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

OFFICE OF CHEMICAL
SAFETY AND
POLLUTION PREVENTION

MEMORANDUM

DATE: May 12, 2011

SUBJECT: Transmittal of the Meeting Minutes of the FIFRA SAP Meeting Held February 15-17, 2011 on the Scientific Issues Associated with "Chlorpyrifos Physiologically Based Pharmacokinetic and Pharmacodynamic (PBPK/PD) Modeling Linked to Cumulative and Aggregate Risk Evaluation System (CARES)"

TO: Steven Bradbury, Ph.D.
Director
Office of Pesticide Programs

FROM: Sharlene Matten, Ph.D.
Designated Federal Official
FIFRA SAP Staff
Office of Science Coordination and Policy

THRU: Laura Bailey 5/17/2011
Executive Secretary, FIFRA SAP
Office of Science Coordination and Policy

Frank Sanders
Director
Office of Science Coordination and Policy

Please find attached to this memorandum the meeting minutes of the Federal Insecticide, Fungicide, and Rodenticide Act (FIFRA) Scientific Advisory Panel (SAP) open meeting held in Arlington, Virginia on February 15-17, 2011. This report addresses a set of scientific issues associated with "Chlorpyrifos Physiologically Based Pharmacokinetic and Pharmacodynamic (PBPK/PD) Modeling Linked to Cumulative and Aggregate Risk Evaluation System (CARES)."

Attachment

cc:
Stephen Owens
William Diamond
Vicki Dellarco
William Jordan
Margie Fehrenbach
Donald Brady
Joan Harrigan-Farrelly
Jack Housenger
Tina Levine
Richard Keigwin
Keith Matthews
Robert McNally
Lois Rossi
Douglas Parsons
Enesta Jones
Vanessa Vu
Jack Fowle
Anna Lowit
David Miller
OPP Regulatory Docket

FIFRA SAP Members
Kenneth Portier, Ph.D. (FIFRA SAP Chair)
Janice Chambers, Ph.D., DABT, ATS
Carey Pope, Ph.D.
Martha Sandy, Ph.D.

FQPA Science Review Board Members
William Brimijoin, Ph.D.
James Bruckner, Ph.D.
Russell Carr, Ph.D.
Panos Georgopoulos, Ph.D.
Jogarao Gobburu, Ph.D.
Sami Haddad, Ph.D.
Teresa Leavens, Ph.D.
Chensheng (Alex) Lu, Ph.D.
David Macintosh, Sc.D.
Nu-May Ruby Reed, Ph.D., DABT
P. Barry Ryan, Ph.D.
Michael Sohn, Ph.D.

SAP Minutes No. 2011-03

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

Chlorpyrifos Physiologically Based Pharmacokinetic and Pharmacodynamic (PBPK/PD) Modeling Linked to Cumulative and Aggregate Risk Evaluation System (CARES)

February 15-17, 2011
FIFRA Scientific Advisory Panel Meeting
Held at
One Potomac Yard
Arlington, Virginia

NOTICE

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

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

In preparing these meeting minutes, the Panel carefully considered all information provided and presented by EPA and Dow AgroSciences (DAS), as well as information presented in public comment. This document addresses the information provided and presented by EPA and DAS within the structure of the charge.

ED_005731_00001263-00004

TABLE OF CONTENTS

NOTICE **2**

INTRODUCTION **8**

PUBLIC COMMENTERS **10**

SUMMARY OF PANEL DISCUSSION AND RECOMMENDATIONS **11**

DETAILED RESPONSES TO CHARGE QUESTIONS **19**

- CHARGE QUESTION 1: PHYSIOLOGICALLY-BASED PHARMACOKINETIC/PHARMACODYNAMIC (PBPK/PD) MODELING 19
- CHARGE QUESTION 2: LONGITUDINAL DIETARY EXPOSURE ASSESSMENT 30
- CHARGE QUESTION 3: MODEL CALIBRATION AND EVALUATION WITH DIRECT DOSING HUMAN STUDIES 34
- CHARGE QUESTION 4: COMPARISON OF MODEL PREDICTIONS WITH HUMAN MONITORING DATA 38
- CHARGE QUESTION 5: SENSITIVITY ANALYSES, VARIABILITY, AND UNCERTAINTY 48
- CHARGE QUESTION 6: CALCULATING DATA-DERIVED EXTRAPOLATION FACTORS 60
- REVISED MODEL NAME 68
- GRAMMATICAL/TYPOGRAPHICAL ERRORS IN THE DAS WHITE PAPER 68

REFERENCES **69**

ED_005731_00001263-00005

SAP Minutes No. 2011-03

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

Chlorpyrifos Physiologically Based Pharmacokinetic and Pharmacodynamic (PBPK/PD) Modeling Linked to Cumulative and Aggregate Risk Evaluation System (CARES)

February 15-17, 2011
FIFRA Scientific Advisory Panel Meeting
Held at
One Potomac Yard
Arlington, Virginia

Kenneth M. Portier, Ph.D.
FIFRA SAP Chair
FIFRA Scientific Advisory Panel

Date: 5/11/2011

Sharlene R. Matten, Ph.D.
Designated Federal Official
FIFRA Scientific Advisory Panel Staff

Date: 5/11/2011

ED_005731_00001263-00006

Panel Member List for the Meeting of the Federal Insecticide, Fungicide, and Rodenticide Act Scientific Advisory Panel (FIFRA SAP) to Consider and Review Chlorpyrifos Physiologically Based Pharmacokinetic and Pharmacodynamic (PBPK/PD) Modeling Linked to Cumulative and Aggregate Risk Evaluation System (CARES). February 15-18, 2011

EPA-HQ-OPP-2010-0588 **OPP Docket Tel: 703-305-5805**

FIFRA SAP Chair

Kenneth M. Portier, Ph.D.
Program Director, Statistics
American Cancer Society
National Home Office
Atlanta, GA

Designated Federal Official

Sharlene R. Matten, Ph.D.
US Environmental Protection Agency
Office of Science Coordination & Policy
FIFRA Scientific Advisory Panel
EPA East Building, MC 7201M
1200 Pennsylvania Avenue, NW
Washington, DC 20460
Tel: 202-564-0130, Fax: 202-564-8382, matten.sharlene@epa.gov

FIFRA Scientific Advisory Panel Members

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

Carey N. Pope, Ph.D.
Professor, Head & Sitlington Chair of Toxicology
Department of Physiological Sciences
College of Veterinary Medicine
Oklahoma State University
Stillwater, OK

Martha S. Sandy, Ph.D.
Senior Toxicologist and Chief

ED_005731_00001263-00007

Cancer Toxicology and Epidemiology Section
Reproductive and Cancer Hazard Assessment Branch
Office of Environmental Health Hazard Assessment
California Environmental Protection Agency
Oakland, CA

FQPA Science Review Board Members

William S. Brimijoin, Ph.D.
Professor and Chair
Department of Molecular Pharmacology and Experimental Therapeutics
Mayo Clinic and Medical School
Rochester, MN

James V. Bruckner, Ph.D.
Professor
Department of Pharmacology & Toxicology
University of Georgia
Athens, GA

Russell L. Carr, Ph.D.
Associate Professor
Center for Environmental Health Sciences,
College of Veterinary Medicine
Mississippi State University
Mississippi State, MS

Panos G. Georgopoulos, Ph.D.
Professor
Department of Environmental and Occupational Medicine
Environmental & Occupational Health Sciences Institute (EOHSI)
University of Medicine and Dentistry of New Jersey
Robert Wood Johnson Medical School
Piscataway, NJ

Jogarao Gobburu, Ph.D.
Director, Division of Pharmacometrics
Office of Clinical Pharmacology
U.S. Food and Drug Administration
Silver Spring, MD

Sami Haddad, Ph.D.
Associate Professor
Departément des Sciences Biologiques
Université de Montréal
Montréal, Canada

Teresa Leavens, Ph.D.
Research Assistant Professor
College of Veterinary Medicine
Department of Population Health and Pathobiology
Center for Chemical Toxicology Research and Pharmacokinetics
North Carolina State University
Raleigh, NC

Chensheng (Alex) Lu, Ph.D.
Mark and Catherine Winkler Assistant Professor of Environmental Exposure Biology
Department of Environmental Health
Harvard School of Public Health
Cambridge, MA

David Macintosh, Sc.D.
Principal Scientist
Environmental Health & Engineering, Inc.
Needham, MA

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

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

Michael D. Sohn, Ph.D.
Staff Scientist and Group Leader
Airflow and Pollutant Transport Group
Indoor Environment Department
Lawrence Berkeley National Laboratory
Berkeley, CA

ED_005731_00001263-00009

INTRODUCTION

The Federal Insecticide, Fungicide, and Rodenticide Act (FIFRA) Scientific Advisory Panel (SAP) has completed its report of the SAP meeting regarding scientific issues associated with **"Chlorpyrifos Physiologically-Based Pharmacokinetic and Pharmacodynamic (PBPK/PD) Modeling Linked to Cumulative and Aggregate Risk Evaluation System (CARES)."** Advance notice of the SAP meeting was published in the *Federal Register* on **December 6, 2010**. The review was conducted in an open Panel meeting on **February 15-17, 2011** at One Potomac Yard, Arlington, Virginia. Materials for this meeting are available in the Office of Pesticide Programs (OPP) regulatory docket or via www.regulations.gov, Docket No.: EPA-HQ-OPP-2010-0588. Kenneth Portier, Ph.D., chaired the meeting. Sharlene Matten, Ph.D., served as the Designated Federal Official. Steven Bradbury, Ph.D., Director, Office of Pesticide Programs (OPP), and John (Jack) Fowle III, Ph.D., Deputy Director, Health Effects Division (HED), provided opening remarks at the meeting. Agency presentations of technical background materials were provided by Anna Lowit, Ph.D., David J. Miller, and Chester Rodriguez, Ph.D., all of HED. In addition, technical presentations were provided by Michael Bartels, Ph.D., Mr. Paul Price, and Sue Marty, Ph.D., of Dow AgroSciences, LLC and Torka Poet, Ph.D. and Paul Hinderliter, Ph.D., of Battelle, Pacific Northwest National Laboratory.

Chlorpyrifos (*O,O*-diethyl *O-(*3,5,6-trichloro-2-pyridinyl*)* phosphorothioate) is a broad-spectrum, chlorinated organophosphorus (OP) insecticide. In 2000, nearly all residential uses were cancelled voluntarily by Dow AgroSciences (DAS), but agricultural uses remain. The 2000 human health risk assessment was largely based on adult laboratory animal data (rat or dog) for cholinesterase inhibition and the application of default uncertainty factors to address inter- and intra-species differences including susceptible populations. Currently, the Agency is developing a new human health risk assessment expected to be released in June 2011. In 2008, the FIFRA Scientific Advisory Panel (SAP) reviewed a draft science issue paper on the human health effects of chlorpyrifos.

Since that time, DAS has developed a source-to-outcome model, which is composed of existing probabilistic dietary exposure software to produce estimates of longitudinal exposure (CARES model) and a PBPK/PD model of response (LifeStage PBPK/PD model). Coupled together, these model components relate data on pesticide residues on crops to the probability of inhibition of cholinesterase in humans. Chlorpyrifos was used as an example pesticide because there exists a comparatively rich data set on metabolism and mechanism of action of cholinesterase inhibition, multiple markers for exposure and effects, quantitative PBPK/PD models to predict, markers of early effects (i.e., inhibition of red blood-cell and plasma cholinesterases), and human data for model calibration and evaluation.

The linking of the chlorpyrifos PBPK/PD model with a probabilistic exposure model may provide opportunities to calculate distributions of exposure to chlorpyrifos and its metabolites with cholinesterase inhibition levels across the U.S. population. In addition, this approach may allow estimation of data-derived uncertainty factors that consider use of toxicokinetic and toxicodynamic data to inform quantitative extrapolations for interspecies differences and human variability in dose response assessment. The topics covered during the February 2011 SAP are consistent with EPA's Office of Pesticide Programs' continuing efforts to improve the scientific

basis for risk assessment by broadening the application of probabilistic exposure techniques and PBPK models. The Agency has a conceptually similar effort on-going to link PBPK models for pyrethroids with Stochastic Human Exposure and Dose Simulator (SHEDS) exposure software, a probabilistic exposure model developed by the EPA's Office of Research and Development, which was reviewed by the SAP in July, 2010. The current effort by DAS is a research effort, which may, if sufficiently robust, inform future Agency risk assessments; the February meeting is a key milestone in the process.

The purpose of the February 2011 SAP meeting was to request scientific advice from the Panel on technical issues related to the PBPK/PD model, the proposed approach for merging the PBPK/PD model with CARES, and the use of such tools in risk assessment.

PUBLIC COMMENTERS

Oral statements were presented by:

1) Dale Hattis, Ph.D., Research Professor, Clark University, Worcester, MA

2) Wendelyn Jones, Ph.D., Senior Director, Human Health Policy, CropLife America, Washington D.C.

Written statements were provided by:

1) Dale Hattis, Ph.D., Research Professor, Clark University, Worcester, MA

2) Carol Dansereau, Executive Director, Farm Worker Pesticide Project, Seattle, WA

ED_005731_00001263-00012

SUMMARY OF PANEL DISCUSSION AND RECOMMENDATIONS

Charge Question 1: Physiologically Based Pharmacokinetic/Pharmacodynamic (PBPK/PD) Modeling

QUESTION 1.1: Model Structure*. Please comment on the structure of the chlorpyrifos PBPK/PD model with specific consideration of the mechanistic basis for the acetylcholinesterase (AChE) inhibiting mode of action. Please include in your comments consideration of age-dependant metabolism and the proposed approach to assess human variability.*

Panel Response Summary

With certain reservations, the Panel found that the mechanistic basis for the AChE-inhibiting mode of action of chlorpyrifos had been satisfactorily incorporated into the PBPK/PD model, noting that the model did not address non-cholinergic effects of chlorpyrifos. The Panel generally supported the focus on inhibition of AChE and red blood cell (RBC) AChE as the most suitable marker and surrogate for the true target of the anticholinesterase effect of chlorpyrifos, specifically synaptic AChE in brain, spinal cord, and peripheral nervous systems. However, the Panel identified several concerns with how DAS handled age-dependent metabolism in the model and the proposed approach to assess human variability. Among other factors, limited numbers of samples to study age-dependent human metabolism (in particular with very young individuals) and the use of non-physiological conditions to estimate metabolic rates, limited the Panel's confidence in that component of the model. There were concerns about the parameterization of the model based on the unpublished age-dependent metabolism study (Smith et al., 2011 as Attachment A[1] in DAS, 2011), especially with respect to paraoxonase (PON), the enzyme that hydrolyzes chlorpyrifos-oxon. The Panel had difficulty in evaluating the DAS model because the pharmacokinetic and pharmacodynamic components were grouped together in a way that made it hard to assess the validity of the structure and parameters of each component. The Panel recommended that DAS provide the source of and details about all of the model parameters in an effort to be more transparent and complete. Such information would significantly raise the general understanding of and confidence in the model. Overall, the model would be greatly improved if the parameterization for PON was based on more accurate and relevant experimental data.

QUESTION 1.2: Dose Metrics. *In Section 6, DAS proposes a number of dose metrics for use in the PBPK/PD effort. These include peak levels of blood and brain AChE inhibition, peak blood and brain levels of chlorpyrifos and its oxon metabolite, and urinary measures of 3,5,6-trichloro-2-pyridinol (TCPy), a chlorpyrifos metabolite. Please comment on the utility, strengths, and limitations of these proposed internal dose metrics.*

---

[1] DAS Report (2011), Attachment A. Smith, J.N., C. Timchalk, M.J. Bartel, T.S. Poet. 2011. *In vitro* age-dependent enzymatic metabolism of chlopyrifos and chlorpyrifos-oxon in human hepatic microsomes as well as chlorpyrifos-oxon in plasma. *Unpublished manuscript.*

ED_005731_00001263-00013

Panel Response Summary

The Panel indicated that all of the named parameters could be supported scientifically as metrics for chlorpyrifos exposure with some limitations. Peak brain AChE inhibition could be an indicator of potential toxicity, but cannot be measured in humans. Peak blood AChE inhibition could serve both as a biomarker of exposure and as a potential surrogate marker for target tissue toxicity; however, its measure is highly variable among people and therefore results must be interpreted with this variability in mind. Peak chlorpyrifos concentration may be useful for computing acute exposures, but the area under the time-concentration curve may be more useful than the peak level for computing the effects of repeated exposures. The peak chlorpyrifos-oxon concentration is potentially useful because chlorpyrifos-oxon is the active species, although this metabolite is highly reactive and rarely measurable. Consequently, little confidence could be placed in the concentrations measured. Therefore, chlorpyrifos-oxon concentration is probably not a practical input parameter, although it could be useful for the model to predict values for it. Another dose-metric with limited utility is urinary TCPy. Urinary TCPy can be biased by pass-through TCPy that is absorbed as an environmental degradate; therefore, these measurements should be interpreted with caution.

QUESTION 1.3*: *Non-Cholinergic Effects. *Although the current modeling effort focuses on AChE inhibition, there are data to suggest that chlorpyrifos exposure may also result in non-cholinergic effects. As such, and if appropriate, please provide plausible additional and/or alternative dose metrics such as area under the curve (AUC) metrics or other temporal-based internal dose metrics that may be appropriate for evaluating potential non-cholinergic effects.*

Panel Response Summary

The Panel thought it was reasonable to consider non-cholinergic effects in the long term, but did not think it would be appropriate to provide guidance on any specific dose metrics as there are no standard methods to measure non-cholinergic activity and no non-cholinergic mechanism has yet been validated. Obtaining model output for AUC for chlorpyrifos and chlorpyrifos-oxon will be useful for use in simulating non-cholinergic effects. Modeling BChE inhibition may also be useful as a potential measure of non-cholinergic toxicity. While non-cholinergic effects may be of high interest and importance, any such non-cholinergic effects of chlorpyrifos have not yet been confirmed.

Charge Question 2: Longitudinal Dietary Exposure Assessment

QUESTION 2.1: *Please comment on the methods used by DAS to investigate the relationships between dietary exposures and levels of chlorpyrifos and chlorpyrifos-oxon in blood. Please include in your comments discussion on the strengths, weaknesses, impacts, and utility of the DAS proposal to focus on the immediate prior and subsequent days to high-end exposures and provide alternative approaches to assessing longitudinal (multi-day) exposures, if appropriate. Under what conditions and scenarios might this focus be inadvisable and/or lead to incorrect conclusions regarding exposures, doses, and risks?*

ED_005731_00001263-00014

Panel Response Summary

The Panel concluded that the scope of DAS' longitudinal dietary exposure was incomplete. It should include dietary intake over many days (not just five days) and under various, correlated, eating behaviors. The Panel recognized that limitations in the CARES and/or Lifeline model make such analyses difficult.

QUESTION 2.2: *Please comment on the soundness of the conclusions reached by DAS with respect to the impact and importance of accounting for longitudinal dietary exposures to chlorpyrifos at the upper ends (e.g., >99th percentile) of the exposure distribution. To what extent, if any, can the DAS conclusions be generalized to other AChE-inhibiting chemicals and under what conditions? Please suggest how this issue of generalization to other pesticides could be investigated and explored?*

Panel Response Summary

The Panel agreed that it was important to account for longitudinal dietary exposures to chlorpyrifos at the upper ends (e.g., >99th percentile) of the exposure distribution. However, the Panel stated that it would be premature to conclude that the chlorpyrifos analysis can be generalized to other chemicals or when multiple AChE-inhibiting chemicals are present in a single eating event. For example, the CARES-PBPK/PD model predictions for chlorpyrifos may be erroneous if there are other pesticides also present with different affinities for BChE or AChE.

Charge Question 3: Model Calibration and Evaluation with Direct Dosing Human Studies

QUESTION 3: *Please comment on the model evaluation approach comparing the linked CARES-PBPK/PD source-to-outcome model predictions with the Kisicki et al. 1999 study. In what ways could or should the model evaluation approach used by DAS be extended? Are there other model evaluation methods with respect to this aspect of the DAS manuscript that the Panel suggests be performed? To what extent does the Panel agree that the DAS model predictions are reasonably consistent with those of the Kisicki et al (1999) literature-reported values?*

Panel Response Summary

The Panel agreed that the data from the Nolan et al. (1982) and Kisicki et al. (1999) studies were scientifically suitable to be used in the model evaluation. The Panel concluded the Nolan et al. (1982) study was suitable for calibrating the DAS PBPK/PD source-to-outcome model. However, the Panel noted that, for several reasons, the model was not "validated" using the Kisicki data. For one, the Panel commented that since some of the model parameters were calibrated to the Kisicki data, it puts into question the independence of a model-to-data validation.

The model evaluation that this question refers to is only the comparison of the LifeStage PBPK/PD model to the Kisicki et al. (1999) data. In light of the limitations of the Nolan and Kisicki studies, the Panel concluded that the model fit to the Kisicki data was generally

reasonable (by visual comparison), although the model fit for the highest dose was not particularly good, and the model's ability to predict cholinesterase inhibition recovery was not as good as would be desired. The charge question asked the Panel to comment on the "reasonableness" of a model evaluation without providing a compelling description of how the joined models will be used. Without knowing how a model will be used, it is not possible to comment sufficiently on its reasonableness or appropriateness. The Panel recommended that the potential sources of uncertainties identified in these and other comparisons should be captured for assessing the overall model uncertainties if the model is used for risk assessment. This analysis will be useful to ensure goodness of fit at a dose range outside the potential human exposure range and for possible model adjustment that can reduce the uncertainties in model predictions at the range relevant to human exposures.

Charge Question 4: Comparison of Model Predictions with Human Monitoring Data

QUESTION 4: *Please comment on the model evaluation approach used by DAS to compare the linked DAS source-to-outcome model dose predictions to (i) Eaton et al (2008), Whyatt et al. (2009) and Barr et al. (2010), (ii) the Curwin (2007) and Lu (2008) literature values and (iii) NHANES data measurements. To what extent does the Panel agree that the DAS model predictions are reasonably consistent with those of the literature-reported values and the NHANES data? Please suggest, if appropriate, other model evaluation methods or alternative approaches for comparing model predictions with actual human exposure that the Panel recommends to further evaluate the model predictions. Please include in your comments suggestions for what additional datasets should be used for comparison.*

Panel Response Summary

The Panel indicated that DAS made a strong effort to compare the results of the source-to-outcome model and the validation model to data and past models reported in the public record. The DAS model provided similar results to other published models, but when comparing results to direct measures from biological field sampling, there were only two relatively small studies of chlorpyrifos blood and serum levels from two adjacent states, New York and New Jersey (Whyatt et al., 2009 and Barr et al., 2010).

The Panel found the comparison between the DAS model and the NHANES data to be problematic. The source-to-outcome predictions did not agree particularly well with the data. The Panel also commented on DAS' omission of data at the limit of detection (LOD) in the comparisons. Much of the TCPy data in NHANES are below the LOD. Inclusion of the LOD values for NHANES would offer a clearer comparison. Results from two smaller studies, Curwin et al. (2007) and Lu et al. (2008), were used to compare the source-to-outcome model results with field data. While the Panel concluded that the method of comparison itself seemed valid, the conclusion that the predictions were consistent with observed values was overstated. The Panel also found that the DAS comparison of the model to the published data of TCPy in urine was cursory. The urinary concentrations of TCPy from the individuals in the Curwin et al. (2007) study were not appropriate to compare with model predictions because exposures other than dietary were included.

The Panel thought the source-to-outcome model was inadequate to predict urinary TCPy values that could be compared to the published data. Considering that the PBPK-PD model had been calibrated thoroughly, one would expect the model outputs, in terms of chlorpyrifos, chlorpyrifos-oxon, and TCPy in blood and AChE inhibition, to resemble the data from human volunteer exposure studies (see response to Charge Question 3). Therefore, the inability of the CARES dietary exposure model to generate representative input data significantly reduces confidence in the predictability of the PBPK-PD model. The Panel reported on recently published data from Lu et al. (2010) that could be used as potential input values in the PBPK-PD model to assess how the model would perform compared to the observed values.

Several Panel members expressed concern with using a simple multiplicative factor to estimate TCPy levels in pregnant women. The Panel consensus was that using a simple multiplicative factor might provide inaccurate and imprecise TCPy levels in pregnant women.

Concern was expressed with regard to the focus on dietary exposures to chlorpyrifos through the diet to the exclusion of other pathways, and to the direct intake and subsequent excretion of TCPy in the diet itself. The Panel indicated that more work in this area of the model is important. An additional concern focused on the comparison with occupational data. The route of exposure in occupationally-exposed individuals would likely be quite different from the dietary pathway emphasized in the remainder of the report. As a consequence, this analysis may be inappropriate as a demonstration of the validity of the model.

Charge Question 5: Sensitivity Analyses, Variability, and Uncertainty

QUESTION 5.1: *The four step procedure described above was intended to permit DAS to focus on the factors that were most important in determining variation in response. Please comment on the methods used by DAS to assess variation in response (e.g., identification of sensitive factors and collection and integration of empirical data on variation). Please discuss the extent to which the methods described are appropriate and complete?*

QUESTION 5.2: *During the Scientific Advisory Panel meeting held in July 2010 in which EPA's Office of Research and Development SHEDS/PBPK model was presented (see http://www.epa.gov/scipoly/sap/meetings/2010/072010meeting.html), the Panel reviewed ORD's Bayesian Approach to quantitative uncertainty analysis. DAS in its source-to-outcome model did not attempt to perform a formal quantitative uncertainty analysis (QUA), but instead evaluated components of the model by performing model-to-model comparisons (e.g., multiple dietary exposure models and multiple models of longitudinal exposures) and by performing model-to-measurement comparisons for internal dose (e.g., chlorpyrifos in blood and TCPy in urine) and AChE in blood and plasma. Since formal Bayesian QUAs are only rarely conducted, are there other methods (short of rigorous Bayesian approaches) that the SAP can recommend for characterizing uncertainty due to limited data?*

Panel Response Summary

The Panel addressed Charge Questions 5.1 and 5.2 together in one unified response.

The Panel concluded that the sensitivity, variability, and uncertainty analyses were incomplete as outlined in the DAS report, but emphasized the work represented a reasonable starting point. The Panel agreed that a systematic and explicit characterization of both variability and uncertainty, separately and for all individual components/steps of the exposure-to-biomarker modeling system, is needed and would substantially improve the analysis presented in the report. Implementing Bayesian analytical methods to enhance the study under review was also strongly recommended by the Panel, though there was a wide range of suggestions regarding the extent and the specifics of such an implementation. The Panel recognized that gaps in the data make any analysis difficult, including employing the aforementioned Bayesian techniques. The Panel agreed that, despite gaps in the data, it is nonetheless valuable to also analyze components of the models when possible. The Panel also agreed that local sensitivity analysis was inadequate and recommended exploring and applying methods for computationally-efficient global sensitivity analysis. The Panel recommended a larger systematic effort to simplify the model using known assumptions, processes, and interactions in both the exposure and the biological (pharmacokinetic and pharmacodynamic) models.

Charge Question 6: Calculating Data-Derived Extrapolation Factors

QUESTION 6.1: *Please comment on the strengths and limitations of DAS's proposed approach to estimate an animal-to-human extrapolation factor as described in Section 10. The Agency is concerned about a component of the proposed approach involving use of the study design characteristics of a single animal study (See Section 10). The Agency generally recommends a weight of the evidence approach for determining toxicological points of departure when multiple studies are available. For chlorpyrifos, there is a large database of study results performed across different animal life stages and so it may be more appropriate to determine points of departure and extrapolation factors based on an integrated analysis of these multiple studies instead of based on a single study. In your response, please comment on the approach proposed and provide guidance and suggestions for alternative approaches, if appropriate.*

<u>Panel Response Summary</u>

The Panel agreed with the general principle of the weight of evidence approach as commonly applied in many areas of risk assessment. The advantage of the weight of evidence approach is to provide more robust support for a range of uncertainty factors (UFs). However, the Panel expects that when a well-evaluated PBPK/PD animal-to-human model is ready for use in risk assessment, the use of UF_A would no longer be needed.

The UF_A proposed by DAS is based on the single study (Marty and Andrus 2010) with PND11 and adult rats. The Panel noted that while the data from the Marty and Andrus study may be more suitable for conducting benchmark dose analysis than many other studies, there were also limitations in the study's possible application to other toxicity endpoints and human exposure scenarios.

1) The peak inhibitions increased over time in a tissue- and dose-dependent manner based on the modeled pattern for RBC AChE and the steady state was not achieved until many days after repeated dosing.

2) The DAS UF_A only pertains to the dietary exposure estimated from a model that did not include the prenatal period or infants with lactational exposures to chlorpyrifos.
3) The current dietary model needs to fully capture beverage and water consumption.

The Panel was critical of the DAS approach to estimate human variability with eight-person groups and proposed an alternative approach without regrouping. The Panel indicated that the $BMDL_{10}$ can be best estimated by applying the true definition of this concept directly to the results of the combined CARES PBPK/PD model. Once a human $BMDL_{10}$ estimate is predicted via the variability and uncertainty component of the model, it can be compared to the animal $BMDL_{10}$ estimate. The UF_A can then be derived by comparing the animal and human $BMDL_{10}$, both of which are derived from the same model, but with species-specific parameters. This approach uses the model as it was intended, as a tool to integrate the variability and uncertainty in the process of incorporating *in vitro* and *in vivo* animal data to prediction of risks in the human population.

QUESTION 6.2: *Please comment on the strengths and limitations of proposed approach to estimate a human variability extrapolation factor as described in Section 10. Please provide alternative approaches, if appropriate.*

Panel Response Summary

The Panel stated that the model-derived human variability uncertainty factor (UF_H) is estimated not from the uncertainty of the modeled mean response, but from the individual-to-individual variability.

The Panel focused on different areas of model and data uncertainty. Knowledge of target site-specific chemical levels, e.g., concentrations at the active sites of biotransformation enzymes within the endoplasmic reticulum vs. concentrations within the organ as a whole, is limited. This lack of knowledge is a significant cause of uncertainty in predicting the concentrations of the chemical actually available for metabolism. Extrapolating from those *in vitro* conditions to estimating the biotransformation rates in the intact human tissues therefore has considerable uncertainty. In addition, the Panel recognized that limited data from younger individuals is an important source of uncertainty. The use of small numbers of human tissue samples to estimate human-specific metabolic rates, in particular, in very young individuals where the greatest differences from older age groups may exist, raises the uncertainty in the values of model inputs.

The Panel recommended that the current case study analysis be expanded to accurately predict all pertinent human exposures for target populations of interest (e.g., early life stages), groups with specific exposure characteristics (e.g., farm workers and their families), and multiple routes of exposure (e.g., inhalation, dermal, ingestion from food, water, hand-to-mouth activities), These are all important factors for fully and adequately addressing the inter-individual variability and uncertainty. DAS employed only two immature age-groups (six months and three years) in their modeling exercise.

The Panel noted that the possibility of non-cholinergic neurodevelopmental disorders should also be considered in the future as a basis for pharmacodynamic differences in infants and children.

QUESTION 6.3: *The current effort by DAS is limited to food exposure and does not include all relevant exposure routes. Please comment on strengths and weaknesses of using data-derived extrapolation factors described in the current effort for life stages (e.g., gestation or pregnancy) and/or for routes (dermal, inhalation) not considered in the current modeling effort.*

Panel Response Summary

The Panel recommended that data are needed to characterize all sources of exposure.
The current DAS model should be expanded to accommodate all pertinent exposure scenarios (ingestion, dermal, and inhalation) and various life stages or conditions (pregnancy and gestation, lactation, and nursing infants).

DETAILED RESPONSES TO CHARGE QUESTIONS

Charge Question 1: Physiologically Based Pharmacokinetic/Pharmacodynamic (PBPK/PD) Modeling

Question 1.1: Model Structure. *Please comment on the structure of the chlorpyrifos PBPK/PD model with specific consideration of the mechanistic basis for the acetylcholinesterase (AChE) inhibiting mode of action. Please include in your comments consideration of age-dependant metabolism and the proposed approach to assess human variability.*

Panel Response

With certain reservations, the Panel found that the mechanistic basis for the AChE-inhibiting mode of action of chlorpyrifos had been satisfactorily incorporated into the PBPK/PD model, noting that the model did not address non-cholinergic effects of chlorpyrifos. The Panel generally supported the focus on inhibition of AChE and red blood cell (RBC) AChE as the most suitable marker and surrogate for the true target of the anticholinesterase effect of chlorpyrifos, specifically synaptic AChE in brain, spinal cord, and peripheral nervous systems. However, the Panel identified several concerns with how DAS handled age-dependent metabolism in the model and the proposed approach to assess human variability. Among other factors, limited numbers of samples to study age-dependent human metabolism (in particular with very young individuals) and the use of non-physiological conditions to estimate metabolic rates, limited the Panel's confidence in that component of the model. There were also concerns about the parameterization of the model based on the unpublished age-dependent metabolism study (Smith et al., 2011, Attachment A, DAS, 2011), especially with respect to paraoxonase (PON), the enzyme that hydrolyzes chlorpyrifos-oxon.

The Panel had difficulty in evaluating the DAS model because the pharmacokinetic and pharmacodynamic components were grouped together in a way that made it hard to assess the validity of the structure and parameters of each component. The Panel recommended that DAS provide the source of and details about all of the model parameters in an effort to be more transparent and complete. Such information would significantly raise the general understanding of and confidence in the model. Overall, the model would be greatly improved if the parameterization for PON was based on more accurate and relevant experimental data.

General Considerations

Toxicity testing in the 21st century. The Panel's responses to the charge questions are founded in the vision and recommendations embodied in the National Research Council's (NRC) report, *Toxicity Testing in the 21st Century: A Vision and a Strategy* (NRC, 2007), which was completed in response to a EPA request. The DAS modeling efforts under consideration by the SAP represent a major step toward implementing the recommendations of the NRC's report. Understanding the relevance of the NRC report to this meeting is therefore important. In brief, the NRC Committee of Testing and Assessment of Environmental Agents (NRC committee) recommended the use of a new toxicity system that uses cutting-edge technologies such as high-throughput chemical screening and research on substances' interactions with genetic material to predict the toxicity pathway(s) of a chemical rather than use expensive and time-consuming *in*

ED_005731_00001263-00021

vivo animal studies (to extrapolate to humans) as currently required for human health risk assessment of pesticides. The proposed system includes both the assessment of toxicity pathways and "targeted testing," which is designed to clarify and refine information from toxicity pathway tests for use in chemical risk assessments. Other key elements are dose-response and extrapolation modeling, chemical characterization, population and exposure data, and risk contexts. The following is a brief summary of each element found in the NRC report (NRC, 2007). All parts of the toxicity assessment system are inter-related. Dose-response and extrapolation (exposure) modeling, e.g., Dow's LifeStage PBPK/PD model linked to a probabilistic exposure model such as CARES, would be used to estimate environmental exposures that would lead to significant perturbations of toxicity pathways observed in the cellular tests. High quality population and human exposure data (e.g., biomonitoring data) are needed as inputs into dose-response and exposure models. The risk context is used to determine when high-throughput screening is needed and when target toxicity testing is needed. This step would reduce the need for a lengthy step-wise process of chemical characterization to testing to dose-response modeling. Accurately describing the normal physiology of an organism and predicting responses to perturbations, along with ultimate outcomes, are prime goals of modeling and experimental efforts. Bio-mathematical models, such as Dow's PBPK/PD model, need to incorporate relevant pharmacokinetic and pharmacodynamic factors implicated in the mechanism of toxicity and specific toxicity pathway. The PBPK/PD model would be linked to an exposure model that uses high quality population and human exposure data. As a result of the NRC report, EPA established a National Center for Computational Toxicology that is developing new software and methods for predictive toxicology. The National Institute of Environmental Health Sciences, through the National Toxicology Program's Roadmap for the Future, initiated a partnership with the Chemical Genomics Center of the National Institutes of Health to develop and perform high- and medium-throughput screening assays to test more chemicals in less time and at lower costs. The Panel referred to the NRC report many times throughout the meeting because of its relevance to all of the charge questions.

Chlorpyrifos is a data-rich chemical. The Panel noted that there is an exceptionally rich database on chlorpyrifos based on decades of biochemical and toxicological studies. These studies show that the active metabolite of chlorpyrifos, chlorpyrifos-oxon, directly phosphorylates the active sites of several B-esterases including acetylcholinesterase (AChE). This reaction is the basis for the desired insecticidal event, quasi-irreversible inhibition of pest AChE, which is followed by death. Although other pest enzymes also may be affected, there is no indication that such reactions contribute meaningfully to the pesticidal action. In that sense AChE can be considered as the "target enzyme." Mammalian toxicity flows directly from inhibition of the same enzyme (i.e., AChE is the target enzyme in "off target" species as well), which is an "apical event" associated with various undesirable effects collectively defined as a hyper-cholinergic syndrome. Chlorpyrifos may generate additional apical events by inhibiting different B-esterases or other non-cholinergic events.

Dow's chlorpyrifos LifeStage PBPK/PD model. The current LifeStage PBPK/PD model for chlorpyrifos is the product of over a decade of effort by Timchalk and his research group at Battelle/Pacific Northwest Research laboratories. His original model for rats was extrapolated to predict the time courses of blood chlorpyrifos, chlorpyrifos-oxon and a metabolite, TCPy, as well as inhibition of plasma butyrylcholinesterase (BChE), and RBC and brain AChE in adult rats and

humans. The model was then adapted to pre-weanling rats by utilizing measured and allometrically scaled metabolic rate constants. The age-dependence of chlorpyrifos metabolic activation and inactivation by human liver and/or plasma was quantified subsequently and incorporated into the PBPK/PD "Life-Stage model" that was utilized in the currently proposed, exposure dependent risk assessment. The Panel commended the investigators for their modeling efforts to simulate chlorpyrifos dosimetry and effects on infants and children and noted how similar efforts could be extended to simulate dosimetry and effects of any chemical in infants and children.

Mechanistic Basis for Mode of Action

Acetylcholinesterase inhibition is a well-known and measurable effect of organophosphate insecticides and can lead to most of the toxic effects that follow chlorpyrifos exposure (Taylor, 2001). Some of these effects are likely to result from reduced activity of synaptic AChE in brain; others may reflect inhibition of the same enzyme autonomic ganglia and terminals of autonomic and motor nerves. AChE in RBCs is the same gene product as the target AChE in the brain, so AChE inhibition in both brain and RBCs is a logical endpoint for measurement. In practice, RBC AChE inhibition has long been used as a biomarker of exposure, because it is less convenient to sample true target tissues in animals and inappropriate or impossible in humans. With respect to the mechanistic basis for the mode of action there was general support on the Panel for the focus on inhibition of AChE, and with using RBC AChE as the most suitable marker and surrogate for the true target of the cholinergic effect of chlorpyrifos, i.e., synaptic AChE in the central and peripheral nervous systems. Nonetheless, the Panel wished to emphasize the point that inhibition of red cell AChE is most properly viewed as a biomarker of exposure and NOT as a measure of effect.

Plasma BChE is more sensitive than AChE in both animal and human studies, but BChE is subject to more individual variability as a function of body weight, height, sex, and genotype (Sidell and Kaminskis, 1975; Brock and Brock, 1990). In contrast, RBC AChE has negligible variation (Lefkowitz et al., 2007), which strengthens the structure of the model by providing a fairly stable indicator of toxicity. The Panel stated that it was appropriate and desirable for the model to include plasma BChE and RBC AChE inhibition as outputs so there would be a sensitive indicator of exposure, plasma BChE inhibition, in addition to a relevant surrogate marker for toxicity, RBC AChE inhibition.

The Panel had difficulty in evaluating the incorporation of the AChE inhibition in the DAS model because the pharmacokinetic and pharmacodynamic components were not presented separately, and therefore, the validity of the structure and parameters for each component could not be assessed. In particular, insufficient information was given concerning the source and variability of pharmacodynamic parameters such as the rate of AChE synthesis, AChE degradation, and aging of the enzyme-substrate complex. Neither the DAS presentations nor report (DAS, 2011) explained whether these parameters were estimated from a previous model of *in vivo* modeling exercises or were drawn from actual *in vitro* experiments. If the DAS source-to-outcome model for chlorpyrifos is to serve as a case study on how to implement the concepts outlined in the NRC report, then the Panel recommended that more transparency and completeness be provided concerning the source and details for all model parameters, of which

there are 200 or more. This would include the mean, median, and "normal" range values for all parameters. Citations of the research that established these values would also be needed. Such information would significantly improve general understanding and confidence in the model.

Age-dependent Metabolism

The Panel reviewed the DAS unpublished *in vitro* age-dependent chlorpyrifos metabolism study authored by Smith et al. (provided as Attachment A, DAS, 2011). Results from this study revealed no apparent age-dependency of chlorpyrifos dearylation and desulfuration when data were expressed on a per mg protein basis. Likewise there appeared to be no change in ratio of activity between these pathways as maturation progressed. On the other hand, when the data were analyzed in terms of total reaction rates per sample, age became a major factor. This study confirmed what has been observed in other published studies that when the great increase in liver mass during early life is factored into the equation, the metabolism of chlorpyrifos-oxon by PON rises substantially (> 3.5 fold) from six months to adulthood. In other words, chlorpyrifos metabolism showed a powerful age dependency, which appeared to depend entirely on the sheer abundance of hepatic enzymes, including PON. This dependency was factored into the DAS model.

The Panel identified several specific concerns with the parameterization of the PBPK/PD Life Stage model based on Smith et al. (2011).

- The DAS study was based on a small number of samples, which was considered by many Panel members to be a major limitation, even after recognizing the difficulty of obtaining multiple tissue samples at younger ages.

- One panel member stressed that the known genetic polymorphisms in several of the CYP enzymes involved in chlorpyrifos metabolism (e.g., CYP2D6), and the functional polymorphisms in PON will all contribute to variability in chlorpyrifos metabolism. This makes it especially important to consider genotype and phenotype in datasets of small to modest size when assessing age-related differences in metabolism.

- The Vmax values for PON were determined *in vitro* at a pH of 8.5, which is well above the physiological pH of 7.4. A higher pH could have altered the structure and function of PON. Thus, using those Vmax values in the model would lead to errors in estimating the detoxication contribution of that enzyme.

- In addition, the computed Vmax values were attributed to the detoxication contribution of PON alone. Acid-denatured blanks do not distinguish PON-mediated hydrolysis from hydrolysis mediated by B-esterases (such as BChE), liver carboxylesterases, and plasma and liver cholinesterases, as well as any other alternative binding sites. Thus, the Vmax values used in the model are actually a combination of the PON hydrolysis of chlorpyrifos-oxon along with the stoichiometric binding of chlorpyrifos-oxon to the B-esterases. This also results in an erroneous estimate of the detoxication contribution of PON. A blank utilizing EDTA to chelate calcium, the PON co-factor, would have allowed the PON-mediated hydrolysis to be distinguished from the non-PON-mediated

hydrolysis, and this was not done. The brief incubation time used, i.e., 5 minutes, would have allowed a large contribution of product from stoichiometric contributions from these B-esterases. Therefore the reported Km and Vmax values are artificially inflated through the contributions of the stoichiometric detoxication along with the catalytic detoxication.

- As recognized by Timchalk et al. (2002), the stoichiometric binding of chlorpyrifos-oxon to the B-esterases is an important detoxication mechanism. This binding was incorporated into the 2002 model and was also incorporated into the present model as independent parameters (indicated in Table 1 of the DAS report). However, since the paraoxonase assay method used in the Smith et al. (2011) manuscript failed to effectively eliminate the stoichiometric contributions of the B-esterases, as discussed in the point above, the B-esterase contributions to detoxication were also incorporated into the model as a part of the Vmax value for PON. Therefore, the detoxication contribution of the B-esterases has been incorporated into the model twice, which will lead to overestimation of the role of the B-esterases.

- Very high chlorpyrifos-oxon concentrations were used in the *in vitro* assays (i.e., 72-2400 micromolar) because of the insensitivity of the endpoint of the assay (i.e., spectrophotometric detection of TCPy, one of the hydrolysis products). These are not concentrations that could be realistically achieved *in vivo*. These high concentrations have also contributed to overestimations of Km and Vmax. Using these Km and Vmax values would have led to the conclusion that PON activity is the driver of the variability in the model. That conclusion should be reevaluated with more accurate PON kinetic parameters.

- Most of the chlorpyrifos-oxon arises from the relatively inefficient cytochromes P450 (CYPs), which can only deliver substrate (i.e., chlorpyrifos-oxon) to PON after the desulfuration of chlorpyrifos. Since PON is the second enzyme in a two-enzyme pathway, PON cannot produce product any faster than substrate is delivered to it. It is highly unlikely that the *in vitro* concentrations assayed in the Smith et al. (2011) unpublished manuscript or the Km calculated (520 μM) could be provided to PON by the CYP's in any *in vivo* situation. Even though the PON Vmax is theoretically high, it is not likely to be active at that rate because of limited substrate availability. The Panel was unclear whether the CYP and PON reactions were coupled in the model to prevent inaccurate assessment of PON's theoretical velocity. Therefore, it is possible that PON may be the driver only under unrealistic concentrations. Reassessment of the model with more accurate kinetic parameters is needed to re-evaluate this conclusion.

- One Panel member showed data arising from a more sensitive *in vitro* assay technique using a low chlorpyrifos-oxon concentration of less than 200 nM in human serum. If the proposed DAS model is accurate, this nM concentration would slightly exceed the high end of exposure projections for humans. These unpublished results demonstrated that at this lower nM substrate concentration, neither the Q192R genotype nor the PON phenotype (i.e., overall PON activity level with paraoxon or diazoxon as substrates) seemed to influence appreciably the level of chlorpyrifos-oxon hydrolysis.

- To determine the "functional genotype" for the Q192R polymorphism in human serum, 2M NaCl is routinely used in the *in vitro* assay to elevate PON activity and separate the genotypes more fully in a plot of paraoxon hydrolysis vs. diazoxon hydrolysis. The presence of 2M NaCl elevates paraoxon hydrolysis by about 25-50% and elevates diazoxon hydrolysis about 4-fold; it is unknown how much chlorpyrifos-oxon hydrolysis would be elevated by salt. Nevertheless this high concentration of salt should not be used in an assay to determine overall PON velocity and the kinetic parameters for the PBPK model. Use of high salt in the Smith et al. (2011) study probably led to falsely high rates of chlorpyrifos-oxon hydrolysis.

- Normalizing CYP activities to protein in liver and brain to compare the two tissues for kinetic parameters would not be accurate because the microsomal protein content in these tissues will differ. Since serum PON is believed to come from the liver, it is unclear why the serum PON would be age-dependent and the liver PON not age-dependent. Further investigation should be considered to examine the age dependency or the lack thereof in these two tissues.

- Another limitation of the data set is a lack of information on the health, nutritional status, and recent exposures (e.g., pharmaceuticals, alcohol, diet, other xenobiotics) of the donors, or the condition/viability of the tissue itself. Together these limitations generate questions concerning the metabolic activation and inactivation of chlorpyrifos by human infants. Greater immaturity generally confers greater susceptibility to chemical toxicity. Neonates and infants younger than 6 months would be anticipated to be most sensitive to chlorpyrifos acute toxicity and potential neuro-developmental disturbances. Low body fat, low plasma protein binding capacity, an immature blood-brain barrier, inefficient detoxication, nursing from chlorpyrifos-exposed mother, etc., could predispose neonates and infants to injury. Toddlers and other young children may also be at higher risk. Additional literature sources should be consulted for accurate values for these parameters for input into the model.

Other age-related concerns

The stoichiometric binding of any organophosphate compound to B-esterases is an important detoxication mechanism and is a contributing factor in the age-related toxicity of many organophosphate compounds including chlorpyrifos. While the model does address the detoxication contribution of the B-esterases, it does not consider any age-related variability for the B-esterases. With respect to the liver carboxylesterases, multiple animal studies report that juveniles have lower levels than adults. This appears to be the case in human liver as well (Pope et al., 2005; Yang et al., 2009). With respect to human plasma BChE, there is variability in adults, as indicated above, but it is not clear how the level of plasma BChE changes with age.

The Panel indicated that models should span the entire life-stage of humans to be useful for risk assessments that address concerns for vulnerable lifestages and subpopulations. The Panel stated that gestation and neonatal life required explicit attention in the model. The DAS model does not consider pregnancy or early neonatal life (e.g., potential exposure via mothers' milk). The Panel stressed that this omission would be a glaring oversight given the mission of the National

Children's Study (Hirschfeld, et al., 2011), which is to determine how early life exposures to toxicants translate into later effects in life. The DAS report, however, reported that fat volume and blood volume strongly influenced predicted levels of RBC AChE inhibition (DAS, 2011). Research by Pavek et al. (2009) indicated that these parameters change during pregnancy, which supports the need for explicit consideration of this important life-stage. In addition, pharmacokinetic models should also incorporate sex-specific physiological parameters instead of lumping males and females together, as done here. The absence of gestational and lactational stages is a major deficit that prevents the chlorpyrifos PBPK/PD model from being an optimal prototype for other risk assessments as the Agency moves forward in the new toxicity paradigm.

The Panel offered specific comments with regard to estimation of age-related changes in fat volume, noting that the presentation cited three sources for the values of that parameter (Butte et al. 2000, Lafortuna et al. 2005, and Valentin et al. 2002). One of these, Valentin et al. (2002), could not be accessed as reference details were missing. The data from the other two studies were deemed inadequate to represent the human population across all life-stages. Butte et al. (2000) did cover certain early life-stages, but only examined children at six months, 12 months and 24 months. It used the dual energy X-ray absorptiometry (DXA) technique to estimate the Bone Mineral Density and then computed fat-free mass versus fat mass. Lafortuna et al (2005) examined normal weight adults in the range of 30 ± 5 yrs and obese adults in the range of 20 ± 7 yrs. This study used the bioimpedance method to estimate the participant's percentage body fat. These two methods of determining body fat have different degrees of accuracy and precision (Norgan, 2005). Therefore, the fat values in the data may not faithfully represent variability in the population.

The Panel suggested that the potential influence of reduced plasma protein binding in neonates and infants should be examined. Lowe et al. (2009) found that chlorpyrifos was highly (98%) bound to serum albumin (a surrogate protein for non-lipophilic blood proteins) in both rats and humans. These authors demonstrated that the lipid-containing fraction of whole blood increases through late pregnancy and by extension in the fetus.

Assessment of Variability

Model predictions of cholinesterase inhibition include many additional factors such as enzyme turnover rates (tissue specific synthesis and loss), enzyme catalytic rates, rates of dearylation, reactivation rates, and aging rates. From a broad perspective, the Panel acknowledged the DAS had considered a number of factors that affect the rate of chlorpyrifos-oxon formation and metabolism in its "Variation PBPK/PD model." Predictions of the model provide a fair match to data from typical adults, according to the results of Nolan et al (1984). These findings were incorporated into the LifeStage PBPK/PD model for chlorpyrifos. The model output is mainly consistent with other relevant experimental databases. Some significant deviations occur at higher dose levels (e.g., 2 mg/kg). These deviations could be considered "conservative" in the sense that they either match the data or predict higher levels of active compound and higher levels of cholinesterase inhibition than actually observed. One panelist commented on the judicious use of subjective terms like "conservative" since we do not know the full possible uses of the model. The Panel indicated that the deviations probably represented a systematic error of

some sort, possibly the over-emphasized contributions of the B-esterases pointed out in the previous section of this response.

The "Variation PBPK/PD model" used a very limited sensitivity analysis to identify parameters in the LifeStage model that drive variation in response. There was some sense among the Panel that a more rigorous and wider ranging sensitivity analysis would be well justified. Model parameters varied in the initial DAS study were tissue volumes, blood flows and perfusion rates, enzyme activity levels, tissue-partitioning coefficients, binding of cholinesterases, recovery and regeneration, as well as rates of transport of chlorpyrifos from stomach to gut or uptake from gut. This analysis showed that the most important parameters were brain and liver CYP metabolism and their effect on rates of formation of chlorpyrifos-oxon and TCPy, which depended strongly on PON activity in blood and liver, etc. In fact, it appears that PON activity is the major driver of downstream outcomes including the magnitude of AChE inhibition in blood and brain, although there are flaws in the data supporting that conclusion (see earlier discussion). One panel member also mentioned that the chlorpyrifos brain: blood partition coefficient is a sensitive parameter that has been overlooked. Variability of this parameter should have been incorporated to assess variability in brain AChE inhibition. The variability of this parameter may have a significant impact on overall variability in toxicity and estimated uncertainty factors.

Multiple panel members drew attention to the fact that the assessment of variability in the DAS model was only concerned with pharmacokinetic parameters. Pharmacodynamic parameters are likely to be equally important. In particular, variability in rates of AChE (and BChE) synthesis, degradation, regeneration, and aging rates could well impact overall variability in AChE inhibition and, hence, toxic outcomes in the general population. In a PBPK/PD modeling study of diisopropylfluorophosphate (DFP) intoxication, Chen et al. (2009) demonstrated through a global sensitivity analysis that these pharmacodynamic parameters had the greatest influence on brain AChE activity. Nong et al. (2008) used a Bayesian approach (i.e., Markov Chain Monte Carlo methods) to calibrate parameters for a carbaryl PBPK/PD model. This calibration yielded a range of AChE inhibition and regeneration rate constants.

In light of these considerations and the background information, the Panel was reluctant to embrace the conclusions in the DAS report that extrapolation from animals to humans and from adults to children and infants can utilize very small uncertainty factors, on the order of two-fold or less. This point is further considered in the Panel response to Question 6.

Question 1.2. Dose Metrics: *In Section 6, DAS proposes a number of dose metrics for use in the PBPK/PD effort. These include peak levels of blood and brain AChE inhibition peak blood and brain levels of chlorpyrifos and its oxon metabolite, and urinary measures of 3,5,6-trichloro-2-pyridinol TCPy), a chlorpyrifos metabolite. Please comment on the utility, strengths, and limitations of these proposed internal dose metrics.*

Panel Response

The Panel indicated that all of the named parameters could be supported scientifically as metrics for chlorpyrifos exposure with some limitations. Peak brain AChE inhibition should be an indicator of potential toxicity, but obviously cannot be measured in humans. Instead peak blood

AChE inhibition, which is highly variable among people so results must be interpreted with this variability in mind, is a potential surrogate marker for target tissues. Peak chlorpyrifos concentrations may be useful for acute exposures, but the area under the curve may be more useful than the peak level for repeated exposures. While the peak chlorpyrifos-oxon concentration is potentially useful because chlorpyrifos-oxon is the active metabolite and the model could predict values for this parameter, chlorpyrifos-oxon is highly reactive and rarely measurable; therefore, there would be little confidence in any concentrations measured. Urinary TCPy can be biased by pass-through TCPy that is absorbed as an environmental degradate; therefore, these measurements should be interpreted with caution.

Subsequent discussion centered on the following three dose metrics: 1) the potential value of AUC determinations as a supplement (not replacement) for peak level measurements, 2) the value of tracking BChE inhibition, and 3) the utility of TCPy measurements.

Potential value of AUC determinations

The proposed internal dose metrics included peak levels of blood and brain AChE inhibition, peak blood and brain levels of chlorpyrifos and chlorpyrifos-oxon, along with urinary measures of TCPy. These were all appropriate, but some Panel members, who were concerned with the possible risks of sustained low-level exposure, suggested that area under the curve (AUC) estimates would also be valuable. The scale used for AUC estimates should not go to infinity, but rather be truncated to match the period of significant effects. The Panel voiced the opinion that it is not necessary at this early stage of model development to narrow down the dose metrics in terms of AUC or peak values. Rather, at the model validation stage, the dose metrics will be properly focused on the characteristics of data that are available for the purpose. The basic expectation is that a sufficiently validated model would be able to predict those parameters that are of biological and toxicological importance, which cannot be measured accurately in experiments or surveys. At this stage, the model construct should be versatile for such future, realistic, and reasonable uses. While AUC is harder to measure accurately, it should be noted that in human studies there often is no way to be certain that peak inhibitions are captured. Thus, it is premature to limit the model development to peak concentrations and effects.

The Panel also suggested that the model would be very useful in looking at derived metrics, metrics that are not directly measurable but which can be obtained from the final fitted model. The question is which of the many derived metrics would be correlated to key health effects. The Panel recommended that the Agency should look not just at 'dose metrics,' which help to validate a model (e.g., by comparing model predictions to measurable quantities), but look also to 'verification metrics' to help confirm that the model is performing as it was intended. Metrics for verification could include time-to-peak time to reach 90% of peak, or time for nearly complete elimination (e.g., 90% elimination).

Value of BChE Inhibition as a Dose Metric

Many on the Panel argued that BChE inhibition should be tracked in addition to AChE inhibition. This proposal generated lively discussion. Some members took the view that plasma BChE activity is inherently more variable than RBC AChE activity. This variability may stem

ED_005731_00001263-00029

from at least two different sources: 1) genetic variants with altered catalytic properties that occur as balanced polymorphisms in most populations and 2) day-to-day regulation in response to hormones and other factors. For these reasons it is likely that isolated measurements of BChE activity (i.e., random samples from a population) will show substantial variance. It was also pointed out that BChE is probably best seen as a sentinel or surrogate marker for toxic effects since this enzyme has no identified physiological function other than an auxiliary role in acetylcholine hydrolysis and a metabolic role in detoxication of bio-active esters in the diet. Nonetheless, Panel attention was drawn to two reasons why BChE activity may be a useful and even important metric for a PBPK/PD model of chlorpyrifos. First, as the DAS report demonstrates, BChE is critically important in detoxication of chlorpyrifos-oxon. This role does not depend on an ability to catalyze the hydrolysis of the activated compound, but on irreversible stoichiometric binding. The binding is extensive enough to be functionally significant because plasma BChE levels are substantial (about 3 mg/L) and because BChE is a highly sensitive target. A second and closely related reason is that this high sensitivity to chlorpyrifos-oxon means that BChE can be used as a surrogate for other "non-cholinergic" targets that are difficult or impossible to measure, or even unknown.

Another Panel member proposed that, with the well-documented concern for the role of BChE in neurodevelopment, it is unnecessary to preclude the validation of this parameter during model development (to the extent that data are available). When the model is ready for risk assessment applications, all mechanistic data should be assembled. Then, the most appropriate parameters should be chosen based on the concerns for BChE's role in neurodevelopment. In these considerations, the complexity and variability of human plasma BChE would be taken into account together with many other factors. The key to model validation is to make use of as much data available as possible, not to exclude them.

Utility of TCPy data

For dose metric comparison with experimental data, the issue of ready detection needs to be balanced against the relevance to toxic effects. While the urine samples are easy to collect and easy to measure for TCPy, the toxicological relevance of urinary TCPy for the present purpose is highly questionable. Urinary TCPy is not a very useful dose metric because only a portion comes from dietary exposure to chlorpyrifos, for example, 5-20% as reported by Wilson et al. (2003) and 30-35% as reported by Lu et al. (2005). The remainder comes directly from the diet and represents chlorpyrifos breakdown products in food. Obviously the most relevant endpoints are the brain concentrations of the oxon metabolite or AChE inhibition in brain, which is the primary target organ of interest. These measurements are not feasible in humans and hence unavailable as dose metrics. Therefore, the blood concentrations of chlorpyrifos and its oxon metabolite or AChE inhibition in blood become the relevant internal metrics by default. Direct measures of chlorpyrifos and, especially, its oxon metabolite, are often limited by the sensitivity of analytical methods, so at present it would be difficult to rely on these measures as primary dose metrics. Since the studies presented by Marty and Andrus (2010) showed similar dose response profiles for the inhibition of AChE in blood and brain, relying on AChE inhibition in RBC as a dose metric is appropriate.

ED_005731_00001263-00030

Question 1.3: *Although the current modeling effort focuses on AChE inhibition, there are data to suggest that chlorpyrifos exposure may also result in non-cholinergic effects. As such, and if appropriate, please provide plausible additional and/or alternative dose metrics such as area under the curve (AUC) metrics or other temporal-based internal dose metrics that may be appropriate for evaluating potential non-cholinergic effects.*

Panel Response

The Panel thought it reasonable to consider non-cholinergic effects in the long term; however, the current model structure is specifically designed for studying cholinergic effects. Obtaining model output for AUC for chlorpyrifos and chlorpyrifos-oxon will be useful for use in simulating non-cholinergic effects. Modeling BChE inhibition may also be useful as a potential measure of non-cholinergic toxicity. While non-cholinergic effects may be of high interest and importance, there are no standard methods to measure non-cholinergic activity and no non-cholinergic mechanism has yet been validated.

Evidence for non-cholinergic effects includes the discovery that tubulin can be phosphorylated by chlorpyrifos-oxon, with resulting dysfunction of microtubules and the axonal cytoskeleton (Jiang et al. 2010). This work suggests that convalent binding of the Chlorpyrifos-oxon to tubulin and tubulin-associated proteins is a potential mechanism of neurotoxicity. Yang et al. (2008) showed that chlorpyrifos and chlorpyrifos-oxon would inhibit axonal outgrowth of sensory neurons in tissue culture under conditions in which there was no detectable inhibition of AChE activity. Their conclusion was that inhibition of axonal growth by OPs requires AChE, but disrupts its morphogenic functions rather than its catalytic functions. Findings by both Jiang et al. (2010) and Yang et al. (2008) suggest that an exclusive focus on AChE inhibition may overlook important secondary mechanisms of neurotoxicity. That said, the Panel indicated it was premature to recommend specific metrics for such "non cholinergic" toxicity, especially because there is still no convincing proof that these sorts of effects do occur *in vivo* without accompanying cholinesterase inhibition.

The Panel discussed three points worth considering when gathering evidence for non-cholinergic effects.

1) Panel members pointed out that obtaining AUC for chlorpyrifos and chlorpyrifos-oxon will be useful in modeling and predicting non-cholinergic effects when and if that proves feasible and necessary. If the concentration of chlorpyrifos that must be present in a target tissue to exert a non-cholinergic effect is known, then the model can identify what exposure level will be required to obtain toxic levels in that tissue.

2) At least one Panel member pointed out that it might be worthwhile to model BChE inhibition, which is more sensitive to inhibition by chlorpyrifos, as a potential index of "off-target toxicity," despite the greater natural genetic variability in this enzyme.

3) A major concern is the enhanced susceptibility of early life stages to anti-cholinergic effects as there is rapid development of brain, spinal cord, and peripheral nervous system. In addition, there may be unique susceptibility of early life stages to non-cholinergic

effects of chlorpyrifos on the developing nervous system. Gestation and early neonatal life stages should be included in the PBPK/PD modeling of chlorpyrifos and its metabolites.

Charge Question 2: Longitudinal Dietary Exposure Assessment

QUESTION 2.1: *Please comment on the methods used by DAS to investigate the relationships between dietary exposures and levels of chlorpyrifos and chlorpyrifos-oxon in blood. Please include in your comments discussion on the strengths, weaknesses, impacts, and utility of the DAS proposal to focus on the immediate prior and subsequent days to high-end exposures and provide alternative approaches to assessing longitudinal (multi-day) exposures, if appropriate. Under what conditions and scenarios might this focus be inadvisable and/or lead to incorrect conclusions regarding exposures, doses, and risks?*

Panel Response

The Panel concluded that the scope of DAS' longitudinal dietary exposure was incomplete. It should include dietary intake over many days (not just five days) and under various, correlated, eating behaviors. The Panel recognized that limitations in the CARES and/or Lifeline model make such analyses difficult.

The Panel acknowledged the progress DAS has made in the design of the chlorpyrifos PBPK/PD model, given the complexities of assessing dietary exposures and relating the likelihood of exposure at a population level to specific toxicological endpoints. This modeling effort furthers the understanding of the relationship between exposures and toxicity, and improves risk assessment by moving away from chemical specific default values. The utility of the immediate prior and subsequent days to high-end exposures in the model simulation demonstrated changes in chlorpyrifos and chlorpyrifos-oxon levels in blood and provided insight on how high-level chlorpyrifos exposure contributes to the overall AChE inhibition in RBCs and the brain. DAS' exposure scenario coincided with a recent paper by Samsam et al. (2005), which reported on persistent cognitive impairments observed in rats after chronic dietary chlorpyrifos exposure accompanying an acute dose high enough to induce AChE inhibition. The Panel recommended to the Agency and DAS that the scope of a longitudinal dietary exposure assessment should span many days, not merely the consecutive days when measurements were made. Recent studies (Lu et al., 2008; Riederer et al., 2009) have indicated an under-estimation of cross-sectional dietary consumption (in particular to commodities frequently being detected with pesticide residues) when compared to longitudinal data, and a seasonal-dependence for dietary organophosphate pesticide exposures. Since dietary exposure and dose are related to the unknown and uncertain levels of pesticide residues in commodities as well as the probabilities of consuming those commodities, computing such random and episodic eating events would be very difficult to explain in models in a longitudinal study. Although the CARES model, and likely other models such as Lifeline, could be used as a tool for simulating consecutive days of dietary pesticide intakes, its application in the "source-to-outcome" modeling exercise highlights its significant difficulty in simulating "true longitudinal" dietary consumption patterns at the population level.

The Panel noted that the longitudinal dietary exposure estimation from CARES resulted in a pattern of sporadic peak daily exposures among a series of relatively low days of exposure. Applying this pattern to the five-day exposure history may provide a reasonable basis for evaluating longitudinal exposure, but this too contains known limitations. For example, the CARES longitudinal exposure pattern generates random draws of residue data without considering the possibility of any day-to-day correlation. The possibility of eating the same batch of food (with potentially high residue levels) for consecutive days was not computed. Additionally, individuals receiving high exposures, through unique, but specific patterns of food consumption, were not computed; therefore, the predicted exposure or dose would be inaccurately low. Some of the realistic scenarios to consider include: consecutive days of similar exposures from consuming leftovers or eating from food supplies purchased in a batch. These, and many other issues with the current approach, including the question whether a five-day chlorpyrifos exposure history is a reasonable basis for evaluating longitudinal exposure, have been the subject of many past SAP discussions. Those discussions should have been considered in this study. For example, the Panel stated that it would be premature to conclude that a five-day chlorpyrifos exposure history is a reasonable basis for evaluating longitudinal dietary exposure based on discussions held during other SAP meetings concerning probabilistic dietary exposure models (e.g., Lifeline™, Calendex™, CARES™, Stochastic Human Exposure and Dose Simulation Model for Multimedia, Multiroute/Pathway Chemicals (SHEDS-Multimedia), see http://www.epa.gov/scipoly/sap/meetings/index.htm).

The Panel pointed out that the methods outlined in Section 5 of the DAS report focused on two possible relationships between successive-day exposures with one method resulting in essentially no correlation between day-to-day exposures and the other assuming a fairly strong correlation (See Figure 44 in DAS, 2011). The panelist stated that both were incorrect and that the reality lies somewhere in between, perhaps closer to the highly correlated version than the completely random one. The CARES approach has been adopted and given perhaps too much weight in the conclusion that single-peak exposure drives everything. If day-to-day correlation were actually present, then longitudinal exposure modeling may offer an avenue for predicting continued AChE depression with associated chronic effects. These effects may differ from those found in acute "poisoning" events and may affect the pharmacokinetics and pharmacodynamics of chlorpyrifos metabolism.

The Panel suggested an alternative longitudinal dietary exposure scenario that would implement an admixture of the above-described uncorrelated and strongly correlated states (see also the discussion below on Markov chain approaches). One could envision a drift from the highly correlated to the uncorrelated states over a period of several days as illustrated in the following scenario. Assume that foodstuffs in a single purchase have unusually high levels of chlorpyrifos, and over the course of several days these foodstuffs would be consumed on a daily basis. Consider that the next purchase of foodstuffs would have low levels of chlorpyrifos. If flour was the food stuff purchased on both occasions and flour from the first purchase was mixed with flour from the second purchase then the levels of chlorpyrifos would be reduced in the flour mixture. Over time, but not instantaneously, the effects of consuming flour with high levels of chlorpyrifos from the first purchase would be mitigated.

Other examples abound, such as purchasing a carton of strawberries, all of which come from the same grower and field. The residues on these strawberries will be very similar (i.e., statistically

clustered). Consumption of the strawberries on a daily basis will continue until they are all gone. If the first strawberry has residues, then the remaining will likely have residues, and *vice versa*. Thus, dietary exposure associated with eating strawberries from a single carton of strawberries will likely be correlated. The Panel noted that Figures 44a and 44c in the DAS report depict uncorrelated doses drawn at random from a distribution of doses found in a population, while Figure 44b shows much more correlated doses. A dietary profile that goes from 44b to 44a, for example, would be a more realistic "carry over" profile. An individual may eat similar foods for consecutive days, although the tendency to do so, and thus the correlation in exposure, will decrease with time from the initial exposure. One panel member stated that "carry over" of a single day exposure to beyond 24-hours is the basic concept that distinguishes acute and repeated exposure toxicity evaluation. Therefore, it would be a good start to consider, for example, a Markov chain approach with admixtures of correlated exposures and random exposures.

The Panel recommended that the dietary exposure model (e.g., CARES, Lifeline) should consider highly correlated consumption behavior of specific foodstuffs. This more generalized exposure model would be relevant to other AChE-inhibiting chemicals and perhaps other chemicals with a different mode of action. One suggestion to improve estimation of longitudinal dietary exposure was to make use of the model output presented in Figure 46 of the DAS report, which showed dose dependent AChE inhibition with constant daily exposure. This simulation indicated that more than five days may be needed to achieve steady-state AChE inhibition. Steady-state AChE inhibition can be substantially higher than that resulting from a single day exposure; i.e., 2.5-fold at 0.0001 mg/kg for infants and adults, 12-fold at 0.1 mg/kg for infants and 20-fold at 0.1 mg/kg for adults. Depending on the exposure level, it could take more than 10 days to achieve steady state AChE inhibition. Thus, instead of modeling a 5-day dietary exposure for the longitudinal analysis with only one peak, an upper bound AChE inhibition to address the day-to-day linkage of high exposures can be estimated by extending the single day peak level from CARES output to the number of days that it takes to achieve a steady state.

The Panel commented that only dietary exposure is considered in the current form of the chlorpyrifos PBPK/PD model. Ultimately, the model should consider other exposure scenarios, including single and various frequencies of multiple days of exposure coupled with pertinent endpoints of action for each exposure duration, which were not presented in the DAS report. The DAS report focused on a single exposure scenario and a single set of toxicity endpoints. Instead, investigation should examine long-term exposure matched with long-term endpoints identified in the current toxicity database and these will be used in modeling for both animals and humans.

The Panel encouraged DAS to develop innovative methods for displaying the model results. For example, Figure 59 in the DAS report (DAS, 2011) was very difficult to follow. DAS plotted dose on the 5th day on the x-axis, but the y-axis plots predictions over the entire five-day dietary period. It is very difficult to interpret the figure without knowing the amount of "carry over" from the previous day. One panelist suggested computing a Markov Chain transition matrix (see description in Parzen, 1962; Bailey, 1964) between the previous and current day's dose. With a low dose on the previous day, chlorpyrifos concentration or inhibition on the current day will depend on the current dose. With high dose on the previous day, the transition matrix will show how the current day's dose increases the chlorpyrifos concentration (or inhibition). The source-

to-outcome model would then be run to populate the transition matrix, as well as computing a transition matrix for days 3, 4, 5, or steady-state.

One panelist had difficulty in responding to this charge question because the DAS report did not provide sufficient information to confirm or refute the conclusions made by DAS. For example, this panelist pointed out that very little detailed information is available to confirm the conclusions drawn from Figures 51 and 52 (DAS, 2011, pp. 119-120). This panelist did not understand the statement made in the Report that "...the top half of the population are not substantively affected by the uncertainty introduced by the absence of longitudinal data on dietary intakes.", and therefore questioned what preliminary model output can be used to confirm this statement.

QUESTION 2.2: *Please comment on the soundness of the conclusions reached by DAS with respect to the impact and importance of accounting for longitudinal dietary exposures to chlorpyrifos at the upper ends (e.g., >99th percentile) of the exposure distribution. To what extent, if any, can the DAS conclusions be generalized to other AChE-inhibiting chemicals and under what conditions? Please suggest how this issue of generalization to other pesticides could be investigated and explored?*

Panel Response

The Panel agreed that it was important to account for longitudinal dietary exposures to chlorpyrifos at the upper ends (e.g., >99th percentile) of the exposure distribution. However, the Panel stated that it would be premature to conclude that the chlorpyrifos analysis can be generalized to other chemicals or when multiple AChE-inhibiting chemicals are present in a single eating event. For example, the CARES-PBPK/PD model predictions for chlorpyrifos may be erroneous if there are other pesticides also present with different affinities for BChE or AChE.

The Panel indicated that the CARES-PBPK/PD model that DAS developed for chlorpyrifos could, conceptually, be applied to other AChE-inhibiting chemicals under the provision that those chemicals share some very similar physical, chemical, metabolic, and pharmacokinetic characteristic with chlorpyrifos. Certainly some of the pharmacodynamic variables, when carefully validated, can be expected to be useful for chemicals with the same mode of action. However, chlorpyrifos is a unique case study because there are extensive data available for deriving the model input variables, which will not be the case for many other AChE- or BChE-inhibiting chemicals. The Panel noted that different exposure periods may be needed for other AChE-inhibiting chemicals. For example, a five-day exposure period may be appropriate for chlorpyrifos, but not for malathion.

The Panel noted the importance of considering the real-life scenario in which dietary exposures often involve multiple OP pesticides, or other chemicals, during a single eating event. Under such circumstances, the outcomes from the CARES-PBPK/PD model for chlorpyrifos may need to be modified to account for the simultaneous presence of other pesticides with different affinities for binding BChE or AChE such as malathion. Malathion has a stronger binding affinity for BChE than chlorpyrifos, and therefore may lead to a greater inhibition of AChE by the co-presence of chlorpyrifos. The degree of AChE binding would also be different depending upon whether there are other cholinesterase-inhibiting chemicals present with chlorpyrifos.

Ultimately, a model to facilitate a full risk assessment required under the Food Quality Protection Act would have the capacity to address cumulative risks from aggregate exposures of organophosphate pesticides. The current CARES-PBPK/PD model, as discussed, does not assess the cumulative risks posed by multiple OP pesticides. It would have to consider the impact from both the pattern and level of exposure, which may warrant an alternative approach to the one used to assess exposures to a single chemical alone. Therefore, the Panel considered it important that the source-to-outcome model development encompass multiple OP pesticides.

Furthermore, pesticide risk assessment would also need to address the risk of non-dietary exposure to workers and their families and to children from hand-to-mouth exposures (indoors). A recent study (Beamer et al., 2009) has shown children, especially those younger than two years of age, receive a considerable non-dietary exposure from mouthing activities and from their close proximity to the floor. Their upper end exposures from aggregate and cumulative standpoints should be estimated separately. Similarly, as the DAS team made comparisons of CARES-PBPK/PD model output to several human biomonitoring data sets, it is important that these aspects of their longitudinal exposure (i.e., other routes of exposure, in addition to diet) also be considered.

Charge Question 3: Model Calibration and Evaluation with Direct Dosing Human Studies

QUESTION 3: *Please comment on the model evaluation approach comparing the linked CARES-PBPK/PD source-to-outcome model predictions with the Kisicki et al. 1999 study. In what ways could or should the model evaluation approach used by DAS be extended? Are there other model evaluation methods with respect to this aspect of the DAS manuscript that the Panel suggests be performed? To what extent does the Panel agree that the DAS model predictions are reasonably consistent with those of the Kisicki et al (1999) literature-reported values?*

Panel Response

The Panel agreed that the data from the Nolan et al. (1982) and Kisicki et al. (1999) studies were scientifically suitable to be used in the model evaluation. The Panel concluded the Nolan et al. (1982) study was suitable for calibrating the DAS PBPK/PD source-to-outcome model. However, the Panel noted that, for several reasons, the model was not "validated" using the Kisicki data. For one, the Panel commented that since some of the model parameters were calibrated to the Kisicki data, it puts into question the independence of a model-to-data validation.

The model evaluation that this question refers to is only the comparison of the LifeStage PBPK/PD model to the Kisicki et al. (1999) data. In light of the limitations of the Nolan and Kisicki studies, the Panel concluded that the model fit to the Kisicki data was generally reasonable (by visual comparison), although the model fit for the highest dose was not particularly good, and the model's ability to predict cholinesterase inhibition recovery was not as good as would be desired. The charge question asked the Panel to comment on the "reasonableness" of a model evaluation without providing a compelling description of how the joined models will be used. Without knowing how a model will be used, it is not possible to comment sufficiently on its reasonableness or appropriateness. The Panel recommended that the

potential sources of uncertainties identified in these and other comparisons should be captured for assessing the overall model uncertainties if the model is used for risk assessment. This analysis will be useful to ensure goodness of fit at a dose range outside the potential human exposure range and for possible model adjustment that can reduce the uncertainties in model predictions at the range relevant to human exposures.

Model Calibration

The EPA Human Studies Review Board (HSRB) is a Federal Advisory committee tasked with reviewing any studies conducted by third parties involving the intentional dosing of humans with pesticides that EPA might use in its regulatory decisions. The HSRB reviewed both the Nolan et al. (1982) and the Kisicki et al. (1999) studies in 2009, and had access to the original reports of data provided to the Agency (HSRB report of the June 2009 meeting provided to the Panel). The HSRB found that there were scientific limitations to both studies, but found the studies to have been conducted ethically and, therefore, there were no ethical barriers to EPA's use of the studies.

In the current source-to-outcome modeling effort, the Nolan study was used to calibrate the DAS PBPK/PD model. The HSRB viewed the Nolan study as largely useful and reliable; however, they expressed several concerns with the scientific quality of the data. The HSRB questioned the reliability of the chlorpyrifos data because levels in blood and urine were at the limit of detection. They also questioned the accuracy of the TCPy measurements in blood because of the technique used and its ability to detect the appropriate TCPy conjugate. The HSRB considered the cholinesterase measurements to be reliable, although they pointed to the lack of an untreated control group to assess inter-day variability, which was, not unexpectedly, quite high. The HSRB also raised several scientific concerns with respect to the Kisicki study. There were no detailed descriptions of the methods, e.g., whether there was urinary hydrolysis to release TCPy, how the data were analyzed, and why there was just 35% absorption. The HSRB's findings are detailed in the final report of the June 2009 HSRB meeting (see OPP Regulation Docket: EPA-HQ-OPP-2010-0588).

During the meeting, DAS responded that two of the major issues raised by the HSRB were addressed in a written response provided to the Office of Pesticide Programs. The response was not provided to the HSRB so there were no corrections to the HSRB report at the time of the SAP meeting. DAS provided a copy of their response to the SAP (see OPP Regulatory Docket: EPA-HQ-OPP-2010-0588), which addressed two of the major HSRB concerns: 1) there was hydrolysis of urinary conjugates prior to quantification of TCPy, and 2) disparities in proportion of dose absorbed between the Kisicki (35%) and Nolan (70%) studies could have been due to differences in chlorpyrifos particle size between the two studies. The DAS response presented experimental detail that was not present in the original reports to EPA. Therefore, only a few minor concerns continue to persist from the HSRB review regarding the reliability of the experimental data. Based on the additional information provided in the DAS response to the June 2009 HSRB report, the Panel agreed that the data from the Nolan et al. (1982) and Kisicki et al. (1999) studies were scientifically adequate to be used in the model evaluation.

Even though the Kisicki study was deemed scientifically useful, there were important limitations in the data. The Panel discussed some of the limitations. The red blood cell AChE inhibitions

were calculated compared to each person's baseline, but there was insufficient documentation on sample batching procedures. No total mass balance analysis was performed. There was some question about the dosage delivered to one individual. The red blood cell cholinesterase depression (25%) in one individual at the 1 mg/kg dosage suggested inhibition, but this occurred at a later than expected time point (96 hours) and all samples were low on that day, so this observation probably did not reflect inhibition. One participant in the high (2 mg/kg) dosage level experienced 28% inhibition at 12 hours, so this was a biologically plausible time and level of inhibition; unfortunately, this subject was lost to the study after this time point so did not provide information regarding the time course of the inhibition and potential recovery. Despite these limitations, data reported by Kisicki et al. (1999) were useful to compare with the predictions of the CARES-PBPK/PD source-to-outcome model.

Modeling Evaluation

The Panel commented on the model evaluation approach comparing the linked CARES-PBPK/PD source-to-outcome model predictions with the Kisicki et al. (1999) study. According to the model file, *LS_Human_Parameters.m Vd and Kel* (lines 133 and 134) were optimized to the Kisicki data, so the model was not independently validated" on the Kisicki data as some of the model parameters were optimized to the same data. While the predictions of the cholinesterase inhibition seemed fairly good, there were relatively few observations of cholinesterase inhibition. Some significant deviations occurred at higher dose levels (e.g., 2 mg/kg) and, to a lesser extent, for cholinesterase inhibition. The model predictions appeared to agree better with the TCPy data, although it was difficult to draw conclusions from the agreement at this stage of the model evaluation. The model description presented did not employ the exposure data from CARES, which might have provided better exposure estimates. These deviations could be considered "conservative" in the sense that they either match the data or predict higher levels of active compound and higher levels of cholinesterase inhibition than actually observed. One panel member commented on the judicious use of subjective terms like "conservative" since we do not know the full possible uses of the model. The Panel indicated that the deviations probably represented a systematic error of some sort, possibly the over-emphasized contributions of the B-esterases pointed out in the previous section of this response.

The Panel recommended that the potential sources of uncertainties identified in these and other comparisons should be captured for assessing the overall model uncertainties if the model is used for risk assessment. This analysis will be useful to ensure goodness of fit at a dose range outside the potential human exposure range and for possible model adjustment that can reduce the uncertainties in model predictions at the range relevant to human exposures.

This charge question asked the Panel to comment on the "reasonableness" of a model evaluation without providing a compelling description of how the joined models will be used. The report showed examples/illustrations of how a joined model may be used but it did not address how this specific model will be applied. Assessment of reasonableness and appropriateness of a model (or modeling platform) goes hand-in-hand with its intended use. As such, one cannot answer whether the reported model evaluation is appropriate. Nonetheless, the following comments and suggestions were made to illuminate general principles of model evaluation.

1) The term model evaluation should refer to a process that includes: 1) comparison of predictions from the entire modeling platform to data, when possible, 2) comparison of model components, e.g., the LifeStage PBPK/PD model, to data, and 3) diagnostic analysis of outputs from the entire model platform, and intermediate outputs from its subsidiary components. To their credit, DAS did show many examples of comparing model components.

2) The LifeStage PBPK/PD model evaluation was very sequential. A model was coded, it was tuned to data, and then the model was compared to other data. This is rarely how model development is done. Model development and reconciliation to data is usually very iterative. A conceptual model is developed; data are found; model parameters are proposed; predictions are compared to data; the model is refined; parameters are refined, and the process is repeated. The Panel encouraged the developers of case studies to bring to light the messiness of this process (i.e., bouncing between model development, calibration, analysis of fidelity, and prediction). An example would be the development and analysis shown in Dr. Hattis' presentation during the public comment period (see OPP Regulatory Docket: EPA-HQ-OPP-2010-0588.

Specifically, the Panel thought it would have been useful to have seen results for:

- Model parameters developed for Kisicki and predictions compared to Nolan;
- Model parameters developed for both Kisicki and Nolan;
- Model predictions as derived from so called default parameters and predictions compared to data;
- Model predictions derived for each individual;
- Effect of the parameter uncertainty and variability on outcome predictions;
- Alternative model formulations calibrated to data and compared to the LifeStage model; and
- Separate modeling of males and females to determine if there are sex differences in disposition and response.

The Panel listed a number of additional comments and questions.

- How was the calibration of the model performed using the Nolan study?
- What numerical method do you used to solve the mathematical equations? What is the convergence accuracy? The time steps for chlorpyrfos-oxon are fast compared to TCPy formation, which results in numerically stiff equations. What is the assurance that numerical errors are minimized?
- Raw data should be plotted, but when there are just a few samples, the standard deviation can be misleading.

- The discussion on the importance of activity was very interesting and well performed. How can this discussion be applicable to the general approach to source to dose analysis (DAS, 2011, p. 71)? In other words, how can the overall activity period be determined?
- The assumption of 100% absorption is conservative, but it does not yield a distribution.
- Some parts of the model may be able to be validated independently. The model needs an evaluation of assumptions and algorithms independently.
- Perhaps Figure 56 in the DAS report (DAS, 2011) could be expanded to a matrix to make the process clearer.
- There are questions about the estimated variability.
- There are many variables in the model adding to its complexity. The model was parameterized with Monte Carlo techniques to show variability across the population. Inter-individual variability in enzymatic activity had the biggest influence on the outputs. The Panel suggested that modeling efforts focus on the more sensitive parameters to increase the efficiency of these efforts.
- Although the peak AChE inhibition resulting from a specific chlorpyrifos dose is the current interest in the regulatory agency, the "carry-over" AChE inhibition from one exposure to the next would be as important as the peak inhibition under the circumstance in which a longitudinal exposure scenario is considered.
- All models of this complexity will have some compromises. The model needs to be validated to its ultimate purpose. The model will be valid only for a range of applications, and cannot be used outside that range of applications.
- If human exposure studies are feasible in the future, repeated dosing studies would yield useful information.
- The terms used in the source-to-outcome modeling should be explained in a manner that does not use jargon and proprietary terms.

Charge Question 4: Comparison of Model Predictions with Human Monitoring Data

QUESTION 4: *Please comment on the model evaluation approach used by DAS to compare the linked DAS source-to-outcome model dose predictions to (i) Eaton et al (2008), Whyatt et al. (2009) and Barr et al. (2010), (ii) the Curwin (2007) and Lu (2008) literature values and (iii) NHANES data measurements. To what extent does the Panel agree that the DAS model predictions are reasonably consistent with those of the literature-reported values and the NHANES data? Please suggest, if appropriate, other model evaluation methods or alternative approaches for comparing model predictions with actual human exposure that the Panel recommends to further evaluate the model predictions. Please include in your comments suggestions for what additional datasets should be used for comparison.*

Panel Response

The Panel indicated that DAS made a strong effort to compare the results of the source-to-outcome model and the validation model to data and past models reported in the public record. The DAS model provided similar results to other published models.

The Panel discussed the DAS model in the context of the overall model evaluation approach proposed by DAS. In an attempt to validate the modeling results, DAS proposed a direct comparison of the modeled results, scaled when appropriate, to observations made in field investigations as shown in Figure 1 (Figure 56 in the DAS report, DAS, 2011, p. 128). The first step in their process was to evaluate the results of three different modeling systems. Although not the point of this charge question, the fact that the results of the various modeling systems gave similar results was encouraging. The Panel noted that comparing the calibrated model to the data used for calibration was a simple, perhaps preliminary step in the validation process.

With regard to direct comparison with biological field sampling, the Panel responded that data available for comparison and validation studies were quite sparse. Only two relatively small studies of blood and serum levels from two adjacent states (New York and New Jersey) have been published (Whyatt et al., 2009 and Barr et al., 2010). [Note: Eaton et al. (2008) use the same data as Whyatt et al. (2009).] More data exist for urinary markers of chlorpyrifos exposure. The newest NHANES study, when released, will include data from about 1000 subjects and thus, may offer better validation data. Until that time, the data from these two studies are the best available. It is difficult to argue with soft terms like "reasonably consistent," especially when a large fraction of the data reported in the two investigations is below the reported detection limit. The Panel would have preferred a more quantitative statement, but this was not possible given the data at hand.

The Panel found the comparison between the DAS model and the NHANES data to be problematic. The source-to-outcome predictions did not agree particularly well with the data. The Panel also commented on DAS' omission of data at the limit of detection (LOD) in the comparisons. Much of the TCPy data in NHANES are below the LOD. Inclusion of the LOD values for NHANES would offer a clearer comparison. Results from two smaller studies, Curwin et al. (2007) and Lu et al. (2008), were used to compare the source-to-outcome model results with field data. While the Panel concluded that the method of comparison itself seemed valid, the conclusion that the predictions were consistent with observed values was overstated. As a result, the Panel found that the DAS comparison of the model to the published data of TCPy in urine was cursory. The urinary concentrations of TCPy from the individuals in the Curwin et al. (2007) study were not appropriate to compare with model predictions because exposures other than dietary were included.

The Panel thought the source-to-outcome model was inadequate to predict urinary TCPy values that could be compared to the published data. Considering that the PBPK-PD model had been calibrated thoroughly, one would expect the model outputs, in terms of chlorpyrifos, chlorpyrifos-oxon, and TCPy in blood and AChE inhibition, to resemble the data from human volunteer exposure studies (see response to Charge Question 3). Therefore, the inability of the CARES dietary exposure model to generate representative input data significantly reduces

confidence in the predictability of the PBPK-PD model. The Panel reported on recently published data from Lu et al. (2010) that could be used as potential input values in the PBPK-PD model to assess how the model would perform compared to the observed values.

Several Panel members expressed concern with using a simple multiplicative factor to estimate TCPy levels in pregnant women. The Panel consensus was that using a simple multiplicative factor provided inaccurate and imprecise TCPy levels in pregnant women.

Concern was expressed with regard to exposures to TCPy from the environment through the diet, and with non-dietary intake of chlorpyrifos, e.g., through inhalation. The Panel indicated that additional work in this area of the model is important. An additional concern focused on the comparison with occupational data. The route of exposure in occupationally-exposed individuals would likely be quite different from the dietary pathway emphasized in the remainder of the report. As a consequence, this analysis may be inappropriate as a demonstration of the validity of the model.

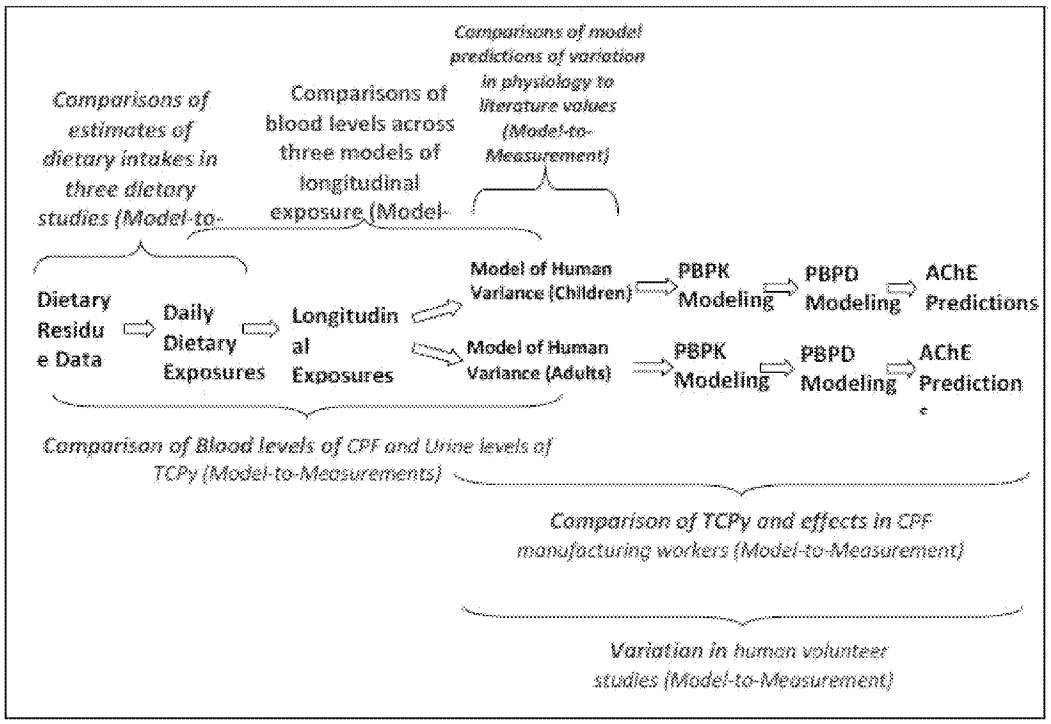

Figure 1. Steps in the source-to-outcome model and anlyses that provide a basis for the evaluation of the model outputs. This figure is a reproduction of Figure 56 in the DAS report (DAS, 2011, p. 128).

Discussion of the Blood Chlorpyrifos Measurements and Modeled Results

The Panel noted that the Whyatt, Eaton, and Barr studies were relatively small, with less than 100 maternal and cord bloods measurements in the former and less than 150 in the latter (Whyatt et al., 2009 and Barr et al., 2010). Table 1 (Table 7 in the DAS report, DAS, 2011, p.131) provides a summary of the measured chlorpyrifos blood levels in the Whyatt et al. (2009), Barr et al. (2010), and Eaton et al. (2008) studies. Inspection of the table indicated that the majority

of the measured values were less than the level of detection (LOD) for both maternal and cord blood samples. In the Whyatt study, at least 75% of the samples were below the LOD, while in the Barr study, at least 50% of the cord blood samples were below the LOD, while at least 95% of the maternal blood samples were below the LOD. In both the Whyatt and Barr studies, LOD values were variable and ranged from 0.5 to 1 ng/L (Dana Barr, personal communication). Barr's lab analyzed both sets of samples. Only summary data were listed in the Eaton presentation of the Whyatt data. In comparison, the modeled values were not limited by laboratory analytical limits of detection and values thus were estimated for all individuals. The DAS model tended to under predict the observed values by a factor of 2-3. DAS deemed that the model fit "reasonably well." The Panel concurred both with the model fit to the studies, given the intrinsic variability of results near the LOD, and with the general approach of direct comparison.

Table 1. Comparison of measured blood chlorpyrifos levels (ng/L) from selected biomonitoring studies with predictions from source-to-outcome model. This figure is a reproduction of Table 7 in DAS report (DAS, 2011, p.131).

	Source-to-outcome model[1]	Eaton *et al.* (2008)[2]	Whyatt *et al.* (2009)[2]		Barr et.al (2010)	
Exposure Date[3]	*2002-2010*	*2001-2002*	*2001-2004*		*2003-2004*	
Measurement	MB[4]	MB and CB[5]	CB	MB	CB	MB
Number of Individuals[6]	1000	76	65	92	148	138
Range						
Min	0.007	<LD[7]	<LD	<LD	<LD	<LD
Max	7	16	-[8]	-	1.8	10.1
Geometric Mean	0.2	0.5	-	-	1.5	4.2
Percentile						
10	0.05	-	<LD	<LD	<LD	<LD
25	0.09	-	<LD	<LD	<LD	<LD
50	0.2	-	<LD	<LD	<LD	<LD
75	0.4	-	<LD	<LD	1.3	<LD
90	0.7	-	2.3	1.5	1.7	<LD
95	1.0	-	2.5	2.5	1.8	<LD

[1] Model simulations were conducted using the adult data. Values based on randomly selected data points on day 5. Blood levels were adjusted up by 1.3 to reflect impact of pregnancy on CPF levels in blood as described by Lowe *et al.*, (2009).
[2] Summary data from the Whyatt studies were reported by both Eaton *et al.*, 2008 and Whyatt *et al.* 2009. Data for 2001 and 2002 are reflected in both sets of values.
[3] Date of that samples were taken for measured data or date of residue data used in dietary simulation
[4] Maternal Blood
[5] Cord Blood
[6] Number of simulated blood levels or number of measured values.
[7] Less than level of detection. Level of detection was reported by authors to range from 0.5 to 1.0 ng/L
[8] Indicates data not available

The Panel reasonably argued that an alternative approach to the comparison would be to exclude modeled values below the analytical limit of detection and compare the results only with those exceeding this value, but such is a small collection. Although no statistical test-of-fit was made, by and large, a visual comparison suggested that there was no difference in the data. One panel member pointed out that it is possible to use some statistical tests, such as the Shapiro-Francia Goodness of Fit test[2] (Royston, 1982a, b) to compare the predicted distribution of AChE concentration (in RBC or brain) to the left-censored data from animal or human studies.

One panel member pointed out that some of the biomonitoring data used for comparison may not be completely appropriate as some, if not all, of the cohorts in the surveys were exposed to chlorpyrifos through other routes in addition to diet. For example, chlorpyrifos was detected in the indoor and personal air among the inner city cohorts in the Whyatt et al (2009) study. This was echoed by other members of the panel and reflected in early work by MacIntosh et al., 2001.

Discussion of the Urinary Chlorpyrifos Metabolite Measurements and Modeled Results

The Panel agreed that the NHANES data offers the most comprehensive dataset of urinary TCPy measurements. DAS provided their comparison of the modeled results to urinary TCPy levels in Figure 57 of the report (DAS, 2011, p. 133) and reproduced below in Figure 2. The Panel found the comparison problematic and did not think the source-to-outcome model results fit very well with the observed data

Much of the TCPy data in the NHANES investigation is below the LOD. Inclusion of the LOD values for NHANES could offer a better comparison. Although the distribution would be flat, or artificial depending on the way the LOD values were treated, it could allow for a better estimate of the "lower tail." DAS assumed that only 5-20% of urinary TCP was from ingestion of chlorpyrifos (Figure 2). Therefore the model's predictions of TCP urinary concentrations from just chlorpyrifos exposure were adjusted upward to compare with measured levels.
The dotted line representing the source-to-outcome model does not fit very well with the "observed" data, which are also the results of a "model" albeit a simple one relating urinary concentration via total modeled urinary output to total TCPy output.

[2] The Shapiro-Francia goodness-of-fit test statistic is a simplification of the Shipiro-Wilks' test (Shapiro and Wilks, 1965) that works for larger sample sizes. Both tests exploit the fact that if a set of data come from a normal distribution then the relationship between the observed order statistics (y(i)) and the corresponding expected order statistics (z(i)) will follow a straight line of the form y(i) = m + s z(i).

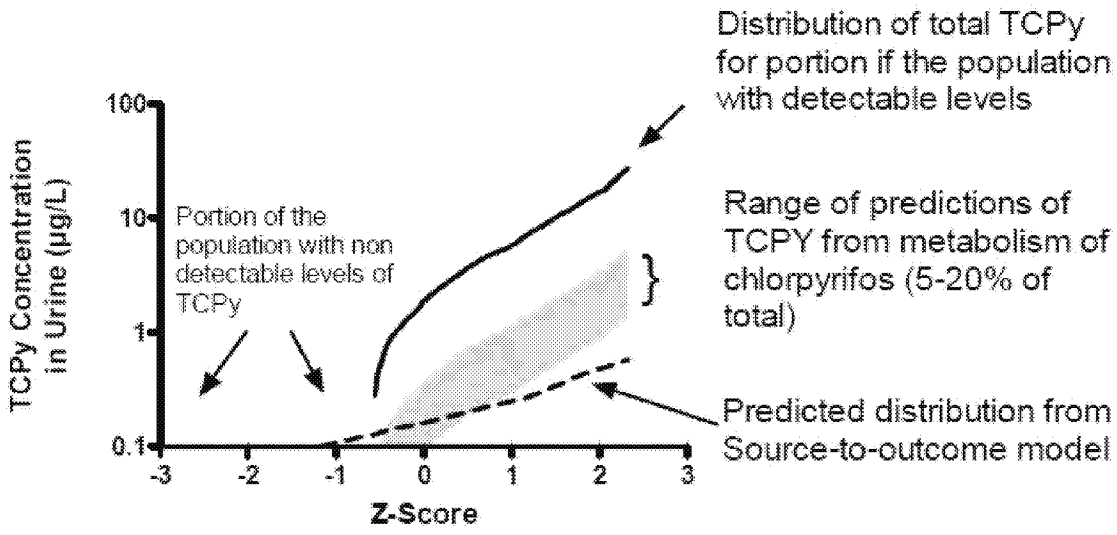

Figure 2. Cumulative distributions of predictions of current urinary levels of TCPy from the source-to-outcome model for adults, observed levels in the U.S. population in the 2002 NHANES survey (30% of population have levels below the NHANES level of quantification and are not shown), and estimates of the portions of NHANES levels that come from chlorpyrifos dietary exposures. This figure is a reproduction of Figure 57 in the DAS report (DAS, 2011, p. 133).

DAS compared the predicted levels of urinary TCPy from the source-to-outcome model to the observed urinary TCPy levels from two monitoring studies, Curwin et al. (2007) and Lu et al. (2008). These comparisons are shown in Table 2 below (Table 8 in the DAS report, DAS, 2011, p. 134). The modeled predictions (at the 90% CI) range from 0.08-0.3 µg/L (adults) and 0.3-1.0 µg/L (children). Both the Curwin et al. (children and adults) and Lu et al. (children) studies reported mean TCPy levels that were greater (an order of magnitude) than the median TCPy levels predicted by the source-to-outcome model. In children, observed mean TCPy levels (Lu et al., 2008; Curwin et al., 2007) were higher than the predicted TCPy levels by a factor of 10 and 30, respectively, and for adults, by a factor of 55 to 60 (Curwin et al., 2007).

DAS adjusted the predicted urinary TCPy levels upward due to the complicating presence of TCPy in the environment and as dietary residues on food. Several researchers have indicated that only 5-20% of TCPy in urine is from dietary intake of chlorpyrifos residues (Barr et al., 2005; Eaton et al., 2008; Wilson et al., 2003). Following the adjustment, the predicted levels were much closer to the observed levels in children (within a factor of 1.6 to 6.4) and adults (within a factor of 2.9 to 11) as shown in slide #43 of the presentation by Mr. Paul Price (DAS) ("Source-to-Outcome Modeling" presentation is available in the OPP Regulatory Docket: EPA-HQ-OPP-2010-0588). DAS concluded that the source to outcome predictions were consistent with 5-20% of observed levels, which they decided was consistent with the reported literature on environmental levels of TCPy. One panelist indicated the degree of environmental degradation of chlorpyrifos is approximately 30-35% (Lu et al., 2005) and disagreed with the use of 5-20% used by DAS. Overall, the Panel concluded that the method of comparison itself seemed valid,

but considered the conclusion that the modeled values were consistent with the observed values in the literature to be overstated.

Table 2. Comparison of selected measured urinary TCPy concentrations (µg/L) from selected biomonitoring studies with source-to-outcome model predictions. This figure is a reproduction of Table 8 from the DAS report (DAS, 2011, p. 134).

	Source-to-outcome model[1]	Curwin *et al.* (2007)	Lu *et al.* (2008)
	Median (90% CI)	Geo. Mean (±Std. Dev.)	Mean (Range)
Date[2]	2008-2009	2001	2003-2004
N	1000	Varies[3]	701
Men	-	12 (3.8-47)	-
Women	-	11 (1.8-35)	-
Women & Men	0.2 (0.08-0.3)	-	-
Children	0.5 (0.3-1.0)	16 (5.4-54)	5.1 (0-32)

[1] Source-to-outcome model simulations were conducted using the CARES model data. Day 5 values are presented.
[2] Date of sampling for measured data or date of exposure data simulated for source-to-outcome model
[3] 24 adult women, 24 adult men and 51 children
[4] Data not reported in published study

The Panel indicated the real limitation of the source-to-outcome model for simulating population-based TCPy urinary levels may not be embedded in the PBPK/PD model component, but rather in the dietary exposure predicted by the CARES model. That is, the CARES model used 5-day dietary exposure values (see discussion on longitudinal dietary exposure in Charge Question 2). Therefore, the source-to-outcome model's predicted range of urinary TCPy, 0.3 – 1.0 µg/L (90% confidence interval) would not be representative of the much broader range of chlorpyrifos exposure in the general population. It would also under-predict the mean urinary TCPy levels measured by Curwin et al. (2007) and Lu et al. (2008). As reported by NHANES and Lu et al. (2008), there were significant portions of the study participants in the respective studies with no measured (below the LOD) TCPy levels. In this regard, the source-to-outcome model over-estimated the baseline or low-end chlorpyrifos exposures.

The Panel stated that the urinary concentrations of TCPy from the individuals in the Curwin et al. (2007) study were not appropriate to compare with model predictions because exposures other than dietary were included. The study participants were farm workers and their families and therefore the urinary levels of TCPy represent aggregate exposures, not solely dietary exposures. The concept of aggregate, multimedia exposure to chlorpyrifos and cumulative exposure to chlorpyrifos and other AChE inhibitors was a recurring theme in the Panel's discussion. The Panel noted the current report was limited in this regard (aggregate exposure) and recommended that attention be given to include aggregate exposure in the modeling effort to provide a more complete understanding of the exposure-response matrix.

The Panel noted that the Curwin et al. (2007) and Lu et al. (2008) studies provided not only mean and median numbers, but the distributions of the dataset as well, see Table 2 below. For

example, the study published by Lu et al. (2008) also contained seasonal distributions of urinary TCPy levels. The Panel thought it would be a good exercise for DAS to compare the model simulation outcomes to those distributions of urinary TCPy levels under other circumstances, such as different seasons and spot vs. volume-weighted average measurements.

The Panel drew attention to a recent paper with content relevant to the topics being discussed by this SAP (see Lu et al., 2010). Lu et al. (2010) reported on the use of the Exposure-Related Dose Estimation Model (ERDEM) which used the data from Nolan et al. (1982) and Timchalk et al. (2002) studies, In this regard, the Lu et al. ERDEM model is similar to the source-to-outcome model. The ERDEM model estimated predicted TCPy levels and these values were compared to the observed data. The authors concluded that the ERDEM model grossly under-predicted the 24-hour cumulative TCPy levels in urine as well as the subsequent absorbed dose estimates for most of the 26 cases except for two cases in which chlorpyrifos was detected in the 24-hour duplicate food samples. The predicted urinary TCPy levels and the absorbed chlorpyrifos dose represented 75% of the observed values, which, when environmental degradation of chlorpyrifos is taken into account, becomes nearly coincident with the observed values. Although this study was small, it was designed specifically for the purpose of using a PBPK model for estimating and validating human biomarker data. This panel member recommended that both the Agency and DAS use the data published in the Lu et al. (2010) paper as potential input values in the PBPK-PD model to assess how the model would perform compared to the observed values.

In conclusion, the Panel thought the source-to-outcome model was inadequate (in its current form) to simulate urinary TCPy values comparable to the published data because the CARES dietary exposure model did not generate representative input data. Considering the PBPK-PD model had been calibrated thoroughly so the model outputs, in the forms of chlorpyrifos, chlorpyrifos-oxon, and TCPy in blood and AChE inhibition, would be close to the data from human volunteer exposure studies (see response to Charge Question 3), the inability of the CARES dietary exposure model to generate representative input data significantly reduces the confidence in the PBPK-PD model.

Discussion of Model Predictions of the Relationship between TCPy in Urine and BChE in Occupationally Exposed Adults.

DAS also decided to look at the relationship between urinary TCPy levels and BChE activity in a group of occupationally exposed workers. They made use of both the LifeStage model and the variability model to evaluate this work. The Panel found the relationship between this output and the main model to be unclear as the main model focused on dietary exposure while the principal exposure experienced by these workers was unlikely to be via the ingestion route. The model comparisons suggested a relative good fit of the data at low urinary TCPy and low inhibition, but both the main model and the variability model over-predicted the activity reduction at high doses. However, the value of this work was not apparent.

More General Comments

Observational biomonitoring data. The Panel noted that conducting any new human controlled-exposure studies is unlikely given the ethical considerations with such studies. Thus,

observational biomonitoring data would be the main source for identifying the effects of chemical exposure in the future. There are limitations in the model-to-biomonitoring comparisons and they should be purposefully noted to inform future biomonitoring data collections and give guidance as to what investigations are likely to prove fruitful. Further, farm workers and their families are unique in their exposure pattern to chlorpyrifos from field applications. Characterizing each exposure pattern should accompany the biological sample collections. This would allow more informed comparisons to model predictions

Model structure. DAS made predictions of maternal blood concentrations of chlorpyrifos, but did so with a 1.3-fold adjustment factor to predictions from their model for a reference male to account for changes in pregnancy and cited work from Lowe et al. (2009). The Lowe et al. (2009) manuscript did include a maternal and fetal component to the PBPK model and compared model predictions to maternal and cord blood concentrations of chlorpyrifos after delivery from Whyatt et al. (2005). However, the human model, which was based on rat data, was not able to predict observed human cord blood concentrations. The lack of an explicit description in the current DAS PBPK/PD model for maternal and fetal pharmacokinetics and pharmacodynamics to chlorpyrifos is a major deficit in the DAS approach. Exploration of mechanisms that may explain the cord blood to maternal blood concentration ratio in humans compared to rats is important for developing a PBPK/PD model that will be accurate for use in health risk assessments of chlorpyrifos.

The Panel considered the method of simply multiplying simulated chlorpyrifos blood concentrations by a factor of 1.3 to determine concentrations in pregnant women to be questionable. Ostensibly, this was performed to correct for changes in blood lipid composition during pregnancy and hemodilution. Changes in blood lipid have an impact on all partition coefficients. Partition coefficients have an impact on volume of distribution and hence clearance. Pregnant women also have a different physiology than non-pregnant women, e.g, tissue volumes, that may impact blood levels as well. A single multiplicative factor might be too simplistic.

The Panel noted that Dr. Dale Hattis (Clark University) presented information from an independent modeling of chlorpyrifos pharmacokinetics. Although the model parameters were the same as the DAS model, there appeared to be a disparity between the predicted blood concentrations from Dr. Hattis's work and those based on the work of Timchalk et al. (2002) upon which the DAS source-to-outcome model is based. The source for the discrepancy is unclear, but indicates the need to carefully review the DAS modeling work for accuracy before use in health risk assessments.

.

Model-to-data comparisons. The Panel noted that the while model predictions were "consistent" with the available data and that the explanations for differences were certainly plausible, there were data discrepancies that might be a result of some model misspecification errors. For example, in Figure 3 below (the model uncertainty bounds are tighter than the corresponding bounds in the data for RBC TCPy (the error bars should be 1.63*stdev). These differences may or may not be significant, but this was not clear.

The Panel suggested that DAS should be more even-handed in describing the model's performance. To illustrate this point, the Panel remarked on two examples in the DAS report in

which DAS seemed to understate the possible importance of differences between the predicted values and the observed values. DAS concluded that the predictions shown in Figure 57 of the DAS report (not shown) were "consistent" with NHANES. The Panel did not agree with this interpretation. In Figure 4 (Figure 58 in DAS, 2011, p. 136) differences were reported as "slight overestimates," but the Panel suggested that the differences could be structural.

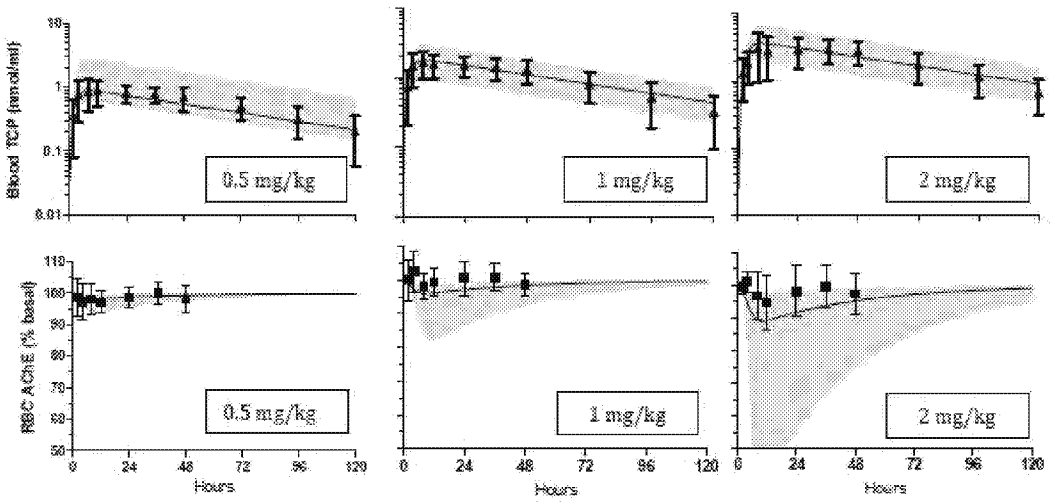

Figure 3. AChE inhibition and TCPy in blood in volunteers (Kisicki et al., 1999) over time following a single oral dose of chlorpyrifos. Triangles and error bars reflect actual data. The black line indicates the mean predictions and the gray area illustrates the 5th and 95th percentiles of the predicted range of response for individuals over time from the Variation model. This figure is a reproduction of Figure 32 in the DAS report (DAS, 2011, p. 80).

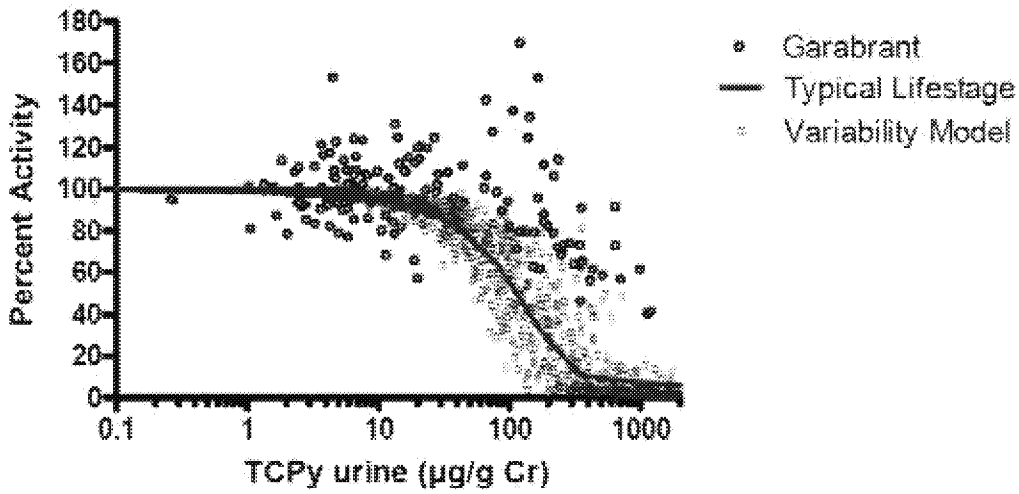

Figure 4. Observed relationships of cholinesterase inhibition and TCPy in urine for workers in a chlorpyrifos manufacturing plant and for control workers. Predicted relationship (assuming major route of exposure is oral) using theVariation PBPK/PD model. This is a reproduction of Figure 58 in the DAS report (DAS, 2011, p. 136).

On page 161, top paragraph (shown below), DAS concluded that the final predictions are within an order of magnitude and that more data are required to be more precise.

> Page 161, DAS (2011)
> *"Because of this, we conclude the final predictions of impacts of AChE in the general populations are within an order of magnitude but additional representative data are required in order to determine if the models are in fact more precise."*

To be more objective, the Panel suggested that DAS state the intended fidelity of the models, in light of available data and other unknowns factors and uncertainties in its report and then discuss whether their goals were achieved (or not) at the end.

The Panel suggested that DAS exercise the model to compute what data are likely to improve model fidelity and performance. In a value of information analysis, the analyst should be able to compute the likely improvement to the model if more data were available. This can be performed probabilistically based on the likelihood of values, and their variability. This would greatly help the Agency determine what data should be obtained for future studies. It also tells the Agency what is limiting in the model. In addition, this panelist noted that there was little discussion in the DAS report to indicate how the chlorpyrifos case study may be extended to a general source-to-output analysis for other chemicals. The Panel recommended that the Agency consider how the model be used for general source-to-outcome analysis.

Charge Question 5: Sensitivity Analyses, Variability, and Uncertainty

QUESTION 5.1: *The four step procedure described above was intended to permit DAS to focus on the factors that were most important in determining variation in response. Please comment*

on the methods used by DAS to assess variation in response (e.g., identification of sensitive factors and collection and integration of empirical data on variation). Please discuss the extent to which the methods described are appropriate and complete?

QUESTION 5.2: *During the Scientific Advisory Panel meeting held in July 2010 in which EPA's Office of Research and Development SHEDS/PBPK model was presented (see http://www.epa.gov/scipoly/sap/meetings/2010/072010meeting.html), the Panel reviewed ORD's Bayesian Approach to quantitative uncertainty analysis. DAS in its source-to-outcome model did not attempt to perform a formal quantitative uncertainty analysis (QUA), but instead evaluated components of the model by performing model-to-model comparisons (e.g., multiple dietary exposure models and multiple models of longitudinal exposures) and by performing model-to-measurement comparisons for internal dose (e.g., chlorpyrifos in blood and TCPy in urine) and AChEI in blood and plasma. Since formal Bayesian QUAs are only rarely conducted, are there other methods (short of rigorous Bayesian approaches) that the SAP can recommend for characterizing uncertainty due to limited data?*

Panel Response

Rationale for a Unified Answer to Charge Questions 5.1 and 5.2

The Panel addressed Charge Questions 5.1 and 5.2 in one, unified response. Indeed, the members of the panel felt that there was an artificial "separation" of sensitivity and variability (or "variation," to use the terminology of the DAS report) on one hand, and of uncertainty on the other, in the formulation of these two questions. The Panel stated that, instead, the issues of sensitivity, uncertainty, and variability should be addressed in an integrated and consistent manner, that would consider explicitly and take into account the relationships and complementarities of these attributes for the system under study. In fact, the Panel agreed that these considerations, and the corresponding analysis, should take place not only within the context of the biological "Variation Model," that is primarily the focus of these particular questions, but within the context of the full multi-model framework (involving a sequence of both dietary exposure modeling and biological kinetics and dynamics modeling) employed in the overall effort presented in the DAS report. The Panel recognized that not only the results of a sensitivity analysis can guide an efficient uncertainty and/or variability analysis, but that the latter can also provide useful sensitivity-relevant information. Finally, the Panel agreed that the phrasing of Charge Question 5.2 was somewhat restrictive, as it appeared to presume that a Bayesian Quantitative Uncertainty Analysis was not an appropriate option for this study.

The Panel concluded that the sensitivity, variability, and uncertainty analyses were incomplete as outlined in the DAS report, but emphasized the work represented a reasonable starting point. The Panel agreed that a systematic and explicit characterization of both variability and uncertainty, separately and for all individual components/steps of the exposure-to-biomarker modeling system, is needed and would substantially improve the analysis presented in the report. Implementing Bayesian analytical methods to enhance the study under review was also strongly recommended by the Panel, though there was a wide range of suggestions regarding the extent and the specifics of such an implementation. The Panel recognized that gaps in the data make any analysis difficult, including employing the aforementioned Bayesian techniques. The Panel agreed that, despite gaps in the data, it is nonetheless valuable to also analyze components of the

ED_005731_00001263-00051

models when possible. The Panel also agreed that local sensitivity analysis was inadequate and recommended exploring and applying methods for computationally-efficient global sensitivity analysis. In relation to not only characterizing, but also reducing the uncertainties in the study, the Panel recommended a larger systematic effort to simplify the model using known assumptions, processes, and interactions in the exposure and the biological (pharmacokinetic and pharmacokinetic) processes.

The concerns and recommendations of the Panel are discussed first, in relation to the general conceptual approach and the terminology employed in the DAS report, and second, in relation to various specific issues involving implementation aspects of this approach.

General Considerations and Issues of Modeling Framework and Terminology

In Charge Question 5, the Panel is asked to "comment on the methods used by DAS to assess variation in response (e.g., identification of sensitive factors and collection and integration of empirical data on variation)" and to "discuss the extent to which the methods described are appropriate and complete." The DAS approach involved the stepwise "expansion" of a human PBPK/PD model, considered as a component of a "source-to-outcome" model from a "Typical Adult Model," to a "LifeStage Model" and then to a "Variation Model" incorporating inter-individual physiological, behavioral and biochemical variability.

Analysis of the inter-individual variation employed the following four steps:

1) Local sensitivity analysis was used to identify the most influential parameters that drive variation in response predicted by the DAS LifeStage Model;

2) Distributions of the identified influential parameters were assembled;

3) Attempts were made to account for and address correlations between input parameters; and

4) Comparisons were made of model-predicted variability to data distributions assembled from human volunteer studies.

The Panel commented on the terminology used in Chapter 4 of the DAS report (DAS, 2011).

1) The modeling framework, though identified as a "source-to-outcome" model, is in reality, a "dietary residue-to-biomarker" model. The fact that important routes and pathways of exposure other than dietary ingestion are not considered can have important impacts on any uncertainty analysis to be performed.

2) Even with critical processes missing from the above mentioned framework, it is still a complex "pipeline" modeling system (i.e., a set of "models in series") that involves computations in sequence, where outputs from one model become inputs for the next one and so on. Therefore, sensitivities, uncertainties, and variability are "propagated" along the "pipeline." Consequently, there is a "convolution" process that can result in either reduction or magnification of the metrics characterizing these attributes (whether these metrics are point values or distributions) as they propagate from model to model.

ED_005731_00001263-00052

Therefore, a systematic analysis must consider not only each component of the modeling system separately, but also the interlinked components together in the model sequence. In particular, "feedbacks" within the sequence of processes from exposure to outcome can be significant and should be incorporated in a comprehensive analysis of the system's dynamics. For example, not only can dietary intake affect physiology and biochemistry (metabolic etc. processes), but physiology and biochemistry can affect dietary intake. It is important that such links within the system (and within the corresponding models describing the system components) are explicitly considered and addressed. Uncertainty analysis for multi-model formulations (often called "integrated models"), however, present various challenges when compared to the analysis for single models (Babendreier and Castleton, 2005; van Asselt and Rotmans, 2002; Oughton et al., 2008).

3) The use of vague and ambiguous terminology in the DAS report raises various concerns. One example is the need to distinguish "variation" from "variability." Calling the expanded version of the "LifeStage Model," the "Variation Model," is rather unconventional. One could only speculate why the authors of the report did not endorse terminology that is standard in the field (such as, e.g., "Population PBPK Model"). Furthermore, usually "variation" is a generic term used to account collectively for all types of changes in a variable or a parameter, whether deterministic or random. The Panel discussed the need for clarifications on how variability and uncertainty are specifically addressed and what aspects of variability are treated as part of an uncertainty analysis. As pointed out in many studies, e.g., Isukapalli et al (2010), there should be a clear distinction between uncertainty and variability, as the treatment of these attributes could have significant policy implications. Insight into specific variabilities and their impact on risk estimates can help in targeting decision-making towards a population-based goal. In contrast, insight into uncertainties and their impacts can help prioritize resources for reducing uncertainty. This distinction has been widely recognized in both human and ecological health risk assessments (Barton et al., 2007; IPCS, 2006; Lester et al., 2007; USNRC, 2007).

   It is generally convenient to consider variability as natural ("irreducible") stochastic heterogeneity and uncertainty as an epistemic concept that reflects various (quantifiable and potentially reducible) limitations in our knowledge regarding the states and processes of the system studied. However, in practice, the "boundary" between variability and uncertainty is actually a moving one, and the location of this boundary reflects a compromise related to the effort that can be allocated towards minimizing uncertainty by collecting and analyzing information that explicitly quantifies variability.

4) The Panel stressed that there are different types of uncertainty that should ideally be considered and addressed separately; specifically, *model uncertainty* (associated with lack of knowledge that affects the model formulation itself) and *parametric/data uncertainty* (associated with inadequate and/or noisy data for estimating parameters of the model).

   a) *Model Uncertainty.* Model uncertainty can arise due to the choices made regarding the logical structure and detail of the model, the computational approach used to solve the model equations, and the model resolution. This type of uncertainty is

significant when alternative sets of assumptions and approximations used for developing a model can result in significantly different conclusions. For example, different types of PBPK models can be developed depending on the choice of compartment lumping, assumptions regarding mixing within a compartment (e.g., flow limited or diffusion limited), and assumptions regarding transport of the chemical of concern within a compartment. If variability has already been incorporated in the model via probability distributions of the variables and parameters, model uncertainty may include inappropriate selection of such distributions and omission of correlations among inputs/parameters (Nilsen and Aven, 2003; Sander et al., 2006). Uncertainties also arise when models or model components that are developed using assumptions and data appropriate for one set of variables are applied to other, substantially different scenarios (e.g., *in vitro* to *in vivo* extrapolation, interspecies extrapolation, high dose to low dose extrapolation). Another type of model uncertainty, *completeness uncertainty*, arises from factors that have not been identified (unknown processes affecting the output of the model) (Reinert and Apostolakis, 2006). Model uncertainties are reduced by refining the model through improved understanding of the system being studied. Their impact on output metrics is quantified through various available methods that include the "adjustment factor approach," the application of the best available model, or by assigning degrees of confidence in each of model and weighting the simulation model outputs (Zio and Apostolakis, 1996; Moschandreas and Karuchit, 2002; Nilsen and Aven, 2003; Linkov and Burmistrov, 2003; Sohn et al., 2004; Reinert and Apostolakis, 2006; Barton et al., 2007; Redding et al., 2008).

b) *Parametric/Data Uncertainty.* Parametric/data uncertainty reflects incomplete knowledge regarding the values of model parameters or input data (USNRC, 2007). This may be due to: random errors corresponding to limitations in the precision and accuracy of measurements, systematic biases, assumptions used to infer an actual quantity of interest from observed readings of a "surrogate" or "proxy" variable, and non-representativeness of data supporting the parameters. Parametric uncertainties can also arise from uncertainties in measurement data used for model evaluation and can also result when model parameters are estimated from available data by deterministic or stochastic parameter estimation techniques (Moles et al., 2003). In principle, parametric/data uncertainty can be reduced by gathering more data. The impact of data and parameter uncertainties on output metrics is typically addressed by "propagating" these uncertainties through the model and identifying significant contributors to overall uncertainty (e.g. Isukapalli and Georgopoulos, 2001).

5) The uncertainty, as well as sensitivity and variability analyses, can take place (and traditionally have indeed taken place) at substantially different levels of detail, depending on the objectives of the study and the resources (including laboratory and field data) available for the study. Traditionally, four tiers of analysis are distinguished (see, e.g., Isukapalli et al., 2010). When uncertainties are not explicitly characterized, a safety factor is used for characterizing upper bounds of the metrics of exposures or outcomes and risks (Stedeford et al., 2007), i.e., a Tier 0 Analysis. A Tier 1 Analysis involves conducting uncertainty analysis in a local manner: a point estimate of exposure, dose, or risk etc., combined with local sensitivity analysis. A Tier 2 Analysis involves

probabilistic characterization of uncertainties (with natural variability often lumped in the uncertainty distribution) and global sensitivity analysis. A Tier 3 Analysis involves separate and explicit characterization of uncertainties and variabilities by techniques such as two-stage (also called two-dimensional) Monte Carlo methods. Improvements in the performance of the uncertainty propagation can be achieved by using efficient sampling techniques, such as the Latin Hypercube Sampling (Helton and Davis, 2003; Rubinstein, 1981). More recent advances in efficient sampling techniques include Sequence Generators based on the methods of Sobol (Bratley and Fox, 1988), Halton (1960), and Faure (Kocis and Whiten, 1997). These efficient sampling methods can cover the space of the input distributions in a more "representative" fashion with a smaller number of samples. Efficient sampling is critical when a model is complex or when a relatively limited number of model simulations can be performed due to constraints in time and computational resources.

The Panel agreed that the Agency should consider a Bayesian approach for quantitative uncertainty analysis of the exposure-to-biomarker model. Panel members differed in their perception of the effectiveness of Bayesian methods for this type of problem. One panelist suggested considering the uncertainty on point estimates of each of the important parameters as a multivariate distribution from which to draw samples for simulations. This would be a poor substitute for a full Bayesian analysis, which might be cumbersome given the limited availability of data. Nevertheless, the Panel recommended that the Agency study the discussion contained in the 2005 SAP report on use of the Bayesian approach to uncertainty analysis in models (http://www.epa.gov/scipoly/sap/meetings/2010/072010meeting.html). Bayesian approaches can be used to support "calibration" of the model, ensuring that variability and uncertainty are incorporated into the model at the design stage. The use of Bayesian here involves more the overlaying of distributions for variability and uncertainty on an essentially "mean" (or "typical" human) model. An example of how Bayesian calibration has been used by the Agency can be seen in the draft Toxicological Review of Trichloroethylene for the Integrated Risk Assessment System (IRIS) (go to http://cfpub.epa.gov/ncea/cfm/recordisplay.cfm?deid=215006, see Chapter 3. Toxicokinetics, Section 3.5.5. Bayesian Estimation of Physiologically Based Pharmacokinetic (PBPK) Model Parameters, and Their Uncertainty and Variability)

Many recommendations and suggestions were made by the Panel to clarify and enhance the scientific and technical approach outlined in the DAS report (DAS, 2011). In some cases, the report was not explicit or sufficiently clear on the details of specific steps or calculations, the results of which are presented in a very concise manner. In some other cases, it was obvious that only a portion of the results obtained by the study are presented explicitly in the report. An example is visible in Figure 32 (on page 80 of the DAS report, DAS, 2011) where there is a large amount of information "encapsulated" in the results/comparisons; however, there is a substantially larger amount of information that should have been included in the report to allow a more comprehensive interpretation of each panel in that figure. Perhaps, this information could have been provided in the Appendices. The comments that follow include across-the-board recommendations that in some instances involve issues that may have been partially addressed in the report.

ED_005731_00001263-00055

Recommendations and Suggestions by the Panel Members for Clarification and Improvement of the DAS Report with a Focus on Sensitivity, Variability, and Uncertainty Analysis

1) *Consider Explicitly Exposure Modeling and Two-Way Coupling of Models in the Analysis of Sensitivity and Variability/Uncertainty*

The Panel stressed the need to consider explicitly the limitations of the (dietary) exposure models used in the study. The Panel further recommended the systematic and explicit characterization of the variability and the uncertainty in these exposure models (e.g., the CARES model), the propagation of this variability and uncertainty (represented by distributions of the outputs of the exposure model) through the biological (pharmacokinetic and pharmacodynamic) models, and the consideration of two-way coupling (including model "feedbacks") in this variability/uncertainty analysis. For example, an enhanced PBPK/PD model would include a lactation pathway, where the output of the PBPK/PD model for the mother would provide information for the dietary intake of the newborn.

2) *Perform Systematic Model Evaluations with Biomarker and Exposure Data for Uncertainty Analysis and Reduction*

The Panel strongly recommended the consideration of available exposure and biomarker data on chlorpyrifos and its metabolites from studies such as NHANES and NHEXAS for model evaluation and uncertainty reduction analyses (e.g., Xue et al., 2010).

3) *Expand and Refine the Modeling Framework of the Study in order to Better Understand and Reduce Uncertainties*

The Panel pointed out that real-world uncertainties in the system under study and in the models used to describe this system, are associated with many simplifying assumptions (and omissions of relevant processes) in the formulation and the application of both the exposure and the biological (pharmacokinetic and pharmacodynamic) modeling components. To understand and reduce these uncertainties in the modeling, other routes of exposure in addition to dietary ingestion and simultaneous exposures to other chemicals that share common biological pathways with chlorpyrifos could be considered. An example would be carbaryl and other cholinesterase inhibitors. It would then be possible and informative to model the pharmacokinetics and pharmacodynamics of chlorpyrifos in the presence of these chemicals. Characterization of uncertainties in the pharmacodynamics of the system due to the presence of other toxicants and food components is a potentially critical challenge in the overall uncertainty characterization effort. Biological models also would be enhanced by adding the gestation phase in the model and consideration of both the pregnant subject and the fetus. Other LifeStage PBPK/PD model formulations, which also incorporate the physiology of gestation are available, e.g., Sasso et al., (2010), are directly adaptable to chemicals relevant to the present study. Such enhancements and associated analyses of the modeling components will eventually improve the credibility and context for the modeling framework of the study.

4) *Consider Alternative Descriptions and Metrics of Uncertainty*

The Panel recommended the consideration of other techniques, in addition to classical probability distributions, for describing uncertainties in the modeling components of the

framework employed in the DAS study. Whereas classical probabilistic analysis is the most widely used method for characterizing uncertainty in physicochemical and biological systems, especially through probability distributions of uncertain parameters, various alternative ways of representing uncertainty are available in the literature and could in fact be particularly useful in the context of a complex system with multiple components, linking exposures to biological outcomes (Oberg and Bergback, 2005; Lester et al., 2007; Ionescu-Bujor and Cacuci, 2004; Cacuci and Ionescu-Bujor, 2004; Dorne et al., 2005; Isukapalli and Georgopoulos, 2001). For example, interval analysis is an appropriate choice when only rough estimates of "bounds" on uncertain parameters are available (i.e., "unknown, but bounded"), which happens when studies report only a range for parameters without specifying a distribution (Hickey et al., 2001; Granvilliers et al., 2004). This approach involves estimating the bounds in model outputs of interest based on corresponding bounds in the inputs and model parameters. Several computational toolboxes exist for interval analysis (e.g., Jaulin, 2001). The main limitation of this technique is that all the uncertainties are forced into one arithmetic interval and no further insight (e.g., into high percentiles) can be obtained.

Fuzzy logic theory facilitates uncertainty analysis of systems where uncertainty arises due to vagueness or "fuzziness" in system attributes rather than due to randomness alone (e.g., when qualitative information is used in modeling). This extends conventional logic to handle "degrees of truth" and maps specific variables from a discrete set to a continuous form. DAS should consider fuzzy logic methods where classical set theory is too restrictive. For example, when values of parameters are explained as "high" or "low" rather than with actual values, there is no easy translation of such terms into classical set theory; classical set theory requires unique membership in well-defined bins or value ranges. A useful reference to fuzzy logic methods is McKone and Deshpande (2005). Fuzzy logic theory might be particularly useful because its applicability to PBPK modeling has already been demonstrated.

Classical set theory has a "crisp" definition as to whether an element is a member of a set or not. On the other hand, fuzzy theory allows for a gradual degree of membership. This approach was introduced by Zadeh (Zadeh et al., 1996; Zadeh, 1965) and in recent years has seen renewed interest in PBPK modeling (Nestorov, 2001; Seng et al., 2007; Gueorguieva et al., 2004), especially as a means of translating between fuzzy logic and probabilistic analysis. Other approaches for uncertainty representation include fuzzy measure theory (Wang and Klir, 1992), fuzzy rough set theory (Klir and Yuan, 1995), possibility theory (Wolkenhauer, 1998), and belief measures (Gordon and Shortliffe, 1984; Halpern, 2003; Parsons, 2001; Buchanan and Shortliffe, 1984).

5) Consider Uncertainty Reduction through Bayesian Model Calibration

The objective of a systematic sensitivity, variability, and uncertainty analysis is not only to qualitatively and quantitatively characterize these attributes of the system studied and of the corresponding integrated model, but also to use these characterizations for the purpose of uncertainty reduction through incorporation of new data and refinements of the component models of the integrated framework. This can be seen as a systematic "model calibration." Panel members recommended that it could be pursued via Bayesian techniques such as Markov Chain Monte Carlo methods (Gamerman and Lopes, 2006; Gelman, 2004; Balakrishnan et al., 2003).

ED_005731_00001263-00057

6) *Consider Correlations of Input Variables in the Sensitivity/Uncertainty Analysis*

A majority of sensitivity/uncertainty analysis methods can only be applied directly when the inputs are independent. When input variables are correlated, e.g., the organ/tissue volumes in a PBPK model, these variables can be represented as functions of independent random variables with appropriate transformation techniques that have been described and demonstrated in the literature (Devroye, 1986; Isukapalli and Georgopoulos, 2001).

7) *Clearly Distinguish Variability from Uncertainty, Improve its Characterization, and Perform Systematic Combined Two-Stage Variability/Sensitivity Analyses*

Separate characterization of uncertainty and variability is important for population exposure-to-outcome modeling. The Panel agreed that uncertainty and variability should be characterized explicitly and separately through the application of methods such as "two-stage" or "nested" Monte Carlo (Cohen et al., 1996; Simon, 1999; Frey and Rhodes, 1996). In this approach, each uncertainty simulation includes a set of Monte Carlo runs that address variability. In the context of exposure modeling, this is typically accomplished as follows (see Xue et al., 2006). For each uncertainty simulation, the parameters for distributions that describe variability of parameters within the population are sampled from their corresponding distributions. Subsequently, for each variability run, samples are generated from the parameter distributions in the uncertainty simulation. The uncertainty and variability simulations are tracked separately so that they can be characterized separately. For each uncertainty simulation, a cumulative distribution function (cdf) can be developed to characterize the variability. Aggregation over the different uncertainty runs provides a "range of confidence" in the distributions describing the variability within population. The two-stage Monte Carlo analysis allows identification of subgroups or conditions that could be of particular concern by focusing on a specific region of the distribution such as the 95th percentile. Likewise, it is possible to investigate the combination of factors that produce such a high output and approaches to mitigate those factors.

The Panel suggested that the change of partition coefficients with area should be explicitly incorporated into the model, e.g., the different adipose tissue composition between children and adult (see Price et al. 2003). Another suggestion was to explicitly account for the effect of factors such as age, gender, weight and height (as described in Haddad et al., 2006 and Verner et al., 2008) on the variability of organ/tissue physiology. The physiological variability in the DAS model is determined by body weight only, except for fat tissue which varies further with the variability of body height.

One panelist pointed out that using such a method would improve the model's ability to capture inter-individual variability in organ volume/weight as observed in brain weight >25 kg (DAS, 2011, Figure 14, p. 51). Though the brain weight was given as an example during the discussion, it was subsequently suggested that liver weight (DAS, 2011, Figure 15, p.52)) is another appropriate example. The simulation model does not seem to include the additional variability around the line in that figure (the line being a "deterministic" mean fit of the body weight versus tissue volume data). The "scatter" around the line represents variability in tissue size among individuals of the same bodyweight. This same panelist also pointed out that the description of metabolism requires additional microsomal data, including data specifically for the early days of life for both males and females (reported separately).

8) Consider Methods for Expanding from Local to Global Sensitivity Analysis

The Panel agreed that the local sensitivity analysis employed in the DAS approach has substantial limitations and suggested the consideration/exploration of global sensitivity analysis methods. Limitations of the local methods are summarized below, along with suggestions for available global methods to consider from the literature. The Panel suggested an alternative option that combines globally sensitivity and uncertainty analysis with the use of surrogate models.

As mentioned earlier, a gradient-based, local sensitivity analysis was used to estimate how the LifeStage PBPK/PD model behaves within a narrow range of parameter values and which parameters contribute most to sensitivity in the vicinity of a baseline (see pp. 67-68 in DAS, 2011). This approach obtains a "snapshot" of how the model behaves around the baseline and can describe general model behavior only if the model is linear. When normalized sensitivity of a model output with respect to an input is large, care must be exercised in making decisions based on that output. In some cases, this can even shed light on the validity of some underlying assumptions. The Panel suggested that it would be very helpful to consider the sensitivity of other model outputs, such as AUC, time to reach peak or 90% of peak or 90% of final evacuation to model parameters. The Panel also remarked that the local analysis was reporting on the Jacobian, but not on the Hessian sensitivity matrix. Another Panel member performed additional local sensitivity analysis for a partition coefficient using the ACSLX model that was provided to the Panel; specifically, a 20% variation in brain: blood partition coefficient to study the corresponding impact on brain AChE inhibition. The 20% variation resulted in a16% change in the brain chlorpyrifos-oxon concentration and 12% change in AChE inhibition, which suggested that brain AChE inhibition is sensitive to the values and variations of the partition coefficients. The Panel recommended that parameter sensitivity should be tested for other endpoints important to effects, e.g., RBC and brain AChE inhibition, as reported here; chlorpyrifos brain levels for potential non cholinergic effects, and biomonitoring for chlorpyrifos in blood and TCPy in plasma and urine.

In addition to the "direct" method that was used to calculate local sensitivities in the DAS report, there are various other techniques that offer various computational advantages by automatically calculating local sensitivities at various baseline points. These techniques include Automatic Differentiation (Isukapalli et al., 2000; Bischof et al., 1996), Symbolic Differentiation (Sandu et al., 2003), and the Decoupled Direct Method (Dunker, 1984). However, when the model is nonlinear, these methods may not be able to capture the impact of an input on an output of interest, since different sensitivity patterns occur in different regions of "input/parameter space" or "input/parameter domain." Thus, local sensitivity analysis must be performed for a large set of points in the parameter domain including different combinations of parameter inputs. In performing a sensitivity analysis, there is a continuum of possible sets of runs to consider. For example, assume there are N key parameters as the focus of the sensitivity analysis. At one extreme, one could perform runs with each parameter varying between an upper and a lower base value while all other parameters are held at their nominal "expected" value. This would require $2*N$ runs of the model. This approach is conceptually close to the analysis performed by DAS. At the other extreme, one could run the full factorial analysis, a Latin Hypercube (LH) design, which would require 2^N runs. With N greater than 5, the number of runs using the LH design becomes prohibitively large. Many panel members focused on the use of a Bayesian approach to

perform the sensitivity analysis, Markov Chain Monte Carlo sampling, because with a few thousand runs more information could be obtained than the full 2^N runs of the LH deterministic approach.

For nonlinear models, local methods can only be used to rank the parameters in order of importance for the single set of input values on which the analysis was performed. In contrast, a global first-order index considers the effect of varying a single parameter both individually and in conjunction with all combinations of parameters for all possible parameter sets. An analysis that does not consider interactions between input parameters implicitly assumes that the effect of varying two parameters simultaneously is given by the sum of their individual sensitivities. This assumption holds for linear models, but is not valid for PBPK models, which incorporate nonlinear expressions, e.g., to represent rates of metabolism. A local analysis is only applicable within a small "distance" from the point in the "parameter space" at which the analysis was performed. A global (domain-wide) sensitivity analysis (GSA) is needed to study the system behavior over the entire range of parameter variation, often taking the uncertainty in the parameter estimates into account. However, the results from GSA are difficult to interpret when the distributions of underlying input/parameters are changed significantly.

The Panel noted that both the sensitivity and uncertainty analyses need to be considered together when studying simulation models. A parameter may have large uncertainty, but the output may not be sensitive to the parameter. Likewise, a parameter may have small uncertainty but the output may be highly sensitive to the parameter. The collective impact of uncertainty and sensitivity depends on many factors. A wide range of GSA techniques and applications is available in the literature (Saltelli, 2008; Saltelli et al., 2000). These techniques include regression of Monte Carlo or model simulation runs based on efficient sampling (Cullen and Frey, 1999), High Dimensional Model Representation (HDMR) techniques (Li et al., 2001; Rabitz and Alis, 1999), Fourier Amplitude Sensitivity Test (FAST) (McRae et al., 1982), Extended FAST method (Saltelli et al., 1999), the method of Sobol (Sobol, 2001), Bayesian Analysis (Jonsson et al., 2007), and Probability Bounds Analysis (Ferson and Tucker, 2006; Nong and Krishnan, 2007). Techniques used for PBPK models need to be numerically efficient, able to deal with non-monotonic output, and differential forms. Most of the sensitivity analysis methods listed above satisfy these criteria and allow for variance decomposition and estimation of main and interaction effects (Saltelli and Bolado, 1998). Generally, PBPK models have well in excess of 20 input parameters, which is a challenge for global methods. A possible alternative, whose accuracy should be tested, is to use a global screening method that is computationally less expensive, such as the Morris' Test (Morris, 1991), to determine the most influential parameters and then to run a global quantitative method on only those parameters. Morris' Test provides a "one-at-a-time" design for screening model inputs that do not have significant impact on model outputs. For a model with n inputs, this method involves statistical analysis of a set of r trajectories, each containing $n + 1$ points. Efficient design of these trajectories for better representation of the input space has also been proposed (Campolongo et al., 2007). Efficient factorial design based techniques can be used in conjunction with Morris' Test (Pujol and Corre, 2007).

9) Consider Combined Global Sensitivity/Uncertainty Analysis with a Surrogate Model Approach

The Panel also suggested considering another general approach for performing combined global sensitivity and uncertainty analyses called the "surrogate" or "fast equivalent" model approach. This approach has been demonstrated with applications involving both exposure and PBPK models. This approach is based on a representation of the original model in terms of an approximation that captures the essential elements of its behavior while aiming for the maximum possible simplicity. This can be an approximation of either the deterministic model or of the probabilistic model that incorporates probability distributions of inputs and parameters. Examples of methods following this general approach, which have been demonstrated in exposure and PBPK modeling applications include: the Stochastic Response Surface Method (SRSM) (Isukapalli et al., 2000, 1998) and the High Dimensional Model Representation (HDMR) (Li et al., 2002; Wang et al., 2005). In the case of the SRSM, the uncertain model inputs are represented using a basis of standard normal distributions, since there is a direct, one-to-one transformation between any two probability distributions (Devroye, 1986; Isukapalli and Georgopoulos, 2001). The outputs are then approximated in terms of orthogonal series expansions of standard random variables with coefficients that can be calculated from a limited number of model runs. The SRSM provides a set of sample points (combinations of different inputs) for running the model. Once the model is run at these points, the model outputs can be used to estimate the coefficients of the orthogonal series expansions. This is followed by the production of statistically equivalent polynomial approximation of the model outputs, which can be used to estimate statistical properties of the outputs as well as correlations among outputs, and among outputs and inputs. The coefficients of the series expansions in the SRSM provide sensitivity information of model outputs with respect to the inputs.

In case of the HDMR method, there are several variants possible. The Cut-HDMR method (Li et al., 2001) aims to reduce a complex model in terms of look-up tables based on a limited number of model simulations. The reduced models are fast models that can be used to perform Monte Carlo type analyses. The Random-Sampling HDMR method (Wang et al., 2003; Li et al., 2002), on the other hand, relies on a set of random samples of model outputs in order to decompose the variance in terms of individual contributions. Both the SRSM and the HDMR method use orthogonal series expansions that allow for relatively straightforward global sensitivity analyses. Orthogonol expansions can also be used as "reduced" or "surrogate" models approximating the original model in computationally intensive applications (e.g., for Bayesian Markov Chain Monte Carlo analysis) (Balakrishnan et al., 2003).

10) Use the Outcomes of Sensitivity and Uncertainty Analysis to Parameterize and Simplify the Original Models in a Rationa l Manner

The Panel remarked that the sensitivity and uncertainty analysis can provide valuable guidance in estimating parameters as well as in simplifying the models used in the study under review. One Panel member argued that the results of the local sensitivity analysis presented in the DAS report (DAS, 2011, p.67-68) suggested that the LifeStage model is insensitive to many parameters, which could mean that the model may be through the removal of non-influential processes. For example, by identifying the most sensitive model parameters/inputs, resources can be focused to reduce uncertainty where it is most appropriate. Furthermore, the

ED_005731_00001263-00061

identification of non-influential parameters would help in calibrating other model parameters (without expending resources on estimating the non-influential parameters). Lack of sensitivity of one output with respect to an input implies that the data gathered for that output should not be used for estimating the parameters for the input. Therefore, sensitivity analysis should ideally be performed before estimating model parameters from available data. This is especially important in the case of Bayesian parameter estimation techniques. The performance of the estimation method can be improved significantly by using the most appropriate combination of data for estimating a given set of input parameters or by iteratively estimating appropriate parameters based on available data and sensitivity information. Furthermore, based on the results of a global sensitivity analysis, a simulation model can be simplified by removing the non-influential parameters (and corresponding detailed descriptions of non-influential processes). This is especially important when the removal of some processes makes the model easier to specify, to calibrate, or to run faster. This is especially true when the number of such inputs is large and there is significant uncertainty in their estimates (Jonsson et al., 2007). This caution applies to the processes of parameter estimation and of model reduction.

Charge Question 6: Calculating Data-Derived Extrapolation Factors

QUESTION 6.1: *Please comment on the strengths and limitations of DAS's proposed approach to estimate an animal-to-human extrapolation factor as described in Section 10. The Agency is concerned about a component of the proposed approach involving use of the study design characteristics of a single animal study (See Section 10). The Agency generally recommends a weight of the evidence approach for determining toxicological points of departure when multiple studies are available. For chlorpyrifos, there is a large database of study results performed across different animal life stages and so it may be more appropriate to determine points of departure and extrapolation factors based on an integrated analysis of these multiple studies instead of based on a single study. In your response, please comment on the approached proposed and provide guidance and suggestions for alternative approaches, if appropriate.*

Panel Response:

The Panel agreed with the general principle of the weight of evidence approach as commonly applied in many areas of risk assessment. The advantage of the weight of evidence approach is to provide more robust support for a range of uncertainty factors (UFs). However, the Panel expects that when a well-evaluated PBPK/PD animal-to-human model is ready for use in risk assessment, the use of UF_A would no longer be needed.

The UF_A proposed by DAS is based on the single study (Marty and Andrus, 2010) with PND11 and adult rats. The Panel noted that while the data from the Marty and Andrus study may be more suitable for conducting benchmark dose analysis than many other studies, there were also limitations in the study's possible application to other toxicity endpoints and human exposure scenarios.

1) The peak inhibitions increased over time in a tissue- and dose-dependent manner based on the modeled pattern for RBC AChE and the steady state was not achieved until many days after repeated dosing.

2) The DAS UF_A only pertains to the dietary exposure estimated from a model that did not include the prenatal period or infants with lactational exposures to chlorpyrifos.

3) The current dietary model needs to fully capture beverage and water consumption.

The Panel was critical of the DAS approach to estimate human variability with eight-person groups and proposed an alternative approach without regrouping. The Panel indicated that the $BMDL_{10}$ can be best estimated by applying the true definition of this concept directly to the results of the combined CARES PBPK/PD model. Once a human $BMDL_{10}$ estimate is predicted via the variability and uncertainty component of the model, it can be compared to the animal $BMDL_{10}$ estimate. The UF_A can then be derived by comparing the animal and human $BMDL_{10}$, both of which are derived from the same model, but with species-specific parameters. This approach uses the model as it was intended, as a tool to integrate the variability and uncertainty in the process of incorporating *in vitro* and *in vivo* animal data to prediction of risks in the human population.

The Panel stressed that there should be no limitations in accommodating the various sets of available chlorpyrifos data to derive a range of UF_A for selecting dose metrics that are most pertinent as toxicity indicators for endpoints and scenarios. A UF_A selected for risk assessment should be carefully matched to the human exposure scenarios. An obvious additional advantage of applying the model to many available data sets is that it can facilitate the identification of model limitations, both for future model improvement as well as for beginning to address model uncertainties throughout the process of model development.

Perhaps the merit of this charge question is that the standard point of departure (POD)/UF approach will be employed before adoption of the PBPK/PD approach advocated by DAS. However, it is unclear why the PBPK/PD model can be valid for estimating the UFs, but not for directly estimating the human equivalence dose. During the discussion, the Agency expressed the desire to estimate a set of UF_A or UF_H values for the purpose of transparency in risk communication at the transition stage going from the default UF paradigm to the PBPK/PD model application.

The Panel encouraged the development of a well-evaluated and robust PBPK/PD model. Such a model has the potential to advance the risk assessment process closer toward the direction envisioned in the NRC's "Toxicity Testing in the 21st Century" (NRC, 2007). As illustrated here, a PBPK/PD model can be used to directly estimate a range of human exposure doses at a pre-determined dose metric for the point of departure. In this case study it was 10% RBC AChE inhibition. As long as the dose metric is determined, having a full set of animal testing data for establishing a point of departure would no longer be needed. Taking AChE inhibition as an example, the focus of the current pesticide data requirements on *in vivo* toxicity testing and *in vitro* studies can be shifted more toward delineating the biological and toxicological significance of tissue- or site- specific levels of AChE inhibition for modeling acceptable human exposure dose levels, characterizing the model variables and parameters that significantly impact the modeling outcome, and understanding the mode of action that is crucial for the selection of the most appropriate model dose metrics for the endpoints of concern.

Marty and Andrus Study

The Panel recognized several areas of strength in the animal study by Marty and Andrus (2010) used for determining points of departure in adult and PND 11 rats. These included:

1) a relatively wide range (0.05-5 mg/kg) of chlorpyrifos doses, varying from no effect doses to toxic doses eliciting >50% AChE inhibition,

2) sensitive measures of adverse effects relevant to the key mechanism of action,

3) adequate numbers of animals per dose,

4) a quite immature age group (11-day-old),

5) potentially susceptible male and female pups for comparison with adults,

6) single and repetitive dosage regimens, and

7) an adequate statistical analyses.

Data on blood chlorpyrifos, chlorpyrifos-oxon and TCPy concentrations were included in Summary Tables 1 – 2 (p. 43-48), Table 3 (p. 68), and Table 4 (p. 71) in the DAS report (DAS, 2011). The Panel agreed that the data derived from this study were suitable for calculating a benchmark dose. Other studies have evaluated dose-response relationships for chlorpyrifos in immature and adult animals, but the availability of the raw data was limited, and some of these other studies used a limited number of animals. As provided in the DAS presentations, there does appear to be some correspondence in dose-related AChE inhibition among studies including the Marty and Andrus data, and studies by Moser et al. (1998) and Timchalk et al. (2006) with relatively similar age groups (PND 12-17). Relatively similar dose-related inhibition was also noted in another study with somewhat younger animals, seven days of age (Zheng et al., 2000).

The Panel noted that while the data from the Marty and Andrus study were more suitable for conducting benchmark dose analysis, there were also limitations. During the discussion, a DAS scientist stated that these measurements were made at or around the time of maximum AChE inhibition. Thus, the post-dosing blood sampling times were apparently varied (no peak AChE inhibition), which might make the data less suitable for inter-age or dose comparisons. However, the acutely dosed 11-day-old pups consistently exhibited significantly higher blood CPF and TCP levels than adults, but no or relatively minor differences in AChE activities.

Another limitation in this set of data is the relatively few points to develop the curve of inhibition and only two dose levels (2 and 5 mg/kg) eliciting significant group differences in degrees of inhibition, although the other studies, Moser et al. (1998) and Timchalk et al. (2006), which reported relatively similar dose-related inhibition, also had limited dose levels in which to

establish this relationship. If substantial, usable data are available with similar endpoints across different life-stages, meta-analysis of these combined data may provide a more reasonable basis for a weight-of-evidence approach for establishing points of departure.

DAS' proposed UF_A is based on the single study (Marty and Andrus, 2010) with PND11 and adult rats. The Panel stated that it was important to keep in mind the limitations of the DAS proposed UF_A when considering its possible application to other toxicity endpoints and human exposure scenarios. The Panel described three realistic scenarios which the DAS proposed UF_A does not address.

1) The peak inhibitions increased over time in a tissue- and dose-dependent manner based on the modeled pattern for RBC AChE inhibition in adults and children presented in Figure 46 of the DAS report (DAS, 2011, p. 113). Steady state inhibition was not achieved until many days of repeated dosing. This pattern from repeated exposure would be a better basis for deriving a single UF_A for RBC AChE inhibition than the outcome of a single exposure scenario.

2) The DAS UF_A only pertains to the dietary exposure estimated from a model that did not include the prenatal period or infants with lactational exposures to chlorpyrifos. The Panel indicated that a complete dietary exposure model should include fetal exposures to chlorpyrifos through the mother's diet, and post-natal lactational exposures due to the mother's diet.

3) It was unclear whether the current dietary model captures beverage consumption, including water consumption.

Alternative to Estimating the Human Variability without Further Regrouping

The Panel was critical of 8-person groups used to estimate human variability. UF_A was defined as the ratio of animal-to-human $BMDL_{10}$ (the lower bound of BMD at 10% RBC AChE inhibition). The animal $BMDL_{10}$ was derived from a BMD curve-fitting approach, whereas the human $BMDL_{10}$ was derived from PBPK/PD model simulation. The model simulating distribution of individual human AChE inhibition was further grouped into 8-person groups to match the sample size in the animal data from the Marty and Andrus (2010) study. The DAS report noted that this regrouping narrowed the range of human AChE inhibition as compared with the range of the original distribution of individual values. The Panel questioned the necessity of associating human modeling outcome with the sample size of the animal study and suggested that the simulation results of an individual can be used to estimate the human variability without further regrouping. A more detailed implementation of this alternative is provided below.

The $BMDL_{10}$ can be best estimated by applying the true definition of this concept directly to the results of the combined CARES PBPK/PD model. The $BMDL_{10}$ represents the lower 95th percentile of the benchmark dose for the health outcome. In this case, the dose in which there is 10% RBC AChE inhibition is the key biological parameter for establishing the point of departure for the risk assessment. The Panel used Figure 65 in the DAS report (DAS, 2011, p. 153) as the

starting point to compute a $BMDL_{10}$ from the dose response output from the linked CARES PBPK/PD chlorpyrifos model shown in Figure 5. The scatter of points ("blue cloud") is assumed to represent the relationship between the dose of chlorpyrifos experienced by each simulated individual and the predicted percentage of RBC AChE inhibition. In one sense, the model result does describe the population and expected outcome. Direct application of the $BMDL_{10}$ definition to these data involves first describing the mean relationship between dose and response. The black line in Figure 5 represents the minimized sums of the squared error in response to each dose. The line presented is for illustration purposes only and is not the true least squares fitted line. If the points represent individual-to-individual variability in response at given dose, then the vertical spread of points at any specific dose, e.g., 0.1 mg/kg, would produce a distribution of expected responses in the population at that dose. If instead a target percent RBC AChE inhibition level is chosen, say 10%, then the horizontal spread of doses at the 10% inhibition line would represent what R.A. Fisher described as the "fiducial distribution" of doses. Using parametric or non-parametric distribution fitting tools, one could estimate the dose at which p% ($0<p<100$) or less of the population would be expected to experience a 10% inhibition of AChE (Figure 5). This is one estimate of the BMD_{10}.

To get the $BMDL_{10}$, the model uncertainty has to be incorporated into the assessment. Conceptually, one has to estimate the uncertainty in the mean response relationship to get the "confidence interval" for the BMD_{10}. Using a Monte Carlo (2-D probabilistic risk assessment) approach, the uncertainty in model parameters (involving both fixed parameter uncertainty and the uncertainty in the hyper parameters for random parameter distributions) would be used to generate alternate "realizations" of the dose response relationship (an example of a second such realization is shown as the dashed line in Figure 5). Each realization results in a slightly different mean dose response line which also impacts the fiducial distribution at the 10% inhibition line. At the end of this exercise one has a "distribution of dose response relationship curves". The exercise of computing the BMD_{10} described above can be performed for each realization, resulting in a distribution of BMD_{10} values. From this distribution of BMD_{10} values, one computes the lower *p*-percentile, in this case the lower 10 percentile and this is the estimate of the $BMDL_{10}$.

This approach directly uses the model as a representation of what is truly happening in the population. If the model has biases, for example, those resulting from the lack of data from nursing infants or lactating or pregnant women, then the $BMDL_{10}$ estimates would also be biased. The more realistic and believable the linked CARES PBPK/PD model, the more believable the $BMDL_{10}$ estimate.

Once a $BMDL_{10}$ estimate is provided via the variability and uncertainty model, this can be compared to the animal $BMDL_{10}$ estimate computed in the traditional way. The linked CARES-PBPK/PD model integrates all of the age and other factors that typically confuse the comparison between animals and humans. The extrapolation factor, computed as the ratio of the human to animal $BMDL_{10}$ estimates, truly describes the required cross-species multiplier. This approach uses the model as it was intended, as a tool to integrate the variability and uncertainty in the process of incorporating *in vitro* and *in vivo* animal data to prediction of the risks in the human population.

ED_005731_00001263-00066

The Panel noted that this alternative approach to estimating the human variability without further regrouping highlights the flaws in the approach taken by DAS. The DAS approach starts with the animal data rather than with the model. It attempts to duplicate in the simulated human population what was done in the animal population in an attempt to get a BMD_{10} and eventually a $BMDL_{10}$ estimate. But grouping humans into samples of eight, as was done with the rats, is unnecessary, confusing, and misses the whole point of developing the human model.

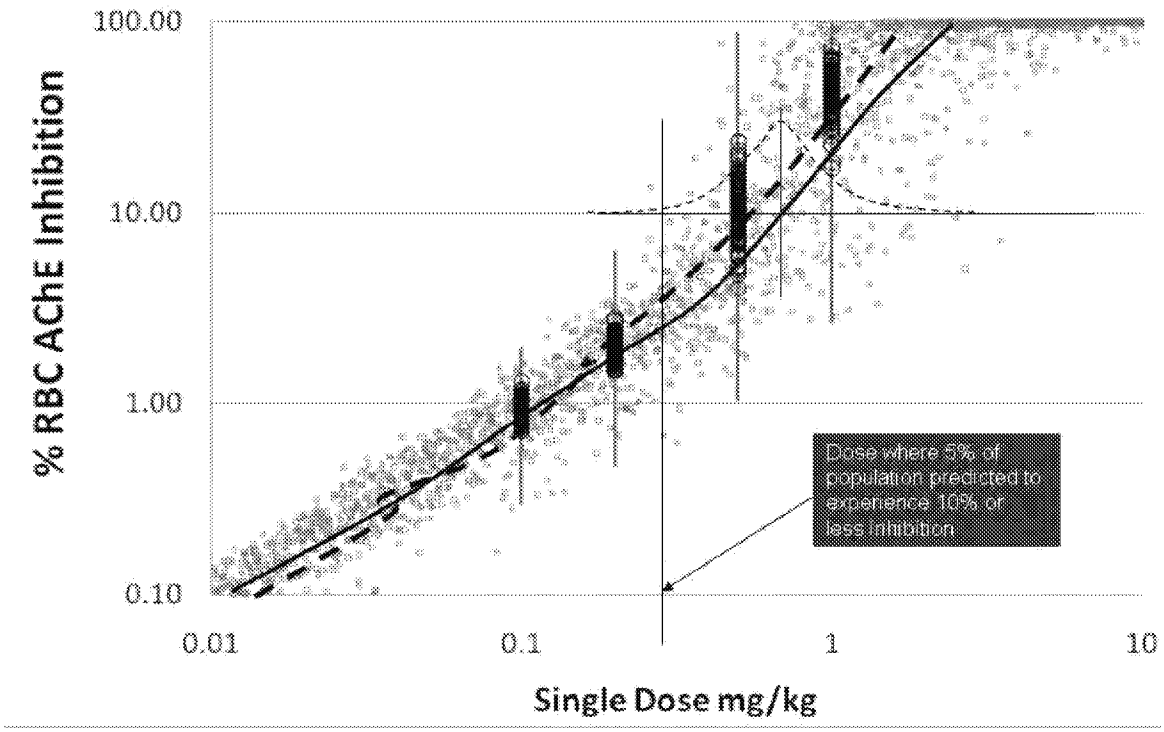

Figure 5. Determination of the BMD10 using the linked CARES PBPK/PD model. The blue cloud represents the outline of chlorpyrifos dose response in populations of infants exposed to a range of doses. The ranges of predicted responses for 5,000 simulated infants at four dose rates (0.1, 0.2, 0.5, and 1.0 mg/kg) are given by the vertical light blue lines. The black line illustrates the least squares fitted line (not a true least squares fitted line). The dashed line represents alternate "realizations" of the dose response relationship using a Monte Carlo approach in model parameters (involving both fixed parameter uncertainty and the uncertainty in the hyper parameters for random parameter distributions). The black circles show the range of mean responses in groups of eight infants (corresponding to the groups of eight test animals used in Marty and Andrus, 2010). The vertical lines with heavy dark circles are not relevant to this discussion. The dose where 5% of the population is predicted to experience 10% or less inhibition is approximately 0.4 mg/kg. This figure is adapted from Figure 65 in the DAS report (DAS, 2011, p. 153).

Alternative Approach for Deriving the Animal $BMDL_{10}$

Along with the above approach for deriving human variability, the Panel suggested an additional approach for deriving the animal $BMDL_{10}$. The Panel noted that the term "$BMDL_{10}$" used by DAS may carry un-matched meanings between animals and humans. On the one hand, the

human $BMDL_{10}$ was derived from the DAS PBPK/PD model that incorporated biological variability, i.e., pharmacokinetic and pharmacodynamic variability. On the other hand, the animal $BMDL_{10}$ was derived from the curve-fitting bounds of the eight data points per group via a mathematical model in which the pharmacokinetic and pharmacodynamic variation among these test animals was not explicitly modeled. Since the DAS PBPK/PD model structure and development originated with fitting model predictions to animal study data, the model should easily be applied to the animal data from this single dosing study. The UF_A can then be derived by comparing the animal and human $BMDL_{10}$, both of which are derived from the same model, but with species-specific parameters.

QUESTION 6.2: *Please comment on the strengths and limitations of proposed approach to estimate a human variability extrapolation factor as described in Section 10. Please provide alternative approaches, if appropriate.*

Panel Response

The Panel indicated that derivation of the appropriate extrapolation factor depends on using the linked CARES-PBPK/PD model to compute the critical value, e.g., UF_A, as noted in Charge Question 6.1. In Charge Question 6.2, the critical value, UF_H, is estimated not from the uncertainty of the mean response relationship model, but from the individual-to-individual variability in the response as shown in the scatter of points in Figure 65 of the DAS report (DAS, 2011, p. 153). More than 1000 simulated individuals would be needed to ensure accurate estimates of extreme events, for example, the 1% or 0.1% tail percentiles (the values at which 99% or 99.9% of individuals would be unlikely to present the target response, e.g., a 10% RBC AChE inhibition). The Panel commented that the model represents an impressive effort to consider all variables, which allows estimation and integration of variation among humans in their sensitivity to AChE inhibition following exposures to an anti-cholinesterase chemical; however, there were a number of comments on both model and data uncertainties. Panelists expressed a common conceptual concern with regard to modeling human variability based on the uncertainty raised concerning the use of *in vitro* metabolism data from a single study (Marty and Andrus, 2010) used in evaluating the range of responses in humans. Another consideration was that differential activation and inactivation of the ultimate toxicant is an important determinant in response to an organophosphate pesticide. As suggested before, the input values for human-relevant activation and inactivation rates used in this model were developed from non-physiological conditions with substrate, cofactor, and ionic concentrations often outside those which would be feasible in a biological system. Knowledge of site-specific chemical levels, e.g., concentrations at the active sites of biotransformation enzymes within the endoplasmic reticulum vs. concentrations within the organ as a whole, is limited. This lack of knowledge is a significant cause of uncertainty in predicting the concentrations of the chemical actually available for metabolism; therefore, extrapolating from those *in vitro* conditions to estimating real-world biotransformation rates in the intact human tissues would have considerable uncertainty.

Moreover, the use of small numbers of human tissue samples to estimate human-specific metabolic rates, particularly in the very young individuals where the greatest differences from older age groups may exist, raises the uncertainty in the constants used as inputs into the model. Although small numbers of samples were analyzed, Pope and coworkers (Pope et al., 2005)

reported that *in vitro* carboxylesterase activity in human liver at two months of age appeared both lower in total amount and higher in sensitivity to chlorpyrifos-oxon than carboxylesterase activity in tissues from older individuals (3 months to 36 years of age). The Panel recognized that limited data from younger individuals is an important deficiency that should be addressed. Little is known about the potential exposures to pharmaceutical or environmental chemicals that may have influenced the metabolic capacity in the individuals from whom the samples were collected or the disease state of an individual that could have potentially influenced physiological factors and metabolic rates, etc. The Panel remarked that difficulties with data from the direct human exposure studies available (see Panel response to Charge Question 3) along with the unlikely possibility of more human direct dosing data, make it difficult to further evaluate the impact on functional endpoint changes. Thus, estimation of human variability and consideration of an UF_H under these conditions remains a difficult task with high uncertainty. Again, the Panel emphasized the importance of identifying factors, which may significantly contribute to model uncertainties during the model development so that they can be adequately accounted for when the model is ready for use in risk assessment.

The Panel recommended that the current case study analysis be expanded to accurately predict all pertinent human exposures for target populations of interest (e.g., early life stages), groups with specific exposure characteristics (e.g., farm workers and their families), and multiple routes of exposure (e.g., inhalation, dermal, ingestion from food, water, hand-to-mouth activities), These are all important factors for fully and adequately addressing the inter-individual variability and uncertainty.

In comparison, DAS employed only two immature age groups (six months and three years) in their modeling exercise. Specific to the younger individuals, it would be preferable to include fetuses, neonates and infants younger than six months. Studies in the literature have demonstrated that the younger the individual, the more immature he/she is physiologically and biochemically. As a general rule, the most immature individuals are the most likely to differ from adults in susceptibility to injury by many drugs and other chemicals (NRC, 1993). Many age-related processes contribute to increased deposition of chlorpyrifos in the immature brain: increased gastrointestinal absorption due to immaturity of the gastrointestinal mucosa, diminished plasma protein binding due to low albumin levels, increased concentrations of bilirubin and fatty acids, relatively low body fat stores, reduced metabolic detoxication potential (e.g., PON activity), low hepatic blood flow, and an immature blood:brain barrier. The Panel recommended that the current case study analysis should be expanded to more accurately reflect different age- or lifestage-specific endpoints, which include developmental toxicity and various exposure scenarios of concern.

Finally, the Panel noted that the possibility of non-cholinergic neurodevelopmental disorders should also be considered in the future as a basis for pharmacodynamic differences in infants and children.

QUESTION 6.3: *The current effort by DAS is limited to food exposure and does not include all relevant exposure routes. Please comment on strengths and weaknesses of using data-derived extrapolation factors described in the current effort for life stages (e.g., gestation or pregnancy) and/or for routes (dermal, inhalation) not considered in the current modeling effort.*

ED_005731_00001263-00069

Panel Response

The Panel agreed that the DAS model for chlorpyrifos would eventually need to consider all routes of exposure. Moreover, the model should be extended to cover other life stages and conditions such as gestation and pregnancy and lactation and nursing infants.

The current DAS case study is limited to dietary exposure, excluding nursing infants. The calculated UF_A and UF_H pertain only to the dose response relationship from dietary exposure with the exposure level in mg/kg. It does not address oral exposure through hand-to-mouth activity, which is of particular importance for small children, nor does it address dermal or inhalation pathways. Obviously, the UFs derived for one route cannot be directly applied to other routes of exposure (i.e., inhalation, dermal). The important kinetic differences between oral dosing and other routes of exposure, first-pass metabolism, can have a dramatic influence on the extent and timing of the functional changes under consideration following OP exposure. More data are needed to better characterize all potential sources of chlorpyrifos exposure. One panelist thought it might be worthwhile to conduct a comprehensive monitoring survey, to learn whether chlorpyrifos persists in the homes of the general population (as compared to occupational workers).

Revised Model Name

One panelist stressed the point that the model name/descriptor should clearly convey what it is and thereby make clear what it is not. The current model is really a PBPK/PD model linked to a low-dose dietary exposure model of chlorpyrifos inhibition of cholinesterase. It is not a model of high dietary exposure, as might occur through a food contamination incident (e.g., adulteration of milk with melamine), nor is it a model of non-cholinergic effects of chlorpyrifos.

Grammatical/Typographical Errors in the DAS White Paper

1) Page 27, 2nd paragraph, line 4. B-esterases are not just cholinesterases but have to also include carboxylesterases and other sensitive proteins.
2) Page 78, line 2. Should be "…population was…"
3) Page 78, line 3-4. Should read "…were 0.15 for…"
4) Figure 33. Children are listed in title but adults in legend.
5) Figure 34. The reverse is shown, i.e., adults in title but children in legend.
6) Page 110, last line. Delete "were" at end.
7) Figure 60. Typo in y-axis label.
8) Page 148, 1st paragraph, line 3. Delete "…that a…"
9) Page 148, last line. Should be "The animal data reflect…"
10) Page 156, line 4. Should be "calculate".

REFERENCES

Armstrong N, Hibbert DB. 2009. An introduction to Bayesian methods for analyzing chemistry data: Part 1: An introduction to Bayesian theory and methods. Chemometrics & Intelligent Laboratory Systems 97(2):194-210.

Bailey NTJ. 1964. Chapter 5, Markov chains in the elements of stochastic processes with applications in the natural sciences. John Wiley & Sons, Inc, NY. pp 38-57.

Balakrishnan S, Roy A, Ierapetritou MG, Flach GP, Georgopoulos PG. 2003. Uncertainty reduction and characterization for complex environmental fate and transport models: An empirical Bayesian framework incorporating the stochastic response surface method. Water Resources Research 39(12):-.

Barr DB, Ananth CV, Yan X, Lashley S, Smulian JC, Ledoux TA, Hore P, Robson MG. 2010. Pesticide concentrations in maternal and umbilical cord sera and their relation to birth outcomes in a population of pregnant women and newborns in New Jersey. Science of the Total Environment 408:790-795.

Barton HA, Chiu WA, Setzer RW, Andersen ME, Bailer AJ, Bois FY, et al. 2007. Characterizing uncertainty and variability in physiologically based pharmacokinetic models: State of the science and needs for research and implementation. Toxicological Sciences 99(2):395-402.

Beamer PI, Canales, RA, Bradman A, Leckie JO. 2009. Farmworker children's residential non-dietary exposure estimates from micro-level activity time series. Environmental International 35(8): 1202–1209.

Bischof C, Khademi P, Mauer A, Carle A. 1996. ADIFOR 2.0: Automatic differentiation of Fortran 77 programs. Computational Science & Engineering 3(3):18-32.

Bratley P, Fox BL. 1988. ALGORITHM 659: Implementing Sobol's quasirandom sequence generator. ACM Transactions on Mathematical Software (TOMS) 14(1):88-100.

Brock, A. and Brock, V. 1990. Plasma cholinesterase activity in a healthy population group with no occupational exposure to known cholinesterase inhibitors: relative influence of some factors related to normal inter- and intra-individual variations. Scandinavian Journal of Clinical & Laboratory Investigation. 50:401-408.

Buchanan BG, Shortliffe EH. 1984. Rule-based expert systems: the MYCIN experiments of the Stanford Heuristic Programming Project. Reading, Mass.: Addison-Wesley.

Butte NF, Hopkinson JM, Wong WW, Smith EO; Ellis KJ. 2000. Body composition during the first 2 years of life: an updated reference. Pediatric Research 47: 578–585.

ED_005731_00001263-00071

Cacuci DG, Ionescu-Bujor M. 2004. A comparative review of sensitivity and uncertainty analysis of large-scale systems - II: Statistical methods. Nuclear Science and Engineering 147(3):204-217.

Campolongo F, Cariboni J, Saltelli A. 2007. An effective screening design for sensitivity analysis of large models. Environmental Modelling & Software 22(10):1509-1518.

Chen K, Teo S, Seng KY 2000. Sensitivity analysis on a physiologically-based pharmacokinetic/pharmacodynamic model for diisopropylfluorophosphate induced toxicity in mice and rats. Toxicology Mechanisms & Methods 19:486-497.

Chevret S, ed. 2006. Statistical methods for dose-finding experiments. Chichester, West Sussex, England ; Hoboken, NJ: John Wiley and Sons.

Clewell RA, Clewell HJ. 2008. Development and specification of physiologically based pharmacokinetic models for use in risk assessment. Regulatory Toxicology and Pharmacology 50(1):129-143.

Cohen JT, Lampson MA, Bowers TS. 1996. The use of two-stage Monte Carlo simulation techniques to characterize variability and uncertainty in risk analysis. Human & Ecological Risk Assessment 2(4):939-971.

Cullen AC, Frey HC. 1999. Probabilistic techniques in exposure assessment: a handbook for dealing with variability and uncertainty in models and inputs. New York: Plenum Press.

Curwin BD, Hein MJ, Sanderson, WT, Striley C, Heederik D, Kromhout H, Reynolds SJ, Alavanja MC. 2007. Urinary pesticide concentrations among children, mothers and fathers living in farm and non-farm households in Iowa. The Annals of Occupational Hygiene 51: 53-65.

Devroye L. 1986. Non-uniform random variate generation. New York: Springer-Verlag.

Dorne JLCM, Walton K, Renwick AG. 2005. Human variability in xenobiotic metabolism and pathway-related uncertainty factors for chemical risk assessment: a review. Food and Chemical Toxicology 43(2):203-216.

Dow Agrosciences (DAS). 2011. Source-to-outcome modeling physiologically based pharmacokinetic/pharmacodynamic (PBPK/PD) model linked to a dietary exposure model: chlorpyrifos as a case study. Science issues paper for FIFRA SAP Meeting, February 15–18, 2011, Washington, D.C., Office of Pesticide Programs Regulatory Docket: EPA-HQ-OPP-2010-0588.

Eaton DL, Daroff RB, Autrup H, Bridges J, Buffler P, Costa LG, Coyle J, McKhann G, Mobley WC, Nadel L, Neubert D, Schulte-Hermann R, Spencer PS. 2008. Review of the toxicology of chlorpyrifos with an emphasis on human exposure and neurodevelopment. Critical Reviews in Toxicology 38(2):1-125.

Dunker AM. 1984. The decoupled direct method for calculating sensitivity coefficients in chemical-kinetics. Journal of Chemical Physics 81(5):2385-2393.

Ferson S, Tucker WT. 2006. Sensitivity analysis using probability bounding. Reliability Engineering & System Safety 91(10-11):1435-1442.

Frey HC, Rhodes DS. 1996. Characterizing, simulating, and analyzing variability and uncertainty: An illustration of methods using an air toxics emissions example. Human & Ecological Risk Assessment 2(4):762-797.

Gamerman D, Lopes HF. 2006. Markov chain Monte Carlo: Stochastic simulation for Bayesian inference. 2nd ed. Boca Raton: Taylor & Francis.

Gelman A. 2004. Bayesian data analysis. 2nd ed. Boca Raton, Fla.: Chapman & Hall/CRC.

Gordon J, Shortliffe EH. 1984. The Dempster-Shafer theory of evidence. In: Rule-based expert systems: the MYCIN experiments of the Stanford Heuristic Programming Project (Buchanan BG, Shortliffe EH, eds). Reading, Mass.: Addison-Wesley, xix, 748 p.

Granvilliers L, Cruz J, Barahona P. 2004. Parameter estimation using interval computations. Siam Journal on Scientific Computing 26(2):591-612.

Gueorguieva II, Nestorov IA, Rowland M. 2004. Fuzzy simulation of pharmacokinetic models: Case study of whole body physiologically based model of diazepam. Journal of Pharmacokinetics & Pharmacodynamics 31(3):185-213.

Haddad S, Tardif GC, Tardif R. 2006. Development of physiologically based toxicokinetic models for improving the human indoor exposure assessment to water contaminants: trichloroethylene and trihalomethanes. Journal of Toxicology & Environ Health A 69(23):2095-2136.

Halpern JY. 2003. Reasoning about uncertainty. Cambridge, Mass.: MIT Press.

Halton JH. 1960. On the efficiency of certain quasi-random sequences of points in evaluating multi-dimensional integrals. Numerische Mathematik 2(1):84-90.

Helton JC, Davis FJ. 2003. Latin hypercube sampling and the propagation of uncertainty in analyses of complex systems. Reliability Engineering & System Safety 81(1):23-69.

Hibbert DB, Armstrong N. 2009. An introduction to Bayesian methods for analyzing chemistry data: Part II: A review of applications of Bayesian methods in chemistry. Chemometrics & Intelligent Laboratory Systems 97(2):211-220.

Hickey T, Ju Q, Van Emden MH. 2001. Interval arithmetic: From principles to implementation. Journal of the ACM 48(5):1038-1068.

ED_005731_00001263-00073

Hirschfeld S, Songco D, Kramer BS, Guttmacher AE, 2011. National Children's Study: Update in 2010. Mt Sinai J Med 78:119–125.

International Programme on Chemical Safety (IPCS). 2006. Guidance on Uncertainty in Exposure Assessment. World Health Organization, International Programme on Chemical Safety. http://www.who.int/ipcs/methods/harmonization/areas/exposure_assessment/en/index.html

Ionescu-Bujor M, Cacuci DG. 2004. A comparative review of sensitivity and uncertainty analysis of large-scale systems - I: Deterministic methods. Nuclear Science & Engineering 147(3):189-203.

Isukapalli SS, Georgopoulos PG. 2001. Computational methods for sensitivity and uncertainty analysis for environmental and biological models (EPA/600/R-01-068). EPA/600/R-01-068. Research Triangle Park, NC: Prepared for the US Environmental Protection Agency National Exposure Research Laboratory.

Isukapalli SS, Roy A, Georgopoulos PG. 1998. Stochastic Response Surface Methods (SRSMs) for uncertainty propagation: Application to environmental and biological systems. Risk Analysis 18(3):351-363.

Isukapalli SS, Roy A, Georgopoulos PG. 2000. Efficient sensitivity/uncertainty analysis using the combined stochastic response surface method and automated differentiation: Application to environmental and biological systems. Risk Analysis 20(5):591-602.

Isukapalli SS, Spendiff M, Georgopoulos PG, Krishnan K. 2010. Uncertainty, variability and sensitivity analyses in simulation models. In: Quantitative Modeling in Toxicology (Krishnan K, Andersen ME, eds): John Wiley, 429-458.

Jaulin L. 2001. Applied interval analysis. London ; New York: Springer.

Jelliffe R, Schumitzky A, Bayard D, Leary R, Botnen A, Guilder MV, et al. 2009. Human genetic variation, population pharmacokinetic - dynamic models, Bayesian feedback control, and maximally precise individualized drug dosage regimens. Current Pharmacogenomics and Personalized Medicine 7:249-262.

Jiang W, Duysen EG, Hansen H, Shlyakhtenko L, Schopfer LM, Lockridge O. 2010. Mice treated with chlorpyrifos or chlorpyrifos oxon have organophosphorylated tubulin in the brain and disrupted microtubule structures, suggesting a role for tubulin in neurotoxicity associated with exposure to organophosphorus agents. Toxicological Sciences 115(1):183-193.

Jones DR, Peters JL, Rushton L, Sutton AJ, Abrams KR. 2009. Interspecies extrapolation in environmental exposure standard setting: A Bayesian synthesis approach. Regulatory Toxicology and Pharmacology 53(3):217-225.

ED_005731_00001263-00074

Jonsson F, Jonsson EN, Bois FY, Marshall S. 2007. The application of a Bayesian approach to the analysis of a complex, mechanistically based model. Journal of Biopharmaceutical Statistics 17(1):65-92.

Kisicki JC, Seip CW, Combs ML. 1999. A rising dose toxicology study to determine the no-observable-effect-levels (NOEL) for erythrocyte acetylcholinesterase (AChE) inhibition and cholinergic signs and symptoms of chlorpyrifos at three dose levels. Report of MDS Harris (Project# 21438), Lincoln, Nebraska. Sponsored by Dow AgroSciences, Indianapolis, Indiana (MRID 44811002 & MRID 45144101).

Klir GJ, Yuan B. 1995. Fuzzy sets and fuzzy logic : theory and applications. Upper Saddle River, N.J.: Prentice Hall PTR.

Kocis L, Whiten WJ. 1997. Computational investigations of low-discrepancy sequences. Acm Transactions on Mathematical Software 23(2):266-294.

Lafortuna, C. L., et al., 2005. Gender variations of body composition, muscle strength and power output in morbid obesity. International Journal of Obesity (London) 29: 833–841.

Lefkowitz, L.J., Kupina, J.M., Hirth, N.L., Henry, R.M., Noland, G.Y., Barbee, J.Y. Jr., Zhou, J.Y., Weese, C.B. 2007. Intraindividual stability of human erythrocyte cholinesterase activity. Clinical Chemistry 53(7):1358-63. Epub 2007 May 10

Lester RR, Green LC, Linkov I. 2007. Site-specific applications of probabilistic health risk assessment: Review of the literature since 2000. Risk Analysis 27(3):635-658.

Li GY, Wang SW, Rabitz H, Wang SY, Jaffe P. 2002. Global uncertainty assessments by high dimensional model representations (HDMR). Chemical Engineering Science 57(21):4445-4460.

Li GY, Wang SW, Rosenthal C, Rabitz H. 2001. High dimensional model representations generated from low dimensional data samples. 1. mp-Cut-HDMR. Journal of Mathematical Chemistry 30(1):1-30.

Linkov I, Burmistrov D. 2003. Model uncertainty and choices made by modelers: Lessons learned from the international atomic energy agency model intercomparisons. Risk Analysis 23(6):1297-1308.

Loizou G, Spendiff M, Barton HA, Bessems J, Bois FY, d'Yvoire MB, et al. 2008. Development of good modelling practice for physiologically based pharmacokinetic models for use in risk assessment: The first steps. Regulatory Toxicology and Pharmacology 50(3):400-411.

Lowe ER, Poet TS, Rick DL, Marty MS, Mattsson JL, Timchalk C, Bartel MJ. 2009. The effect of plasma lipids on the pharmacokinetics of chlorpyrifos and the impact on interpretation of blood biomonitoring data. Toxicological Sciences 108:258-272.

ED_005731_00001263-00075

Lu C, Barr DB, Pearson MA, Waller LA. 2008. Dietary intake and its contribution to the longitudinal organophosphorus pesticide exposure in urban and suburban children. Environmental Health Perspectives 116(4)537:542.

Lu C, Bravo R, Caltabiano LM, Irish RM, Weerasekera G, Barr DB. 2005. The presence of dialkylphosphates in fresh fruit juices: implication for organophosphorus pesticide exposure and risk assessments. Journal of Toxicology and Environmental Health, Part A 68(3): 209-227.

Lu C, Schenck FJ, Pearson MA, Wong JW. 2010. Assessing children's dietary pesticide exposure: direct measurement of pesticide residues in 24-hr duplicate food samples. Environ Health Perspect 118(11): doi:10.1289/ehp.1002044.

MacIntosh DL, Kabiru CW, Echols SL, Ryan PB. 2001. Dietary exposure to chlorpyrifos and levels of 3, 5, 6-trichlor-2-pyridinol in urine. Journal of Exposure Analysis and Environmental Epidemiology 11(4):279-285.

McRae GJ, Tilden JW, Seinfeld JH. 1982. Global sensitivity analysis--a computational implementation of the Fourier Amplitude Sensitivity Test (FAST). Computers & Chemical Engineering 6(1):15-25.

Moles CG, Mendes P, Banga JR. 2003. Parameter estimation in biochemical pathways: a comparison of global optimization methods. Genome Research 13(11):2467-2474.

Morris M, D. . 1991. Factorial sampling plans for preliminary computational experiments. American Society for Quality Control and American Statistical Association, p. 161-174.

Moschandreas DJ, Karuchit S. 2002. Scenario-model-parameter: a new method of cumulative risk uncertainty analysis. Environment International 28(4):247-261.

Moser VC, Chanda SM, Mortensen SR, Padilla S. Age- and gender-related differences in sensitivity to chlorpyrifos in the rat reflect developmental profiles of esterase activities. Toxicological Sciences 1998 46, 211-222.

National Research Council (NRC). 1993. Pesticides in the diets of infants and children. National Academy Press. Washington, D.C.

National Research Council (NRC). 2007. Toxicity testing in the 21st century: A vision and a strategy. National Academies Press. Washington, D.C.

Nestorov I. 2001. Modelling and simulation of variability and uncertainty in toxicokinetics and pharmacokinetics. Toxicology Letters 120(1-3):411-420.

Nilsen T, Aven T. 2003. Models and model uncertainty in the context of risk analysis. Reliability Engineering & System Safety 79(3):309-317.

ED_005731_00001263-00076

Nolan RJ.; Rick DL; Freshner N. L.; et al. 1984. Chlorpyrifos: pharmacokinetics in human volunteers following single oral and dermal doses. Toxicology and Applied Pharmacology 73: 8-15.

Nong A, Krishnan K. 2007. Estimation of interindividual pharmacokinetic variability factor for inhaled volatile organic chemicals using a probability-bounds approach. Regulatory Toxicology and Pharmacology 48(1):93-101.

Nong A, Tan YM, Krolski ME, Wang J, Lunchick C, Conolly RB Clewell HJ. 2008. Bayesian calibration of a physiologically based pharmacokinetic/pharmacodynamic model of carbaryl inhibition. Journal of Toxicology & Environmental Health Part A 7: 1363-1381.

Norgan NG, 2005. Laboratory and field measurements of body composition. Public Health Nutr 8:1108–1122.

Oberg T, Bergback B. 2005. A review of probabilistic risk assessment of contaminated land. Journal of Soils and Sediments 5(4):213-224.

Parsons S. 2001. Qualitative methods for reasoning under uncertainty. Cambridge, Mass.: MIT Press.

Parzen, E. 1962. Chapter 6, Markov chains: discrete parameter in stochastic processes. In: Stochastic Processes. Holden-Day, Inc., San Francisco. p. 288-274.

Pavek P, Ceckova M, Staud F. 2009. Variation of drug kinetics in pregnancy. Current Drug Metabolism 10:520–529.

Pope, C.N., Karanth, S., Liu, J., Yan, B. 2005. Comparative carboxylesterase activities in infant and adult liver and their in vitro sensitivity to chlorpyrifos oxon. Regulatory Toxicology and Pharmacology 42:64-69.

Price K, Haddad S, Krishnan K. 2003. Physiological modeling of age-specific changes in the pharmacokinetics of organic chemicals in children. Journal of Toxicology & Environmental Health, Part A 66(5):417-433.

Pujol G, Corre B. 2007. Simplex screening. In: Fifth international conference of Sensitivity Analysis of Model Output (SAMO2007). Budapest, Hungary.

Rabitz H, Alis OF. 1999. General foundations of high-dimensional model representations. Journal of Mathematical Chemistry 25(2-3):197-233.

Read AL. 2002. Presentation of search results: the CLs technique. Journal of Physics G: Nuclear and Particle Physics 28(10):2693.

Redding LE, Sohn MD, McKone TE, Wang S-L, Hsieh PH, and Yang RSH. 2008. Population

ED_005731_00001263-00077

physiologically-based pharmacokinetic modeling for the human lactational transfer of PCB 153 with consideration of worldwide human biomonitoring results. Environmental Health Perspectives 116(12):1629-1634.

Reinert JM, Apostolakis GE. 2006. Including model uncertainty in risk-informed decision making. Annals of Nuclear Energy 33(4):354-369.

Riederer AM, Pearson MA, Lu C. 2009. Comparison of food consumption frequencies among HANES and CPES children: Implication for dietary pesticide exposure and risk assessment. Journal of Exposure Science and Environmental Epidemiology doi:10.1038/jes.2009.48. PMID: 19738638.

Royston JP 1982a. An extension of Shapiro and Wilks' test for normality to large samples, Applied Statistics 31: 115-124.

Royston JP 1982b. Algorithm AS 177: Expected normal order statistics (exact and approximate). Applied Statistics 31: 161-165.

Royston JP 1993. A pocket-calculator algorithm for the Shapiro-Francia test for non-normality: an application to medicine. Statistics in Medicine, 12, 181-184.

Rubinstein RY. 1981. Simulation and the Monte Carlo method. New York: Wiley.

Saltelli A. 2008. Global sensitivity analysis: the primer. Chichester, England ; Hoboken, NJ: John Wiley.

Saltelli A, Bolado R. 1998. An alternative way to compute Fourier amplitude sensitivity test (FAST). Computational Statistics & Data Analysis 26(4):445-460.

Saltelli A, Chan K, Scott EM. 2000. Sensitivity analysis. Chichester ; New York: Wiley.

Saltelli A, Tarantola S, Chan KPS. 1999. A quantitative model-independent method for global sensitivity analysis of model output. Technometrics 41(1):39-56.

Samsom TE, Hunter DL, Bushnell PJ. 2005. Effects of chronic dietary and repeated acute exposure to chlorpyrifos on learning and sustained attention in rats. Toxicological Sciences 87(2):460-468.

Sander P, Bergback B, Oberg T. 2006. Uncertain numbers and uncertainty in the selection of input distributions - Consequences for a probabilistic risk assessment of contaminated land. Risk Analysis 26(5):1363-1375.

Sandu A, Daescu DN, Carmichael GR. 2003. Direct and adjoint sensitivity analysis of chemical kinetic systems with KPP: Part I - theory and software tools. Atmospheric Environment 37(36):5083-5096.

ED_005731_00001263-00078

Sasso A, Isukapalli S, Georgopoulos P. 2010. A generalized physiologically-based toxicokinetic modeling system for chemical mixtures containing metals. Theoretical Biology and Medical Modelling 7(1):17.

Seng KY, Nestorov I, Vicini P. 2007. Simulating pharmacokinetic and pharmacodynamic fuzzy-parameterized models: a comparison of numerical methods. Journal of Pharmacokinetics and Pharmacodynamics 34(5):595-621.

Shapiro SS. Wilks MB. (1965), An analysis of variance test for normality (complete samples), Biometrika 52: 591-611.

Sidell FR aKaminskis A. 1975. Temporal intrapersonal physiological variability of cholinesterase in human plasma and erythrocytes. Clinical Chemistry 21:1961-1963.

Simon TW. 1999. Two-dimensional Monte Carlo simulation and beyond: A comparison of several probabilistic risk assessment methods applied to a superfund site. Human and Ecological Risk Assessment 5(4):823-843.

Smith JN, Timchalk C, Bartel MJ, Poet TS. 2011. *In vitro* age-dependent enzymatic metabolism of chlopyrifos and chlorpyrifos-oxon in human hepatic microsomes as well as chlorpyrifos-oxon in plasma. *Unpublished manuscript.*

Sobol IM. 2001. Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates. Mathematics and Computers in Simulation 55(1-3):271-280.

Sohn MD, McKone TE, and Blancato JN. Reconstructing Population Exposures from Dose Biomarkers: Inhalation of Trichloroethylene (TCE) as a Case Study. Journal of Exposure Analysis and Environmental Epidemiology, 14(3):204-213, 2004.

Stedeford T, Zhao QJ, Dourson ML, Banasik M, Hsu CH. 2007. The application of non-default uncertainty factors in the US EPA's Integrated Risk Information System (IRIS). Part I: UFL, UFS, and "Other uncertainty factors". Journal of Environmental Science and Health Part C-Environmental Carcinogenesis & Ecotoxicology Reviews 25(3):245-279.

Taylor P. 2001. Anticholinesterase agents. In Goodman and Gilman The Pharmacological Basis of Therapeutics, McGraw Hill, NY, pp 175-191.

Timchalk C, Nolan RJ, Mendrala AL, Dittenber DA, Brzak KA, Mattsson JL. 2002. A Physiologically based pharmacokinetic and pharmacodynamic (PBPK/PD) model for the organophosphate insecticide chlorpyrifos in rats and humans. Toxicological Sciences 66(1):34-53.

Timchalk C, Poet TS, Kousba AA. 2006. Age-dependent pharmacokinetic and pharmacodynamic response in preweanling rats following oral exposure to the organophosphorus insecticide chlorpyrifos. Toxicology. 220: 13-25

ED_005731_00001263-00079

United States Nuclear Regulatory Commission (USNRC). 2007. Guidance on the Treatment of Uncertainties Associated with PRAs in Risk-Informed Decision Making - Draft Report for Comment. NUREG-1855: US Nuclear Regulatory Commission, Office of Nuclear Regulatory Research. http://www.nrc.gov/reading-rm/doc-collections/nuregs/staff/sr1855/sr1855.pdf

Verner MA, Charbonneau M, Lopez-Carrillo L, Haddad S. 2008. Physiologically based pharmacokinetic modeling of persistent organic pollutants for lifetime exposure assessment: a new tool in breast cancer epidemiologic studies. Environmental Health Perspectives 116(7):886-892.

Wang SW, Georgopoulos PG, Li G, Rabitz H. 2005. Characterizing uncertainties in human exposure modeling through the Random Sampling-High Dimensional Model Representation (RS-HDMR) methodology. International Journal of Risk Assessment and Management 5:387-406.

Wang SW, Georgopoulos PG, Li GY, Rabitz H. 2003. Random sampling-high dimensional model representation (RS-HDMR) with nonuniformly distributed variables: Application to an integrated multimedia/multipathway exposure and dose model for trichloroethylene. Journal of Physical Chemistry A 107(23):4707-4716.

Wang Z, Klir GJ. 1992. Fuzzy measure theory. New York: Plenum Press.

Webb-Robertson B-JM, McCue LA, Lawrence CE. 2008. Measuring Global Credibility with Application to Local Sequence Alignment. PLoS Computational Biology 4(5):e1000077.

Wilson N.K., Chuang J.C., Lyu C., Menton R., and Morgan M.K. Aggregate exposures of nine preschool children to persistent organic pollutants at day care and at home. Journal of Exposure Assessment and Environmental Epidemiology 2003: 13(3): 187–202.

Wolkenhauer O. 1998. Possibility theory with applications to data analysis. Taunton, Somerset, England; New York: Research Studies Press ; Wiley.

Xue JP, Zartarian VG, Ozkaynak H, Dang W, Glen G, Smith L, et al. 2006. A probabilistic arsenic exposure assessment for children who contact chromated copper arsenate (CCA)-treated playsets and decks, part 2: Sensitivity and uncertainty analyses. Risk Analysis 26(2):533-541.

Xue J, Zartarian V, Wang SW, Liu SV, Georgopoulos P. 2010. Probabilistic Modeling of Dietary Arsenic Exposure and Dose and Evaluation with 2003-2004 NHANES Data Environ Health Perspect 118(3):345-50

Yang D, Howard A, Bruun D, Ajua-Alemanj M, Pickart C, and Lein PJ. 2008. Chlorpyrifos and chlorpyrifos oxon inhibit axonal growth by interfering with the morphogenic activity of acetylcholinesterase. Toxicol. Appl. Pharmacol. 228: 32-41.

ED_005731_00001263-00080

Yang, D., Pearce, R.E., Wang, X., Gaedigk, R., Wan, Y-J Y, and Yan, B. 2009. Human carboxylesterases HCE1 and HCE2: Ontogenic expression, inter-individual variability and differential hydrolysis of oseltamivir, aspirin, deltamethrin and permethrin. Biochem Pharmacol. 77:238-247.

Zadeh LA. 1965. Fuzzy Sets. Information and Control 8:338-353.

Zadeh LA, Klir GJ, Yuan B. 1996. Fuzzy sets, fuzzy logic, and fuzzy systems: selected papers. River Edge, N.J.: World Scientific.

Zheng Q, Olivier K, Won YK, Pope CN. 2000. Comparative cholinergic neurotoxicity of oral chlorpyrifos exposures in preweanling and adult rats. Toxicology Sciences 55: 124-132.

Zio E, Apostolakis GE. 1996. Two methods for the structured assessment of model uncertainty by experts in performance assessments of radioactive waste repositories. Reliability Engineering & System Safety 54(2-3):225-241.

ED_005731_00001263-00081