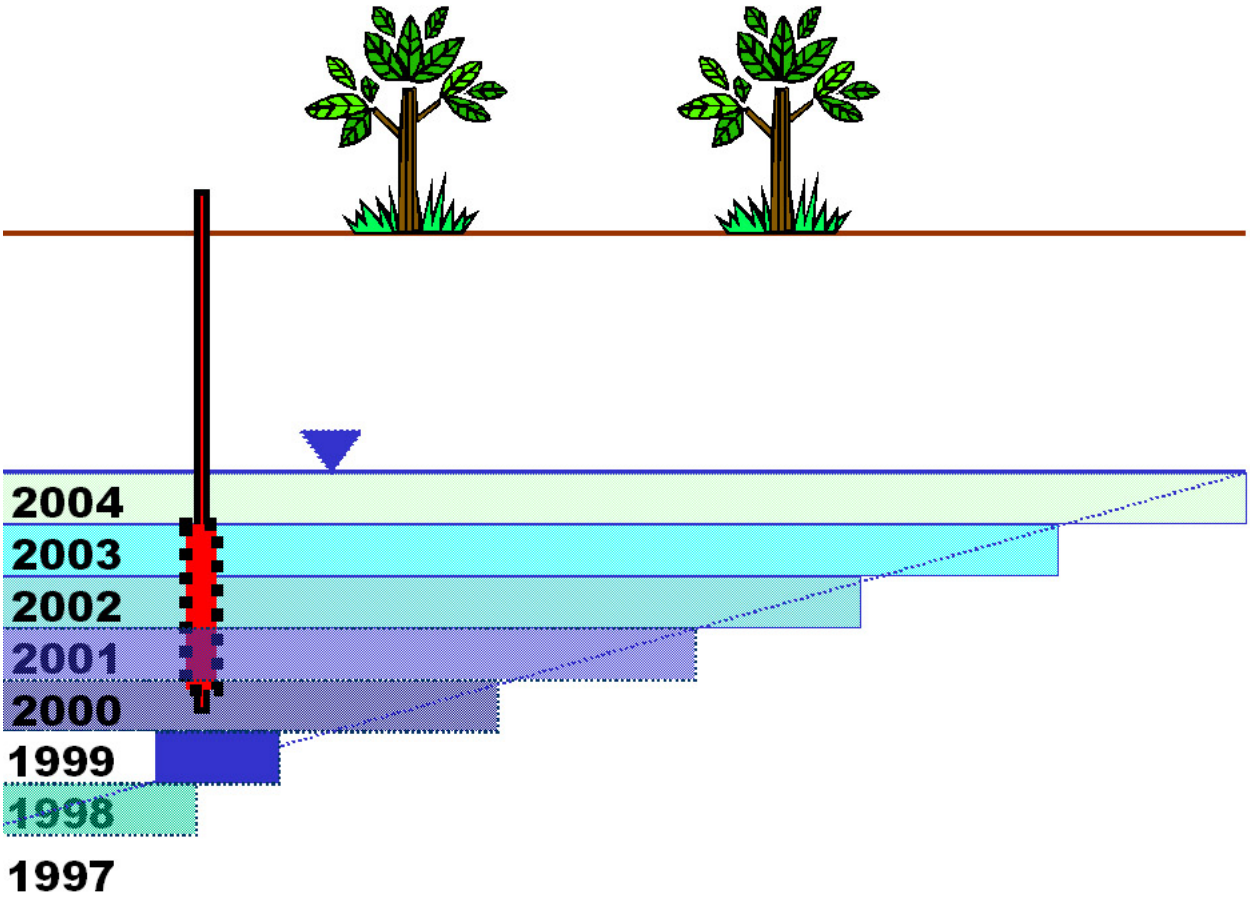

Figure 2. Simple schematic of ground water movement under a pesticide movement.

Question 5

Several carbamates are known to persist for long periods of time after they reach ground water, particularly slightly acidic to acidic ground water (see the drinking water exposure section of the case study). This differs from estimates of exposures in surface water, where pesticide concentrations tend to occur as spikes associated with runoff and decline to low or nondetectable levels after runoff events. Pesticide concentrations in acidic ground water are slower to respond to changes in use patterns and mitigation actions than would be expected in surface water.

What recommendations does the Panel have for addressing carbamate persistence in ground water in order to provide a reasonable, health-protective estimate of residue levels in shallow ground water sources of drinking water?

Panel Response

The Agency is considering three possible approaches:

1) at one extreme, assume no background residues (drinking water exposures would reflect only what is estimated by modeling), i.e., all residues in groundwater are “fresh”;

2) at the other extreme, assume a baseline background concentration (based on available monitoring), with model estimates as additions and no decline;

3) in between, include the background levels with model estimates, but provide an estimate of decline in residues over time (estimate based on long-term trends in monitoring)

The EPA's 3rd suggested approach from the "background document" for Session 3 appears sensible. However, it is unclear how EPA would obtain the "long-term trends in monitoring".

The background document clearly identified that carbamate transformation can be pH specific. Also, the transformations of aldicarb metabolites are concentration specific (Smelt et al., 1995). That is, higher concentrations degrade more slowly than lower concentrations. At least two peer-reviewed sources that studied aldicarb transformation used initial concentrations greater than 150 ug/L. Some studies of aldicarb degradation suggest that anaerobic conditions increase degradation (Smelt et al., 1995; Kraft and Helmke, 1992).

The question of persistence has underlying implications on the travel distance of the pesticide through a continuous aquifer. There are a few references on estimating characteristic travel distance of a persistent pesticide based on its chemical properties (Bennet et al., 1999, 1998).

REFERENCES

Bennett, D.H., McKone, T.E., Kastenberg, W.E., Matthies, M. (1998) General Formulation of Characteristic Travel Distance for Semi-Volatile Organic Chemicals in a Multi-media Environment, *Environ. Sci. Technol.* 32: 4023-4030.

Bennett, D.H., Kastenberg, W.E., McKone, T.E. (1999) General Formulation of Characteristic Time for Persistent Chemicals in a Multimedia Environment, *Environ. Sci. Technol.* 33: 503-509.

Boll, J., Steenhuis, T.S., and Selker, J.S. (1992) Fiberglass Wicks for Sampling Water and Solutes in the Vadose Zone. Soil Sci. Soc. Am. J. 56:701-707.

Boll, J., Selker, J.S., Shalit, G., and Steenhuis, T.S. (1997) Frequency Distribution of Water and Solute Transport Properties Derived from Pan Sampler Data. Water Resour. Res. 33:2655-2664.

Cogger CG, Stark J.D., Bristow P.R., Getzin L.W., Montgomery M. (1998) Transport and Persistence of Pesticides in Alluvial Soils: Ii. Carbofuran. *J. Environ. Qual.* 27(3):551-556.

Cooper, B., Navulur, K., Engel, B., and Hahn, L. (1997) Groundwater vulnerability evaluation to pesticide and nitrate pollution on a regional scale using GIS. http://pasture.ecn.purdue.edu/~aggrass/groundwater

Gelhar L.D. and Williams, J.L. (1974) Ground water quality modeling. Ground Water 12:399-408.

Getzin L.W. and Shanks, C.H. (1990) Enhanced Degradation Of Carbofuran In Pacific-Northwest Soils. *J. Environ. Sci. Health Part B.* 25:433-446

Gish T.J., Kung, K-J.S., Perry, D.C., Posner, J., Bubenzer, G., Helling, C.S., Kladivko, E.J. and Steenhuis, T.S. (2004) Impact of preferential flow at varying irrigation rates by quantifying mass fluxes. *J. Environ. Qual.* 33(3): 1033-1040

Isensee, A.R., Nash, R.G., Helling, C.S. (1990) Effect Of Conventional Vs No-Tillage On Pesticide Leaching To Shallow Groundwater. *J. Environ. Qual.* 19(3):434-440.

Kim, Y.-J., Darnault, C.J.G., Bailey, N.O., Parlange, J.-Y. and Steenhuis, T.S. (2005) Equation for Describing Solute Transport in Field Soils with Preferential Flow Paths. *Soil Sci. Soc. Am. J.* 69. <in press>

Kladivko, E.J., Grochulska, J., Turco, R.F., Van Scoyoc, G.E., and Eigel, J.D. (1999) Pesticide and Nitrate Transport into Subsurface Tile Drains of Different Spacings. *J. Environ. Qual.* 28(3):997-1004.

Kraft, G.J, and Helmke, P.A. (1991) Aldicarb Residue Degradation Rates At 3 Depths of a Wisconsin Central Sand Plain Aquifer. *Pest. Sci.* 33(1):47-55.

Kraft, G.J. and Helmke, P.A. (1992) Dependence Of Aldicarb Residue Degradation Rates On Groundwater Chemistry In The Wisconsin Central Sands. *J. Environ. Qual.* 21(3):368-372.

Kung, K.-J.S., Kladivko, E.J., Gish, T.J., Steenhuis, T.S., Bubenzer, G. and Helling, C.S. (2000) Quantifying Preferential Flow by Breakthrough of Sequentially Applied Tracers: Silt Loam Soil. *Soil Sci. Soc. Am. J.* 64:1296-1304.

Kung, K.-J. S., Hanke, M., Helling, C.S., Kladivko, E.J., Gish, T.J., Steenhuis, T.S. and Jaynes, D.B. (2005) Quantifying Pore Size Spectrum of Macropore-Type Preferential Pathways. *Soil Sci. Soc. Am. J.* (in press)

Malone, R. W., Warner, R.C., Workman, S.R. and Byers, M.E. (1999) Modeling Surface And Subsurface Pesticide Transport Under Three Field Conditions Using PRZM-3 And GLEAMS *Transactions of the ASAE*. 42(5):1275-1288.

Malone, R. W., Shipitalo, M.J., Douglass, L.W., Owens, L.B., Nelsen, T.C., Warner, R.C., and Byers, M.E. (2000) Assessing Herbicide Movement Using Soil Samples Versus Percolate Samples *Transactions of the ASAE*. 43(2):343-348.

Malone, R.W., Ma, L.W., Wauchope R.D., Ahuja, L.R., Rojas, K.W., Ma, Q.L., Warner, R. and Byers, M. (2004) Modeling hydrology, metribuzin degradation and metribuzin transport in macroporous tilled and no-till silt loam soil using RZWQM. *Pest Manage. Sci.* 60(3):253-266.

Nolan, B.T., Hitt, K.J. and Ruddy, B.C. (2002) Map of the probability that nitrate exceeds 4 mg/L in shallow ground waters of the United States, based on the LR model. Accessed April 13, 2004 at URL:http://water.usgs.gov/lookup/getspatial?gwrisk

Pivetz, B.E. and Steenhuis, T.S. (1995) Soil Matrix and Macropore Biodegradation of 2,4-D. *J. Env. Qual*. 24:564-570.

Pivetz, B.E., Kelsey, J.W., Steenhuis, T.S. and Alexander, M. (1996) A Procedure to Calculate Biodegradation During Preferential Flow Through Heterogeneous Soil Columns. *Soil Sci. Soc. Am. J.* 60:381-388.

Ritter, W.F., Chirnside, A.E.M. and Scarborough, R.W. (1996) Movement And Degradation Of Triazines, Alachlor, And Metolachlor In Sandy Soils. *J. Environ. Sci.Health Part A.* 31(10):2699-2721.

Selker, J.S., Steenhuis, T.S. and Parlange, J.-Y. (1996) An Engineering Approach to Fingered Vadose Pollutant Transport. In: Fingered Flow in Unsaturated Soil: From Nature to Model, T.S. Steenhuis, C.J. Ritsema, and L.W. Dekker, Eds. Special Issue of *Geoderma* 70:197-206.

Shalit, G., Steenhuis, T.S., Hakvoort, H.M., Boll, J., Geohring, J.D. and van Es, H.M. (1995) Subsurface Drainage Water Quality from a Structured Soil. ASCE J. Irrig. Drain.121:239-247.

Shipitalo, M.J., Edwards, W.M., Dick, W.A. and Owens, L.B. (1990). Initial Storm Effects On Macropore Transport Of Surface-Applied Chemicals In No-Till Soil. *Soil Sci. Soc. Am. J.* 54(6):1530-1536.

Skopp, J., Gardner, W.R. and Tyler, C.J. (1981) Solute movement in structured soils - 2-region model with small interaction *Soil Sci. Soc. Am. J.* 45(5):837-842

Skopp, J. and Gardner, W.R. (1992) Miscible displacement - an interacting flow region model. *Soil Sci. Soc. Am. J.* 56 (6): 1680-1686

Smelt, J.H., Crum, S.J.H., Teunissen, W. and Leistra, M. (1987) Accelerated degradation of aldicarb, oxamyl and ethoprophos after repeated soil treatments. *Crop Protection* 6:295-303.

Smelt, J.H., Vandepeppelgroen, A.E. and Leistra, M. (1995) Transformation Of Aldicarb Sulfoxide And Aldicarb Sulfone In 4 Water-Saturated Sandy Subsoils. *Pest. Sci.* 44 (4): 323-334.

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

Steenhuis, T.S., Parlange, J.-Y. and Andreini, M.S. (1990) A Numerical Model for Preferential Solute Movement in Structured Soils. *Geoderma* 46:193-208.

Steenhuis, T.S., Kim, Y.-J., Parlange, J.-Y., Akhtar, M.S., Richards, B.K., Dung, K.-J.S., Gish, T.J., Dekker, L.W., Ritsema, C.J., and Aburime, S.O. (2001) An Equation for Describing Solute Transport in Field Soils with Preferential Flow Paths. Preferential Flow, Water Movements and Chemical Transport in the Environment. D.D. Bosch and K.W. King, Editors. ASAE Proceedings. Second International Symposiums. Honolulu, HI. 137-140.

Suett, D.L. and Jukes, A.A. (1988) Accelerated degradation of aldicarb and its oxidation

precuts in previously treated soils. *Crop Protection* 7:147-152.

Thorsen, M., Jorgensen, P.R., Felding, G., Jacobsen, O.H., Spliid, N.H. and Refsgaard, J.C. (1998). Evaluation Of A Stepwise Procedure For Comparative Validation Of Pesticide Leaching Models. *J. Environ. Qual.*. 27(5):1183-1193

Wagenet, R.J. and Hutson, J.L. (1996) Scale-dependency of solute transport modeling/GIS applications *J. Environ. Qual.* 25 (3): 499-510

Walter, M.F., Steenhuis, T.S. and Haith, D.A. (1979) Nonpoint Source Pollution Control by Soil and Water Conservation Practices. *Trans. ASAE* 22:834-840.

SAP Minutes No. 2005-01

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

N-METHYL CARBAMATE CUMULATIVE RISK ASSESSMENT: PILOT CUMULATIVE ANALYSIS

SESSION 4: N-METHYL CARBAMATE EXPOSURE ASSESSMENT: A PILOT CASE STUDY

February 18, 2005
FIFRA Scientific Advisory Panel Meeting held at the Holiday Inn National Airport Arlington, VA

Mr. Joseph E. Bailey
Designated Federal Official
FIFRA Scientific Advisory Panel
Date: April 15, 2005

Steven G. Heeringa, Ph.D.
FIFRA SAP Session Chair
FIFRA Scientific Advisory Panel
Date: April 15, 2005

Federal Insecticide, Fungicide and Rodenticide Act
Scientific Advisory Panel Meeting

February 18, 2005

N-methyl Carbamate Cumulative Risk Assessment: Pilot Cumulative Analysis

Session 4: N-methyl Carbamate Exposure Assessment: A Pilot Case Study

PARTICIPANTS

FIFRA SAP Session Chair

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

Designated Federal Official

Joseph E. Bailey, FIFRA Scientific Advisory Panel Staff, Office of Science Coordination and Policy, EPA

FIFRA Scientific Advisory Panel Members

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

Gary E. Isom, Ph.D., Professor of Toxicology, School of Pharmacy & Pharmacal Sciences, Purdue University, West Lafayette, IN

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

FQPA Science Review Board Members:

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

William Brimijoin, Ph.D.,Chair, Pharmacology, Mayo Clinic and Medical School, Rochester, MN

George B. Corcoran, Ph.D., Professor & Chairman, Department of Pharmaceutical Sciences, Eugene Applebaum College of Pharmacy & Health Sciences, Wayne State University, Detroit, MI

Lutz Edler, Ph.D., Head, Biostatistics Unit C060, German Cancer Research Center, Heidelberg, Germany

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

Scott Ferson, Ph.D., Senior Scientist, Applied Biomathematics, Setauket, NY

Lawrence J. Fischer, Ph.D., Director, Center for Integrative Toxicology, Michigan State University, East Lansing, MI

Natalie Freeman, Ph.D., Associate Professor, Department of Physiological Sciences, College of Veterinary Medicine, University of Florida, Gainesville, FL

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

James P. Kehrer, Ph.D., Director, Center for Molecular & Cellular Toxicology, College of Pharmacy, The University of Texas at Austin, Austin, TX

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

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

David MacIntosh, Sc.D., Senior Associate, Environmental Health & Engineering, Inc., Newton, MA

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

Christopher J. Portier, Ph.D., Director, Environmental Toxicology Program, National Institute of Environmental Health Sciences, Research Triangle Park, NC

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

Michael D. Wheeler, Ph.D., Assistant Professor, Departments of Pharmacology & Medicine, University of North Carolina, Skipper Bowles Center for Alcohol Studies, Chapel Hill, NC

INTRODUCTION

In Session 4 of this meeting, the FIFRA SAP met to consider and review the N-methyl carbamate pesticide cumulative risk assessment: pilot cumulative analysis, specifically issues related to the N-methyl carbamate exposure assessment: a pilot case study. OPP solicited comment from the SAP on a case study which uses the Relative Potency Factor (RPF) approach to perform a cumulative exposure assessment for a group of 10 N-methyl carbamate pesticides that have been previously determined to represent a common assessment group based on the chemicals' similar mechanism of toxicity. For Session 4, Dr. Anna Lowit and Mr. David Miller (Health Effects Division, Office of Pesticide Programs) provided an introduction and background on the pilot case study. Mr. David Hrdy, Mr. Jeff Evans, Dr. Steve Nako, and Mr. David Miller (Health Effects Division, Office of Pesticide Programs) provided a detailed presentation on the pilot exposure assessment case study, including discussions on dietary assessment, residential exposure inputs, residential use data and aggregate exposure, cumulative exposure assessment and model comparison. Opening remarks for Session 4 were provided by Dr. Tina Levine (Acting Director, Health Effects Division, Office of Pesticide Programs).

SUMMARY OF PANEL DISCUSSION AND RECOMMENDATIONS

The Panel members agree that the overall case study is a thoughtful and useful analysis. EPA has gained much knowledge from development of the organophosphate (OP) cumulative risk assessment and the Agency is to be commended on the carefully developed exploratory analysis for the N-methyl carbamates, given the varying quality of the datasets that were available. The level of detail is adequate for a case study, but additional effort is needed in the full assessment, especially with regard to non-food scenarios. It was felt that uncertainty analyses will help to identify the datasets that should be examined most closely.

The default assumption of eating events close together in time, in the context of the pilot case study for the N-methyl carbamate exposure assessment, is a reasonable approach. Allowing for recovery seems difficult given the unknown durations that would have to be assigned to defensibly address this issue. The extent to which eating frequency affects the outcome of this analysis is a testable proposition that could be explored as a distributed analysis. The Panel noted that a supplementary analysis using a stratified sample of high end consumers would help justify the viability of this approach.

The Panel made several recommendations regarding the cumulative N-methyl carbamate pesticide case study in terms of the data availability, reliability, limitations, and evaluations for use in representing patterns of pesticide co-occurence. In addition, recommendations were made about pesticide residue degradation estimates, residue data availability, and the need to use regional and temporal variability in residues in the exposure scenarios chosen to be assessed by the Agency.

The case study presented to the Panel clearly demonstrated that available data, although limited, could be combined to conduct a cumulative exposure assessment for the N-methyl carbamate pesticides. The major concerns of the Panel were related to the age and quality of the data available, and its impact on pesticide use patterns that have been changing. It was recommended that National Health and Nutrition Examination Survey (NHANES) biomonitoring data be used in combination with PBPK models to serve as a bounding analysis for the cumulative risk assessment.

PANEL DELIBERATIONS AND RESPONSE TO THE CHARGE

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

Question 1

There are several key principles for conducting a cumulative risk assessment. One such principle concerns 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?). Both should be adequately considered when performing a cumulative risk assessment so that an individual's exposure is matched with relevant and appropriate toxicological values in terms of duration and timing. There are several important considerations with respect to the temporal characteristics of the exposures and of the cholinesterase inhibitory effects of N-methyl carbamate pesticides in estimating their cumulative risk.

OPP used a Relative Potency Factor (RPF) approach in this case study which is based on cholinesterase inhibition data from acute dosing studies performed in the rat. A similar approach was used in the organophosphorus pesticide cumulative risk assessment several years ago. This RPF approach expresses toxicity of each chemical in terms of "index chemical equivalents". In this approach, all exposure events within a day are adjusted by their RPFs and summed. The three exposure models (DEEM/Calendex, LifeLine, and CARES) used in the case study express exposure as a distribution of 1 day (24 hour totals) exposures within a population.

Since AChE inhibition caused by the N-methyl carbamate pesticides recovers rapidly (minutes to hours), it might be important to consider the intra-day timing of exposure events. Specifically, if the exposure events within a day are distributed sufficiently far apart in time so that significant recovery of AChE inhibition occurs between any two exposure events, then summing exposures over 24 hours might overestimate the risk associated with AChE inhibition. For example, if an individual consumed 200 mL of apple juice in the early morning, an additional 200 mL during the afternoon, and another 200 mL late at night, this could present a very different risk picture than if the total 600 mL were presumed to be consumed at one time.

The current Food Commodity Intake Database (FCID) and the DEEM/Calendex, LifeLine, and CARES models are set up to consider food consumption on a per day (rather than per eating occasion) manner. Thus, the exposures reported in this case study reflect daily (24 hour) exposures. To the extent that a day's eating occasion events leading to high (total) daily exposure are close together in time, the RPF approach described in the case study provides reasonable

estimates of risk. To the extent that eating exposure events leading to high total daily exposures are widely separated in time such that recovery of AChE inhibition occurs, the risks under the RPF approach in this case study may be overstated and a more sophisticated approach which accounts for intra-day eating patterns would be more appropriate.

OPP has investigated the degree to which high-end exposures can be attributed to specific eating occasions (within a day) that occur either closely spaced in time or widely separated by time by looking at the actual individual consumption events as reported in the Continuing Survey of Food Intakes by Individuals (CSFII). Specifically, OPP has looked at the CSFII –based dietary records for individuals at several locations in the upper end of the exposure distribution to determine the extent to which these daily high-end exposures can be attributed to a single eating event, several eating events spaced closely in time (over several hours), or eating events widely separated by time (more than several hours). As described in Section IV.H of the case study document, OPP found that that a large fraction of daily records contributing to the upper tail of the food exposure distribution represent single eating occasions. Assuming that subsequent, more detailed and extensive analyses provide confirmation of these preliminary observations and analyses, OPP believes and that it is unlikely that any more sophisticated, temporal-based approach which better accounts for temporal separation of eating/exposure events will result in substantial or significant changes in OPP's risk estimates.

Part A. EPA requests the SAP provide comments on this exploratory analysis with respect to its adequacy and appropriateness. Please also provide suggestions for future, more detailed analyses.

Panel Response

The Panel agrees that the overall case study is a thoughtful and useful analysis. EPA has gained much knowledge from development of the organophosphate (OP) cumulative risk assessment and the Agency is to be commended on the carefully developed exploratory analysis, given the varying quality of the datasets that were available. The level of detail is adequate for a case study, but more effort is needed in the full assessment, especially the non-food scenarios. Several additional exposure scenarios should be considered and they are detailed below.

The general framing of Question 1 conveys concern about overestimates of risk. The Panel thinks that EPA should be focusing on identifying areas where overestimates or underestimates might occur, and then treat these specific concerns as points to consider in the variability and uncertainty analysis. Clarity is an important consideration as the Agency moves toward a full case study.

The assessment process is complex, with different models contributing estimates to the overall assessment. It is crucial that a stochastic model that implements a quantitative uncertainty analysis be performed. It is an important addition and a necessary adjunct to the overall cumulative assessment that compares the various pathways of exposure.

Overall data quality is highly variable within the assessment. In general, the food assessment has the highest quality and most abundant data, with lower quality and more sparse datasets for water, incidental ingestion, and dermal pathways. Presentation of model calibrations for these other datasets (e.g., measurements of tap water in the Spring season in southeast North Carolina, dermal exposure post-application studies and model outputs compared to urinary biomonitoring levels observed in the US population) are important in order for the output of the assessment to be an accurate representation of the real world.

Another key analytical issue is the choice of exposure distributions to be used. In particular, the Panel believes that the exposure distributions need careful consideration and should use all available data without artificial truncation of what are typically lognormal exposure distributions. The concern that this may produce unreasonable assessments can be tested empirically if these "high" values are included but subjected to sensitivity and uncertainty analyses in the overall cumulative assessments. Many of the input datasets for the analysis may already artificially truncate the exposure distributions by whom they sample and how the samples are taken. Recruiting by telephone, for example, tends to find established people and likely does not reflect current diversity in race, ethnicity and income in the United States. Truncations, if they are used, should be based on physical limits of concentration.

The Panel believes that EPA should focus on appropriate time scales and populations. The use of a 99.9^{th} percentile is important and should be presented such that it is clear this means approximately 4,000 children in the United States may be at risk given that there are approximately 4 million births in any given year.

Comments on Specific Issues

The omission of institutional exposure (i.e., non-residential, non-occupational) leaves a serious gap in the calculation of aggregate exposure. Each age group will spend time in schools, offices, shopping malls, recreation and entertainment facilities, etc., and there will be pesticide use in these places. It is understood that suitable data may not be available for these places, yet it is important that EPA acknowledge these data gaps in the model development. This issue was identified by the September 27-29, 2000 SAP "Session V - Aggregate and Cumulative Assessments Using Lifeline™ - A Case Study Using Three Hypothetical Pesticides" (p. 31).
http://www.epa.gov/scipoly/sap/2000/september/modelsreport.pdf.

Recall surveys can be of dubious quality. This issue was identified by the April

30 - May 1, 2002 SAP "Cumulative and Aggregate Risk Evaluation System (CARES™) Model Review" (p. 22).
http://www.epa.gov/scipoly/sap/2002/april/caresmeetingminutes.pdf

The Panel reminded the Agency that some of the survey data available for the pilot case study exposure assessment of the N-methyl carbamate pesticides are dated and may not closely reflect some current exposure scenarios or be representative of the current population of the U.S.

Relative potency factors (i.e., the ratio of the toxic potency of a given chemical to that of the index chemical) are used to convert exposures of all chemicals in the group into exposure equivalents of the index chemical. The current presentation of this approach for the calculation of the cumulative residue seems more complicated than necessary (e.g. when it is explained how the amount of residue of each chemical is adjusted by multiplying by a Relative Potency Factor (RPF) to get the equivalent residue of an index chemical). Consulting the Toxic Equivalency Factor-Toxic Equivalent (TEF-TEQ) approach for the dioxins and furans, where one determines the TEFs, obtains the TEQ values of each compound in the mixture exposure, and then does the multiplication and the summation, could serve as a good example to communicate this approach. However, this simple multiplication and calculation holds only if it can be shown that the carbamates act according to the principle of simple similar additive action at the chosen endpoint for risk assessment. This means that the isoboles are parallel straight lines as is assumed for the dioxins. This is a crucial assumption underlying all the analyses at this stage.

Several of the contact categories need to be expanded or better justified (see Estimation of Cumulative Exposure from N-methyl Carbamate Pesticides: A Case Study of 10 N-methyl Carbamates, Table 10. Specific Exposure Routes and Pathways/ Scenarios). Additional contact routes for children should include dermal/oral contact with gardens, ornamentals, and trees. Children have been observed having skin contact with vegetation and placing vegetation in their mouths. Whether this is a major source of dermal or non-dietary ingestion is unclear; there may not be enough data to make a determination as to its impact on children's exposure.

In certain scenarios, EPA should provide better support to justify the exclusion of specific exposure routes such as "ingestion of soil and mouthing of grass" (see Estimation of Cumulative Exposure from N-methyl Carbamate Pesticides: A Case Study of 10 N-methyl Carbamates, Section VI -D). EPA states that the reason for not including an exposure route involving oral ingestion of soil and mouthing of grass was justified because of the 'little impact' such routes had on the individual chemical risk assessments. This justification needs better support by further exploring what difference the cumulative effect of several 'little impacts' might make in the cumulative risk assessment.

In terms of clarity of the document, when discussing the residential exposure

scenarios, EPA states that "The full range of exposure values -- expressed as uniform, log-normal, triangular, or cumulative distributions -- are used, where appropriate, rather than relying on point estimates." (see Estimation of Cumulative Exposure from N-methyl Carbamate Pesticides: A Case Study of 10 N-methyl Carbamates, Section VI-A). While the Panel applauds this move away from point estimates, there needs to be clear descriptions as to why a particular distribution was used rather than another one.

The food assessment is vulnerable in terms of extremes in food consumption, home grown vegetables and home pesticide applications. Furthermore, most of the residues are a single record of composites, and deconvolution of residue data should be considered. The food dataset should not be artificially restricted (e.g., exclusion of over tolerance values from the USDA Pesticide Data Program (PDP) dataset).

Part B. Given the results of the initial exploratory analysis, EPA believes that a more sophisticated time-based intra-day model (e.g., PBPK in which the timing of intra-day eating events is explicitly incorporated) for exposures through the food pathway would not substantially change the assessment of potential risks through this pathway compared to the results produced using the RPF methodology used in the case study in which 24 hour food consumption data is used. Please explain why you agree or disagree.

Panel Response

The default assumption of eating events close together in time is a proper approach in the context of this pilot case study for the N-methyl carbamate exposure assessment. Allowing for recovery seems difficult given the unknown durations that would have to be assigned to defensibly address this issue. The extent to which eating frequency affects the outcome of this case study analysis is a testable proposition that could be explored as a distributed analysis. The Panel noted that a supplementary analysis using a stratified sample of consumers at the high end of pesticide ingestion would help justify the viability of this approach. However, a supplementary analysis should only be done if the underlying data are adequate to support the analysis. The Panel is skeptical that this proposition is testable with existing datasets. For example, the assessment presents comparison of records for a number of eating occasions and found few, if any, with four or more per day (see Estimation of Cumulative Exposure from N-methyl Carbamate Pesticides: A Case Study of 10 N-methyl Carbamates, Section IV-H). Food consumption patterns, particularly for fruit juices in some families, is on an *ad libitum* pattern rather than as a specified number of events per day. The structure of the CSFII surveys would not adequately pick up this type of exposure. Therefore, the concern is that the data do not adequately capture human behavior, and exposures may not be seen because of the quality of the available data and not because the potential for exposure did not exist.

Question 2

A key concept that is unique to cumulative risk assessments is the concept of co-occurrence of residues (and thus co- or simultaneous exposure) to members of the Common Assessment Group (CAG). Specifically, a cumulative assessment must appropriately consider residues that co-occur in time and space since these exposures must by combined and considered jointly. This is true for exposures through the food pathway, the drinking water pathway, and the residential pathway. The USDA's PDP data program uses multi-analyte methods and thus simultaneously measures co-occurring residues in samples. The drinking water concentrations generated from the PRZM-EXAMS model considered regional N-methyl carbamate pesticide use and usage practices and thus implicitly considered co-occuring residues.

Exposures through the residential pathways can also co-occur. One of the unique aspects of the N-methyl carbamate pesticide cumulative exposure assessment is the use of the Residential Exposure Joint Venture (REJV) data that provides current information on co-occurring use patterns for residential exposure. The US EPA National Home and Garden Pesticide Use Survey (NHGPUS) can also be used to develop residential use profiles. The PDP, PRZM-EXAMS, and REJV/NHGPUS data were used, to varying degrees, in this case study.

Please comment on the use of the pesticide use/usage data (e.g., REJV and NHGPUS) to account for co-occurring use patterns in assessing residential exposures.

Panel Response

Data for Co-occurrence

The Residential Exposure Joint Venture (referred to hereafter as the Joint Venture) and the National Home and Garden Pesticide Use Survey (referred to hereafter as the Home and Garden Survey) appear to be the two main databases for characterizing the co-occurring pesticide use patterns. However, their use in assessing the co-occurrence of pesticide use would need to be carefully considered and justified as it is possible that neither database fully captures the diversity of the US population as it presently exists. The overall representativeness of both surveys should be evaluated to ensure they accurately reflect the racial, ethnic, and income diversity that currently exists in the U.S. population.

The Panel appreciates the efforts of the Joint Venture task force that surveyed and collected more updated data on pesticide use patterns. However, without access to this proprietary database, the Panel is unable to provide a proper review. The Panel is also unable at this time, to determine if there might be alternative ways to utilize the data

(e.g., the reports from respondents who provided less than 12 months diaries) or any new dimensions to exposure analysis other than what was presented by the Agency.

The Panel understands that the Agency's effort to evaluate the use of Joint Venture data is on-going and encourages a careful review of the validity and reliability of the Joint Venture data, in particular the potential effects of response bias and errors in reporting as indicated by product-application combinations that are not plausible.

Additional Comments on Specific Issues Regarding the Joint Venture Database.

As mentioned earlier in the Panel's response to Question 1 - Part A, telephone recruiting may not provide the most representative sampling of the population. The dependence on telephone for recruiting may bias the sample to a wealthier than average population and may perhaps also bias against younger users that rely solely on cell phones whose numbers are not readily available. The Agency is encouraged to systematically address the overall issues of telephone recruiting and consider a sub-study of non-responders that includes a telephone or personal interview as well as a validity study of responders that includes personal or telephone interviews.

The demonstration of validity should also include response rates, and the adjustments done to them need to be understandable to the Panel and others who will review the cumulative risk assessment. The extreme loss of response rate after the initial screening points to the need for a careful evaluation of the Joint Venture database. The National Family Opinion (NFO) adjusted survey results for 6 demographic factors in an attempt to make the data representative of the overall U.S. population. The factors chosen to make such adjustments should be related to the use of pesticides, and specifically, to the N-methyl carbamate pesticides.

Issues of statistical inference by region or state are important, as is the number of homes represented by the survey in demographic subgroups. The limitation of sample size becomes a greater issue when the database is divided into more refined geographic locations.

The analysis of the database should include the concordance between the inventory data and application records (accounting for labeled application method where appropriate).

Some of the data, as presented by the Agency, stood out and may indicate the need for a more careful evaluation. For example, the 69% derived proportion of applications to trees was by a handwand sprayer (see Estimation of Cumulative Exposure from N-methyl Carbamate Pesticides: A Case Study of 10 N-methyl Carbamates, Table A.1.2 - Residential Use of Chemical A, REJV/1). Also, the 67% usage rate over a year seems low compared to other surveys. Both Adgate et al. (2000) and the Home and Garden Survey reported usage rates in the 80-90% range.

The co-occurrence information provided by the Joint Venture is unique and valuable, yet the limitations of those data should be fully explored and stated. For example, the table of conditional probabilities (see Estimation of Cumulative Exposure from N-methyl Carbamate Pesticides: A Case Study of 10 N-methyl Carbamates, Table A.1.8 - Chemical A Scenario Co-occurance Probability Matrix) for co-occurrence of pesticide applications to lawns must be based on a notably small number of observations (e.g., 1 or 2 homes).

Given the proprietary nature of the Joint Venture dataset, and the need for transparency in Agency decision-making, EPA should carefully consider the pros and cons of using this data source and the specific information it can provide. The Agency is encouraged to investigate the reliability of the co-occurrence data represented in the Joint Venture, to identify the important limitations, weaknesses, and uncertainties associated with the amount of information therein, and to incorporate this knowledge into the N-methyl carbamate pesticide cumulative risk assessment (CRA).

Additional Comments on Specific Issues Regarding Home and Garden Survey

The main concern regarding the Home and Garden Survey database is the lack of assurance that the pattern of pesticide use has remained the same since the survey was taken more than 15 years ago. In many respects, the household pesticide use practices have likely changed over the years. For example, one cannot assume that pesticide use practices and types of pesticides used for various applications are the same now as then. If the amount of advertising that is conducted for marketing residential pesticides is effective, there should be at least increased product purchasing if not usage. The demographic distribution of responders in the study may not reflect current demographics and therefore may either underestimate or overestimate usage patterns. In addition, the choice of pesticide may also have changed in the past 15 years. It would be desirable to have an updated Home and Garden Survey in order to have some assurance that the survey results and extracted co-occurrences are still valid.

The Joint Venture can be expected to reflect more closely the development of better spraying equipment, improvements in use labeling, and the awareness of domestic pesticide use currently. However, it is unclear how much information from that data source is/will be used. Thus, at this stage, the Panel strongly recommends that the Agency compare and contrast the results from Home and Garden Survey and the Joint Venture data where possible. The advantages and disadvantages of using information from one or both surveys should be determined. The comparison also could allow a better assessment of how the Home and Garden Survey data can be utilized in Lifeline to imply current use patterns. An Agency effort to provide easy public access to the two parts of the Home and Garden Survey report is also recommended.

Pesticide Residue Data

The Panel commented on the approach used for pesticide degradation parameters

to define foliar residue in contact exposure (see Estimation Of Cumulative Exposure From N-Methyl Carbamate Pesticides: A Case Study Of 10 N-Methyl Carbamates, Section VI-F). For example, setting the residue to zero for Chemical A 30 days after its application does not appear to be "conservative" because, based on its half life of 9 days, 10% residue is expected to remain. As for chemical G, a linear degradation rate is highly unusual because almost all environmental processes are first-order or quasi first-order. The Agency is encouraged to revisit the assumptions about chemical degradation with respect to realism versus conservatism.

The Panel recommends that the regional and temporal patterns of N-methyl carbamate pesticide residues be explored across all pathways of exposure (not just for pathways associated with applications around the residence and for drinking water) and reflected in the exposure analysis estimates. For example, the concentrations of certain organophosphates in food display a temporal pattern with the highest concentrations occurring in the spring and summer months. It seems reasonable to also expect that N-methyl carbamate pesticides used in domestic agriculture will display a similar pattern. The Agency is encouraged to further explore the potential seasonal patterns of the PDP data that are used in assessing dietary exposures.

Exposure Scenarios

The studies used to develop the unit exposures for each application scenario should be made available to the Panel to evaluate their validity and reliability. Many factors influence the results of such tests and the Panel is not currently aware of the conditions during the test. Some of the important factors include: sprayer-target geometry; meteorology; spray nozzle type and settings, if adjustable; duration of application events; sampling methods and locations; analytical methods and performances; data reduction and analysis methods.

Ready-to-use products have higher upper range dermal unit exposure values (mg exposure per unit of active ingredient applied) than hand wand, etc. Yet, the amount applied with ready-to-use products is relatively low, so this application technique seems the least important (at least for many scenarios). However, questions remain regarding how the duration of exposure is considered; e.g., if the exposure is assumed to be instantaneous or spread over the duration of the application event.

The Panel commented on the following specific exposure scenarios that are not addressed by the Agency:

1. Omitting the hose-end sprayer application method for ornamental/tree sites may not represent a "conservative" approach. Not explicitly accounting for use on trees may thus be an important limitation of the Lifeline model, and could lead to underestimates of upper-end residential use exposures for products applied to trees. The Agency is encouraged to explore the sensitivity of their results to this limitation.

2. The Agency should account for pesticide use by property managers or commercial pesticide applicators.
3. Hand-to-mouth transfer of pesticides by applicators (adults) due to tobacco smoking is a likely exposure pathway, but it is not included in the pilot exposure case study. Whenever the smoking information for adult applicators is available, this oral ingestion route should be incorporated in the assessment. Dermal and oral exposures are also likely to occur in children age 3 to 5 when N-methyl carbamate pesticides are used in home gardens. Studies have indicated that children of parents who use pesticides in their home gardens (either vegetable or flower) have significantly higher urinary OP metabolite concentrations than children whose parents reported no use of pesticides in their garden (Lu et al. 2001). These two likely exposure routes should be included in the assessment.
4. The co-occurrences considered by the Agency in this current analysis are limited exclusively to pesticide occurrence in media in and around the home. As a result, locations frequented by people outside the home are not considered. For children's exposure, important areas not yet considered might include schools and day care centers where a premium is placed on demonstrable hygiene. The Agency should at least discuss the possibility of exposures outside the home and the associated risks.
5. In the current assessment, exposures associated with household outdoor applications of N-methyl carbamate pesticides require contact with the media treated by the applicant (e.g., lawn, garden and vegetables). Several studies in the scientific literature have demonstrated that common residential use insecticides applied outdoors migrate into the home by being tracked inside (Lewis et al., 2001; Stout and Leidy, 2000; Nishioka et al., 2001). The Agency should consider the importance of this potential exposure pathway in their continuing assessment of N-methyl carbamate pesticides.
6. The present analysis correctly puts a large amount of effort into distinguishing between the age groups. This is important and needed because of the possible higher risk of young children. On the other hand, age is just one factor that can be considered. Others that could be considered include vegetarians versus non –vegetarians or other groups of persons with specific behavioral differences which cause heterogeneity. The simple PK model would allow this distinction since it adjusts for the incorporation of a number of experimental factors.
7. It is not immediately clear how certain pathways were omitted from the garden scenarios. The Agency should consider the tendency of children to mimic adult behaviors.

Question 3

The data sources and methodologies used in the N-methyl carbamate case study are similar in many respects to the data sources and methodologies used in the Cumulative Risk Assessment for the OP pesticides. For example, in both assessments the evaluation of exposure of the food pathway relied to a great extent

on the USDA's PDP data, the evaluation of exposure through the water pathway used PRZM-EXAMS modeling data, and the evaluation of exposure through the residential pathway used standard SOP algorithms along with label information, professional judgments, literature values, and survey data (REJV, NHGPUS).

Please comment on the data sources used in the cumulative exposure assessment and on how EPA has considered and incorporated the data. Does the SAP have any suggestions or recommendations regarding additional available data sources that EPA may wish to investigate?

Panel Response

The data sources used in the cumulative exposure assessment were those currently available. The Panel raised a number of concerns about the age and quality of the datasets and urged EPA to test the outcome of its assessment against the NHANES carbamate metabolite data as a means of reality testing. Concerns were raised about secular trends in pesticide use, the age of some of the data, how the data are used, how the data are interpreted, and the types of distributions developed from the data.

Sources Used in Assessment

The case study presented to the SAP clearly demonstrated that available data, although limited, could be combined to conduct a cumulative exposure assessment for the N-methyl carbamate pesticides. All 3 models were built upon the same datasets; therefore, the similar outputs, particularly at the high-end exposure profiles from all models, are not surprising. While the improvement of these models is on-going, the Agency should begin the process of validating all of the models in terms of adequacy and appropriateness by using a different dataset. The immediately available data for such an exercise are data collected by Bayer from their carbaryl turf application study. Assuming the study took place in the southeast region of North Carolina and therefore took into account the component of drinking water exposure in the model simulation, the outputs from these models can then be used as inputs to the PBPK model for estimating 24-hour average 1-naphthol levels. The modeled 1-naphthol levels can then be compared to 1-naphthol levels reported by Bayer. From here, the Agency could decide which model would provide the most robust and accurate prediction, and if not, what modifications are needed in the CRA models so that the gap between predicted vs. observed levels can be brought closer.

Comments Regarding Specific Issues

1. Secular trends: Several Panel members raised the issue of possible change in use patterns and trends away from using cholinesterase inhibiting pesticides. The Panelists question whether the Agency used any factors to adjust for this potential factor. There have been significant secular trends in overall pesticide use with shifts away from

acetylcholinesterase inhibitors because of their inherent toxicity. While there has not been an outright ban or prohibition of any carbamates, we are likely to see far less household use of such pesticides. At the very least, the mix of carbamates used in residential and agricultural settings has likely shifted. It is unclear in the background documents provided to the Panel what correction factors were used.

2. Age of data: There was considerable concern among the Panel about the age of data collected with the NHGPUS 15 years ago. One cannot assume that pesticide use practices and types of pesticides used for various applications are the same now as then. Also, the demographic distribution of responders in that study may not reflect current demographics and therefore may either underestimate or overestimate usage patterns. In addition, the choice of pesticide also may have changed in the past 15 years. Since NHGPUS was published around 1990, and likely represent data collected around 1988, it is quite dated.

3. Data used: Some of the data used are regionally specific (water) while other appear not to be regionally specific. The regional use of data meant to be representative of the nation is inappropriate and likely to be error prone. Since the case study is in the Southeast, it is appropriate to try and use data specific to that region. The authors should be clear on the potential hazards of using such data on regions smaller than that for which they are truly representative.

This assessment uses food residue data collected over a nine year period, 1994 through 2002. A more recent and smaller window would be more appropriate. As stated in the document, the primary reason for this choice is to maximize the number of food commodities. With this design, there is a trade off between precision and bias that needs to be addressed. It would be interesting to see how robust the estimates are when early years are deleted.

Support for some of the assumptions may be made by correlating overall exposure with pesticide sales figures, at least if one collects data over a long period of time.

It was also felt that deconvolution methods should be used to impute individual serving exposure values from composite samples.

4. Data Interpretation: Observational data for children contact with pets are from a very small dataset and do not cover activities such as sleeping with pets. It also is based on a single period of observation per child.

5. Data robustness: It is important to get away from professional judgments that are more difficult to evaluate by outside reviewers and sometimes less transparent in how they were developed, and use existing data bases. Robustness of databases is always a concern, particularly when they were not developed specifically to meet the needs of the Agency for use in modeling exercises such as this one.

6. Uncertainty Analysis: It is very important to accompany the assessment with a substantial uncertainty analysis. The model comparisons of the exposure results are presented in Table 15 of the background document, Estimation of Cumulative Exposure from N-methyl Carbamate Pesticides: A Case Study of 10 N-methyl Carbamates, with six significant digits. It seems unlikely, for a host of reasons, that such precision really attends these numbers. Estimating the 99.9^{th} percentile, for instance, at this precision would require many more replications than were likely deployed in the Monte Carlo simulations. The model uncertainty expressed by the tables suggests only one or two significant digits. There are, perhaps, even fewer after the measurement uncertainties and the sampling error that arises from having few data samples in some of the cells are accounted for.

The Agency attempted to quantify the reliability of the simulation results. It should be pointed out, however, that changing the numerical seed and re-running simulations is *not* a sufficient way to conduct uncertainty analyses. At the very best, this approach can only assess the uncertainty arising from the use of a particular sequence of random numbers. We might call this the 'simulation uncertainty'. This is perhaps the easiest kind of uncertainty to address, but it is likely to be the smallest kind affecting the overall results; or at least can always be made to be the smallest, simply by increasing the number of Monte Carlo replications. The Panel emphasizes that the consequences for the reliability of the results attributable to model uncertainty, sampling uncertainty, and measurement uncertainty in general will likely be much greater, and will probably lead to the results having no significant figures at all. In such a case, an explicit uncertainty analysis is crucial to interpreting the import of the calculations. Measurement uncertainty may be especially hard to estimate because the various model approaches have been based on overlapping and essentially similar datasets.

The Panel voiced concerns about truncation several times. It was felt that the Agency should not truncate a parameterized distribution without a reasoned argument to justify this decision. The same issue arises with *empirical distributions*. Using an empirical distribution is perhaps the most egregious kind of unjustified truncation. As several Panelists mentioned, higher values are almost certain to be obtained than the original data contained if more samples are collected. It is impossible to know the impact of truncations unless an uncertainty analysis is conducted that addresses the

consequences of ignoring the tails. The Panel encourages EPA to continue to study the model stability and sensitivity.

One Panel member suggested that it is important to assess the appropriateness, sufficiency and precision of the outcome of this cumulative risk assessment for the N-methyl carbamate pesticides. For example, are N-methyl carbamate pesticides in drinking water alone capable of elevating the exposure profile during a certain period of time during a year for all ages of modeled individuals at the 95^{th} percentile? The other interesting scenario about the outcome of this cumulative exposure assessment is that, assuming seasonal contribution of N-methyl carbamate cumulative exposure from drinking water is real, and most, if not all, of the N-methyl carbamate pesticide residues come from the seasonal agricultural uses in the modeled region, this seasonal effect is not reflected in the food component.

According to data presented in Table 16 of the background document Estimation of Cumulative Exposure from N-methyl Carbamate Pesticides: A Case Study of 10 N-methyl Carbamates, approximately 65% of total N-methyl carbamate dietary exposure comes from a single food type, which is citrus fruits, including orange, tangerine, and grapefruit. If the availability of citrus fruits to the modeled individuals is limited only to a certain period of time in a year, say winter months, this seasonal effect obviously does not transmit to the total exposure profile. It was felt that longitudinal correlations across seasons are included in the analyses. There was discussion of the longitudinal nature of the PDP studies and the inclusion of foreign food items and items from other areas of the country (i.e., Florida versus California oranges).

Diurnal patterns of exposure are very important for biomonitoring studies of exposure and effect and are not really addressed in the Agency assessments. If we measure a parent/metabolite at one time of day, yet exposures vary across the day, we are likely to get substantially different inferences for exposure and risk depending on when the exposures occurred. A morning exposure may be completely eliminated by late evening. If a urine sample is taken in the late evening, the exposure inferred will be substantially smaller than what might be inferred from a noon urine sample.

Other Data Sources

There are other sources of data that can be used to challenge and evaluate the qualitative accuracy of the data on exposure being generated in this analysis. Specifically, the Centers for Disease Control and Prevention (CDC) has produced the "Second National Report on Human Exposure to Environmental Chemicals" (CDC Publication 02-0716) which determined levels of 1- Naphthol (carbaryl metabolite), 2-isopropoxyphenol (metabolite of propoxur) and carbofuranphenol (metabolite of benzofuracarb and carbofuran) in human urine. Through "best estimate" and "conservative" assumptions and use of the PBPK model presented during this SAP meeting, it would be possible to bound the amount ingested prior to the measurement and compare this with the range seen in the exposure assessment. For example, 1-naphthol is

a metabolic product of carbaryl, naphthalene and some components of cigarette smoke. Assuming all of this is from carbaryl is a "conservative" choice. Calculating a likely urinary level from the exposure estimates and the PBPK model and comparing this to what was seen could provide the maximum percentage in urine that would be due to carbaryl if the exposure estimates are correct. While this still presents some challenges (how long between urinations, what to do about the large number of observations below the limit-of-detection, other sources of the metabolite, etc.), it is a reality check worth doing. The median urinary level of 1-naphthol in children 6-11 years of age was 1.1 micrograms/L, and the 95%-tile 5.6 micrograms/L. From the PBPK model and the data used to form that model, it appears that 1-naphthol will be a minor urinary metabolite of carbaryl, accounting for less than 5% (high metabolism but small urinary excretion) and suggesting about a 20-fold ingestion relative to urinary output; however, this could easily be calculated more accurately by those who are familiar with the model. For 2-isopropoxyphenol, both the median and the 95%-tile are below the limit-of-detection. For carbofuranphenol, the median is below the limit-of-detection and the 95%-tile is 0.43 micrograms/L. The sample sizes for 6-11 year-olds are between 450 and 483 and for the total dataset are between 1917 and 1998. They have the advantage of being representative of the national population and can be divided into broad regions.

In terms of additional data, the Agency noted that a study on lawns and gardens use of Chemical A conducted by Rhone-Poulenc in 1998 showed similar results as estimates from the NHGPUS (see Estimation of Cumulative Exposure from N-methyl Carbamate Pesticides: A Case Study of 10 N-methyl Carbamates, Appendix1 - Page 148, Table A.1.6). It appears that this study may provide a more current context to the NHGPUS and the Agency is encouraged to present this data set in greater detail.

REFERENCES

Adgate, J.L., Kukowski, A., Stroebel, C., Shubat, P.J., Morrell, S., Quackenboss, J.J., Whitmore, R.W. and Sexton, K. (2000) Pesticide storage and use patterns in Minnesota households with children. J. Expos. Anal. Environ. Epi. 10:159-167

Centers for Disease Control and Prevention. National Center for Environmental Health. 2003. Second National Report on Human Exposure to Environmental Chemicals. (NCEH Publication # 02-0716)

Lewis, R.G., Fortune, C.R., Blanchard, F.T., and Camann, D.E. (2001) Movement and deposition of two organophosphorous pesticides within a residence after interior and exterior applications. *J. Air Waste Manage. Assoc.* 51:339-351.

Lu, C., Fisker-Anderson, J., Knutson, D., Moate, T. and Fenske, R. (2001) Biological monitoring survey of organophosphorus pesticide exposure among pre-school children in the Seattle metropolitan area. *Environ Health Perspect*. 109(3):299-303.

Nishioka, M.G., Lewis, R.G., Brinkman, M.C., Burkholder, H.M., Hines, C.E. and Menkedick, J.R. (2001) Distribution of 2,4-D in air and on surfaces inside residences after lawn applications: comparing exposure estimates from various media for young children. *Environm. Health Perspect.* 109(11):1185-1191.

Stout, D.M. and Leidy, R.B. (2000). A preliminary examination of the translocation of microencapsulated cyfluthrin following applications to the perimeter of residential dwellings. *J. Environ. Sci. Health, Part B.* 35(4):477-489.

U.S. Environmental Protection Agency. 2000. Report of the Scientific Advisory Panel Meeting held September 27-29, 2000. "Session V - Aggregate and Cumulative Assessments Using Lifeline™ - A Case Study Using Three Hypothetical Pesticides" (p. 31). FIFRA Scientific Advisory Panel. Office of Science Coordination and Policy. Report # 2000-03. http://www.epa.gov/scipoly/sap/2000/september/modelsreport.pdf.

US Environmental Protection Agency. 2002. Report of the Scientific Advisory Panel Meeting held April 30 - May 1, 2002. "Cumulative and Aggregate Risk Evaluation System (CARES™) Model Review" (p. 22). FIFRA Scientific Advisory Panel. Office of Science Coordination and Policy. Report # 2002-02. http://www.epa.gov/scipoly/sap/2002/april/caresmeetingminutes.pdf