UNITED STATES ENVIRONMENTAL PROTECTION AGENCY
WASHINGTON, DC 20460

OFFICE OF CHEMICAL SAFETY AND POLLUTION PREVENTION

July 11, 2012

**MEMORANDUM**

**SUBJECT:** Transmittal of Meeting Minutes of the FIFRA Scientific Advisory Panel Meeting held April 10-12, 2012 on "Chlorpyrifos Health Effects"

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

**FROM:** Fred Jenkins, Jr., Ph.D. 7/11/12
Designated Federal Official
FIFRA Scientific Advisory Panel
Office of Science Coordination and Policy

**THRU:** Laura Bailey, MS 7/11/12
Executive Secretary
FIFRA Scientific Advisory Panel
Office of Science Coordination and Policy

Frank Sanders 7/11/12
Director
Office of Science Coordination and Policy

Attached, please find the meeting minutes of the FIFRA Scientific Advisory Panel open meeting held in Arlington, VA on April 10-12, 2012. This report addresses a set of scientific issues associated with "Chlorpyrifos Health Effects."

Enclosure

cc:
Jim Jones
Louise Wise
Vicki Dellarco
William Jordan
Margie Fehrenbach
Keith Matthews
Donald Brady
Jack Housenger
Joan Harrigan-Farrelly
Lois Rossi
Jay Ellenberger
Karen Whitby
Richard Keigwin
Anna Lowit
William R. Mundy
Ginger Moser
Carol H. Christensen
Lieutenant Aaron Niman
Dale Kemery
Douglas Parsons
Vanessa Vu
OPP Docket

**FIFRA Scientific Advisory Panel Members**
Kenneth Portier, Ph.D.
Stephen Klaine, Ph.D.
Carey Pope, Ph.D.
Martha Sandy, Ph.D.

**FQPA Science Review Board Members**
Michael Aschner, Ph.D.
Brian Buckley, Ph.D.
Russell Carr, Ph.D.
Abby Collier, Ph.D.
Ellen Gold, Ph.D.
Gaylia Jean Harry, Ph.D.
Stella Koutros, Ph.D.
Chensheng (Alex) Lu, Ph.D.
Bette Meek, Ph.D.
William Popendorf, Ph.D.
Barry Ryan, Ph.D.
Alvin Terry, Ph.D.
John Vena, Ph.D.
Charles Vorhees, Ph.D.
Heather Young, Ph.D.

# SAP Minutes No. 2012-04

# A Set of Scientific Issues Being Considered by the Environmental Protection Agency Regarding Chlorpyrifos Health Effects

**April 10 – 12, 2012**
**FIFRA Scientific Advisory Panel Meeting**
**Held at the**
**Environmental Protection Agency Conference Center**
**Arlington, VA**

## NOTICE

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

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

In preparing these meeting minutes, the Panel carefully considered all information provided and presented by EPA, as well as information presented by public commenters.

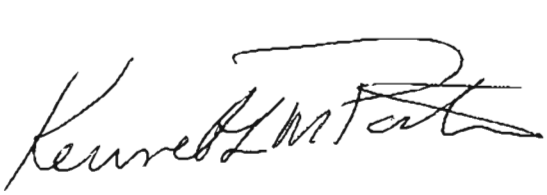
## TABLE OF CONTENTS

# SAP Minutes No. 2012-04

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

## Chlorpyrifos Health Effects

**April 10-12, 2012**
**FIFRA Scientific Advisory Panel Meeting**
**Held at the**
**Environmental Protection Agency Conference Center**
**Arlington, VA**



**Kenneth M. Portier, Ph.D.**
**FIFRA SAP Chair**
**FIFRA Scientific Advisory Panel**
**Date:** JUL 10 2012

**Fred Jenkins, Jr., Ph.D.**
**Designated Federal Official**
**FIFRA Scientific Advisory Panel**
**Date:** JUL 10 2012

**Federal Insecticide, Fungicide, and Rodenticide Act (FIFRA) Scientific Advisory Panel (SAP)**
**Chlorpyrifos Health Effects**
**April 10 – 12, 2012**

**PARTICIPANTS**

**FIFRA SAP Chair**

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

**Designated Federal Official**

**Fred Jenkins, Jr., 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
Phone (202) 564-3327
Fax (202) 564-8382
jenkins.fred@epa.gov

**FIFRA Scientific Advisory Panel Members**

**Stephen Klaine, Ph.D.**
Professor
Environmental Toxicology & Biological Sciences
Clemson Institute of Environmental Toxicology
Pendleton, SC

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

**Martha Sandy, Ph.D.**
Senior Toxicologist and Chief
Cancer Toxicology and Epidemiology Section
Reproductive and Cancer Hazard Assessment Branch
California Environmental Protection Agency
Oakland, CA

**FQPA Science Review Board Members**

**Michael Aschner, Ph.D.**
Director
Division of Pediatric Toxicology
Medical Center
Vanderbilt University Medical Center
Nashville, TN

**Brian Buckley, Ph.D.**
Executive Director of Laboratories
Environmental & Occupational Health Sciences Institute
Rutgers University
Piscataway, NJ

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

**Abby Collier, Ph.D.**
Assistant Professor of Pharmacology
John A. Burns School of Medicine
University of Hawaii at Manoa, Kaka'ako Campus
Honolulu, HI

**Ellen Gold, Ph.D.**
Professor and Chair
Department of Public Health Sciences
University of California-Davis
Davis, CA

**Gaylia Jean Harry, Ph.D.**
Biologist, Head of Neurotoxicology Group
Division of Intramural Research
National Institute of Environmental Health Sciences
Research Triangle Park, NC

**Stella Koutros, Ph.D.**
Research Fellow
Division of Cancer Epidemiology and Genetics
Epidemiology Branch
National Institutes of Health
Rockville, MD

**Chensheng (Alex) Lu, Ph.D.**
Associate Professor
Department of Environmental Health
Exposure, Epidemiology and Risk Program
Harvard School of Public Health
Boston, MA

**Bette Meek, Ph.D.**
Associate Director Chemical Risk Assessment
McLaughlin Institute for Population Health Risk Assessment
University of Ottawa
Ottawa, Canada

**William Popendorf, Ph.D.**
Professor, Emeritus
Department of Biology
Utah State University
Logan, UT

**Barry Ryan, Ph.D.**
Professor
Exposure Assessment & Environmental Chemistry
Department of Environmental and Occupational Health
Rollins School of Public Health
Emory University
Atlanta, GA

**Alvin Terry, Ph.D.**
Professor
Pharmacology and Toxicology
Georgia Health Sciences University
Augusta, GA

**John Vena, Ph.D.**
UGA Foundation Professor & Department Head
B.S. Miller Hall Room 105
Health Sciences Campus
University of Georgia
Athens, GA

**Charles Vorhees, Ph.D.**
Professor
Department of Pediatrics, Div. of Neurology
College of Medicine
University of Cincinnati and
Cincinnati Children's Hospital
Cincinnati, OH

**Heather Young, Ph.D.**
Associate Professor
Associate Chair
Department of Epidemiology & Biostatistics
School of Public Health & Health Services
George Washington University
Washington, DC

# INTRODUCTION

Chlorpyrifos (0,0-diethyl--3,5,6-trichloro -2-pyridyl phosphorothioate) is a broad-spectrum, chlorinated organophosphate (OP) insecticide. Like other OPs, chlorpyrifos binds to and phosphorylates the enzyme acetylcholinesterase (AChE) in both the central (brain) and peripheral nervous systems. This can lead to accumulation of acetylcholine and, ultimately, at sufficiently high doses, to clinical signs of toxicity. In 2011, the Agency released a preliminary human health risk assessment for chlorpyrifos. The focus of this assessment was on the cholinesterase (ChE) inhibiting potential of chlorpyrifos. Consistent with this focus, EPA evaluated the extensive database of ChE data for multiple lifestages and selected points of departure (PoDs) based on consideration of all quality and reliable data. There is, however, a growing body of literature with laboratory animals (rats and mice) indicating that gestational and/or early postnatal exposure to chlorpyrifos may cause persistent effects into adulthood. The results of both *in vivo* and *in vitro* studies on chlorpyrifos have led some research groups to propose that changes in brain connectivity and/or neurochemistry may underlie these changes into adulthood. In addition, there are epidemiology studies evaluating pre- and post-natal chlorpyrifos or other OP exposure in mother-infant pairs that have reported associations with birth outcomes, childhood neurobehavioral and neurodevelopment outcomes in the offspring when evaluated in neonates, infants, and young children.

In 2008, the FIFRA Scientific Advisory Panel (SAP) reviewed a draft science issue paper on the human health effects of chlorpyrifos which provided a preliminary review of the scientific literature on experimental toxicology and epidemiology studies available at that time. In 2010, the Agency developed a draft "Framework for Incorporating Human Epidemiologic & Incident Data in Health Risk Assessment" which provides the conceptual foundation for evaluating multiple lines of scientific evidence in the context of the understanding of the adverse outcome pathway (or mode of action). This draft framework uses modified Bradford Hill Criteria to evaluate the sufficiency of evidence to establish key events within a mode of action(s) and explicitly considers such concepts as strength, consistency, dose response, temporal concordance and biological plausibility.

Since the 2008 SAP on chlorpyrifos, the Agency has performed further analyses on the existing and new epidemiology results in mothers and children, available biomonitoring data, and experimental toxicology studies evaluating proposed adverse outcome pathways in the context of human health risk assessment. Specifically, the Agency is evaluating available literature on the potential for chlorpyrifos to cause long term adverse effects from early life exposure, *in vivo* and *in vitro* studies evaluating mechanistic aspects of chlorpyrifos, and the potential for adverse effects below doses established from ChE inhibition that are used for regulatory purposes. At this time, the Agency is working towards a weight of evidence evaluation integrating the epidemiology studies with the experimental toxicology studies for the neurodevelopmental outcomes. This analysis is complex and multifaceted as it involves different lines of scientific evidence (*i.e.*, *in vivo* & *in vitro* experimental toxicology studies, explicit consideration of adverse outcome pathways, exposure, epidemiology, and biomonitoring data). As such, the Agency believes that peer review on the status of the current analysis is important.

Opening remarks at the meeting were provided by:
Steven Bradbury, Ph.D., Director, Office of Pesticide Programs (OPP), EPA; Karen Whitby, Ph.D., Acting Director, Health Effects Division, OPP, EPA; Anna Lowit, Ph.D., OPP, EPA; William R. Mundy, Ph.D., Office of Research and Development (ORD), EPA; Ginger Moser, Ph.D., DABT, Fellow ATS, ORD, EPA; Carol H. Christensen, Ph.D., MPH, OPP, EPA; Lieutenant Aaron Niman, US Public Health Service, OPP EPA

## Public Comments

Public comments were provided by:
Dow AgroSciences
Julie E. Goodman Ph.D. and Lorenz Rhomberg, Ph.D. of Gradient on behalf of Dow Agrosciences
Abby Li, Ph.D. of Exponent on behalf of Dow Agrosciences
Jennifer Sass, Ph.D. of Natural Resources Defense Council (NRDC)
Dale Hattis, Ph.D. of Clark University on behalf of himself

## Summary of Panel Discussion and Recommendations

### *Charge 1: Mode of action/adverse outcome pathway: Acetylcholinesterase (AChE) inhibition*

Question 1.0

It is well established that AChE inhibition is the primary mode of action/adverse outcome pathway for OPs, like chlorpyrifos. Because AChE inhibition is the initiating event for this mode of action/adverse outcome pathway, using AChE inhibition as a regulatory endpoint is protective of downstream cholinergic effects. Moreover, historically, given the sensitivity of AChE inhibition data for OPs, these data have been considered to be protective of other potential toxicities and/or modes of action for OPs. In 2008, the Agency performed a comprehensive review of the available AChE data from multiple lifestages. This review has been supplemented with the newest studies. Consistent with the recommendations from the 2008 SAP, the Agency believes that AChE data remain the most robust dose-response data for deriving points of departure in *in vivo* experimental toxicology studies with laboratory animals. *Please comment on the Agency's preliminary conclusion that AChE data remain the most robust source of data for deriving points of departure for chlorpyrifos. Please include a discussion of the strengths and uncertainties of this preliminary conclusion.*

The Panel concurs with the Agency's position that AChE data continue to be the strongest resource of data for deriving points of departure for chlorpyrifos. The Panel's conclusion is based on the premise that all studies reporting neurobehavioral changes following *in vivo* prenatal or postnatal exposures to chlorpyrifos have been accompanied by AChE inhibition when measured at an appropriate time following administration of chlorpyrifos.

The Panel additionally notes that studies evaluating neurodevelopmental effects entailed experimental designs that do not permit an efficient means of determining a point of departure for chlorpyrifos. Thus, just as the in the 2008 SAP, this Panel advises that the Agency continue to use AChE data at the most sensitive lifestages for dose-response analysis and deriving points of departure. Also in keeping with the 2008 SAP, this Panel expresses concern about the use of Dimethyl Sulfoxide (DMSO) as a vehicle because of its intrinsic toxicity, its potential influence on absorption and interaction with chlorpyrifos, and the impact of this interaction on the developing organism.

***Charge 2.0: Mode(s) of action/adverse outcome pathway(s): Plausible pathways leading to potential neurodevelopmental outcomes***

Question 2.1

As discussed in Section 3.2.1, although there are numerous mechanistic studies in the scientific literature, the research on different hypotheses does not provide sufficient data to establish causal linkages among different levels of biological organization to show how effects lead to adversity. As such, a mode of action or adverse outcome pathway leading to effects on the developing brain cannot be established at this time. Moreover, although multiple biologically plausible hypotheses are being pursued by researchers, based on the current state of the science, no one pathway has sufficient data to be considered more credible than the others. *Please comment on the Agency's preliminary conclusion that although there are multiple biologically plausible hypotheses being evaluated by research scientists, the mechanistic experimental toxicology data do not yet support a coherent set of key events in a mode of action/adverse outcome pathway.*

The Panel agrees with the Agency's conclusion that based on the current state of the science, no one pathway has sufficient data to be considered more credible than the others with respect to a causal link between chlorpyrifos exposure and neurodevelopmental outcome.

In regard to the Agency's case study demonstrating domoic acid's adverse outcome pathway, the Panel contends that the linear connections of the pathway demonstrated in this case study appear likely to be rare and unique. They also note that it is more likely that other such neurotoxicological pathways are non-linear. Expectations of a linear pathway specifically in the case of chlorpyrifos may be artificially elevated and potentially unrealistic for risk assessment.

Question 2.2

Although a mode of action/adverse outcome pathway has not been established, qualitatively, the growing body of mechanistic studies does demonstrate that chlorpyrifos and/or its oxon are biologically active on a number of processes that affect the developing brain. Some mechanistic studies provide evidence of possible effects which are similarly sensitive or more sensitive than AChE inhibition (*e.g.*, neurite outgrowth, binding to muscarinic receptors, axonal transport; serotonergic nervous system development). Some of these comparisons must be considered with caution since the amount of change in the *in vitro* systems required to elicit an adverse effect *in vivo* is unknown. Moreover, extrapolation from *in vitro* perturbations to *in vivo* effects has not been established, which introduces additional uncertainties. *Given the doses/concentrations evaluated in the in vitro and in vivo mechanism studies, please comment on the degree to which these studies suggest that endpoints relevant to evaluating potential neurodevelopmental outcomes may or may not be more sensitive than AChE inhibition. Please include in your comments a discussion of the strengths and uncertainties. Please also include in your comments a discussion of the scientific understanding of dose-response relationships for biological perturbations and the magnitude, frequency and/or duration of such perturbations that can lead to adverse effects at higher levels of biological organization*

*to support characterization of the likelihood of adverse outcomes in a human health risk assessment (as articulated in NRC 2007).*

The Panel concurs with the Agency that caution should be applied in interpreting the *in-vivo* significance of the changes observed across the various *in vitro* studies. Several uncertainties and limitations are associated with the translation of *in vitro* study results to *in vivo* effects. The inherent complexity of the nervous system presents significant challenges to accomplishing this translation. An additional example of uncertainty is that cells that are isolated in culture within an *in vitro* experiment may be affected differently than they would if they were within their *in vivo* environment.

The Panel recommends continued literature review and analysis of published data with the goal of developing additional hypotheses linking *in vitro* findings to *in vivo* relevance. As an example, the analytical studies of the Lockridge group indicating that chlorpyrifos oxon can covalently modify key cytoskeletal proteins such as tubulin and motor proteins like kinesin, provide information that can contribute to the interpretation of findings of alterations in neurite outgrowth and axonal transport, respectively.

The Panel also recommends that the Agency consider other areas that might be added to the review such as the effect of chlorpyrifos on neurotrophins (growth factors). Several researchers have found early evidence of the potential for these effects (Pope *et al.*, 1995; Slotkin *et al.*, 2007; Betancourt and Carr, 2004).

The Panel cautions the Agency concerning their examination of the dose-response relationships. They particularly note that when evaluating these relationships, pharmocodynamic (PD) analyses should not be uncoupled from pharmacokinetic (PK) models given that PK differences can affect active site concentrations and hence, PD effects. Thus, PK models can significantly affect the magnitude and duration of an effect.

Lastly, the Panel raises concerns about the equivalency of developmental stages between ages of rodents to human. These are not well defined with regard to cell type compositions, brain region, cellular architecture, and physiological or biochemical processes. This lack of equivalence further limits the translation to the *in vivo* situation and the ability to provide a quantitative dose-response relationship that can be compared to that for AChE inhibition.

.

***Charge 3.0 Neurodevelopmental data from laboratory animals***

Question 3.1

As discussed in Section 3.2.2, the experimental toxicology data in laboratory rodents show neurobehavioral effects following developmental exposure with changes in a number of neurological domains. In 2008, the SAP agreed to this preliminary conclusion, and the nine additional studies available since 2008 add further support. *Please comment on the degree to which these studies show changes in a number of neurological domains and support the qualitative conclusion that chlorpyrifos exposure during gestation and/or early post-natal period may result in long-term adverse effects on the developing nervous system. What evidence does and does not support this conclusion? Please also include in your comments a discussion of the strengths and uncertainties. Please also include in your comments a discussion of the scientific understanding of dose-response relationships for biological perturbations and the magnitude, frequency and/or duration of such perturbations that are can lead to adverse effects at higher levels of biological organization to support characterization of the likelihood of adverse outcomes in a human health risk assessment (as articulated in NRC 2007).*

The Panel agrees with the 2008 SAP conclusions that developmental neurobehavioral studies demonstrate adverse effects from chlorpyrifos exposure. However, the number of available neurobehavioral studies is limited leading to caution concerning this finding. Also many of these studies are statistically under-powered and prone to Type I errors and should be discounted in formulating the weight of evidence for or against neurobehavioral effects from developmental exposure to chlorpyrifos. The Panel also expressed caution with the significance of some of the experimental neurotoxicological outcomes that have not been validated. These included the tests of anxiety, depression, and social interactions. The Panel recommends these experimental outcomes be regarded as exploratory, and hypothesis-generating, as opposed to being evidence of toxicity. The lack of specificity in the direction of the neurobehavioral dose response findings is a problematic issue.

Despite the issues raised by the Panel about these studies, the overall evidence across these studies is persuasive in indicating that there are enduring effects on the Central Nervous System (CNS) from chlorpyrifos exposure at or above 1.0 mg/kg. The Panel recommends that future neurodevelopmental studies be focused on testing chlorpyrifos levels below 1.0 mg/kg/day and that these studies be geared towards identifying the correct testing paradigm and neural substrates for detecting possible effects. The Panel advises that cross-laboratory or collaborative studies may provide systematic comparison of the effects of chlorpyrifos on neurodevelopmental domains using unified exposure periods, dosing, age of testing, and methods, combined with urinary analysis of chlorpyrifos' metabolites, and accurate assessments of AChE inhibition.

Question 3.2

The dose-response data in the *in vivo* experimental neurodevelopmental toxicity studies are not amenable to empirical dose-response modeling as many studies use only one or two doses, and in some cases the lower dose, but not higher dose level, produced significant effects. Many studies report effects at a dose of 1 mg/kg/d-- a dose that produces some amount of brain ChE inhibition when given directly to the pups post-natally, but may or may not alter fetal brain ChE activity when given to the dams gestationally. One study (Braquenier *et al.*, 2010) using lower doses, administered to the dam on GD15-LD14, reported a NOEL of 0.2 mg/kg/d. Comparing the NOEL of 0.2 mg/kg/d to a repeated dosing AChE inhibition $BMDL_{10}$ of 0.03 mg/kg/d suggests that AChE inhibition is a sensitive and protective endpoint.

*a. Please comment on the scientific quality and robustness of the animal neurodevelopmental toxicity studies.*

The Panel notes that the quality of these studies vary. The "high quality" studies use multiple doses, adequate sample sizes, controls for litter effects, sound behavioral methods, and appropriate statistical methods to analyze the data. Since all these studies demonstrate long-term neurobehavioral effects, the data generated by them (especially the findings that occurred at doses $\geq$ 1 mg/kg of chlorpyrifos) can be considered robust. The Panel has some concerns even with the high quality studies. For example, the rat strains used in some of the studies are considered by the Panel to be less than preferable. The Panel advises that studies that are considered to support regulatory decisions should be those that use a mainstream rat strain such as Sprague-Dawley from Charles River or Harlan because much more is known about their behavioral characteristics and they do not perform at the extremes of the distribution.

Despite the concerns expressed about the studies, the Panel concurs with the conclusions of the 2008 SAP findings and the EPA White Paper background document, and concludes that the collective weight of evidence from these studies demonstrate that it is probable that there are significant long-term adverse effects from chlorpyrifos exposure.

*b. Please comment on the degree to which studies that measured AChE inhibition and those that measured neurodevelopmental outcomes can be integrated to evaluate whether points of departure based on 10% AChE inhibition provide more sensitive endpoints than endpoints measured in the experimental neurodevelopmental studies (as reviewed in Section 3.2.2). Please include in your comments a consideration of the strengths and uncertainties associated with this assessment.*

The Panel concludes that since AChE inhibition recovers quickly the data are insufficiently refined to allow for a linkage between the mode of action and the neurodevelopmental effects (acute vs. chronic, respectively). They note that since the mode of action of these effects is not established and cannot be presumed to be related to AChE inhibition, these studies do not exclude the possibility that other mechanisms may

be involved, especially concerning long-term effects that may be unmasked at later life-stages. Additionally, since the neurodevelopmental effects may be independent of AChE inhibition, the Agency should consider whether AChE inhibition represents the critical marker for derivation of points of departure for chronic studies.

***Charge 4.0 Epidemiology Regarding Children's Health***

Question 4.1
Section 4.0 and Appendices 5 and 6 provide the Agency's review of the available epidemiology studies from the Columbia Mothers and Newborn study, the Mt. Sinai Child Development study, and the Center for Health Assessment of Mothers and Children of Salinas Valley (CHAMACOS) study. Consistent with the 2008 SAP recommendations, the Agency has considered information offered from each of the three cohort investigations; however EPA acknowledges the primacy of the Columbia cohort data for the purposes of informing risk assessment because researchers measured chlorpyrifos parent compound directly in this study. *Please comment on the sufficiency, clarity, and quality of the Agency's epidemiology review as contained in Section 4.0 and Appendices 5 and 6 of the draft issue paper with respect to identifying the major strengths and limitations of each study.*

The Panel considers the Agency's epidemiology review to be very clearly written, accurate, and to generally provide a very thorough review of the epidemiology literature. In addition, the Panel commends the Agency for putting their epidemiology review in the context of the modified Bradford Hill criteria, as recommended by the 2010 SAP. The Panel believes that the epidemiology review appropriately concludes that the studies show some consistent associations relating exposure measures to abnormal reflexes in the newborn, pervasive development disorder at 24 or 36 months, mental development at 7-9 years, and attention and behavior problems at 3 and 5 years of age, in addition to less consistent results for reduced mental and psychomotor development at 12 and 24 months. Inconsistent results are found for associations between exposure and measures of fetal growth.

The Panel views the Agency's epidemiology review as an excellent description of the strengths and limitations of the studies conducted to examine the relation of chlorpyrifos to children's growth and neurodevelopment. It is noted that studies of this nature are logistically difficult to implement because of the potentially large burden imposed on the study participants in terms of time and effort, often with little or no specific benefit to the participants.

Although in agreement with the Agency that chlorpyrifos could have played a role in the neurodevelopmental outcomes observed in the Columbia cohort, some panel members expressed concern about associating the observed deficits in neurodevelopmental outcomes in children with a single chemical. This is because the studies entail a multi-chemical exposure spanning a multi-year period that encompasses an important period of sequential developmental processes necessary for brain maturation. Thus, panel members caution that it is very difficult to attribute the independent physiological effects to a single chemical in this type of multi-chemical exposure scenario. An additional

concern raised by the Panel is the modest sample sizes of the studies. They deem inadequate sample size as one of the most important limitations of these studies.

Question 4.2

Similar to the initial conclusions from 2008, the Agency has preliminarily concluded that, qualitatively, chlorpyrifos likely played a role in the neurodevelopmental outcomes reported in the epidemiologic studies, and that information available since 2008, including both new etiologic investigations as well as epidemiologic methods papers, strengthens this conclusion. *Please comment on the Agency's preliminary, qualitative conclusion that chlorpyrifos likely played a role in the neurodevelopmental outcomes observed in the epidemiologic studies. Please include in your comments a discussion of the strengths and uncertainties associated with this preliminary conclusion.*

Overall, the Panel concurs with the 2008 SAP and the Agency in concluding that chlorpyrifos likely plays a role in impacting the neurodevelopmental outcomes examined in the three cohort studies. Although exposures to other AChE-inhibiting compounds cannot be excluded as contributing to neurodevelopmental (adverse) outcomes, the potential combination and/or additive effects of these compounds does not rule out the role of chlorpyrifos. As a result, it cannot be concluded that chlorpyrifos is the only contributor to the observed outcomes.

The strengths of the three studies support the Panel's conclusion. There are nine strengths identified by the Panel which are discussed in the detailed response section of this report. Some of the strengths noted include, but are not limited to: 1) the longitudinal designs which permit clear indications of the temporal relation of chlorpyrifos exposure to adverse neurodevelopmental outcomes, 2) the inclusion of biomarkers of exposure as well as self reported exposure, and 3) the relative consistency of findings in different populations while using similar standardized exposure and outcome measures.

Question 4.3

As discussed in Question 2.0, a mode of action/adverse outcome pathway has not yet been fully elucidated for the potential neurodevelopmental outcomes as a result of prenatal chlorpyrifos exposure. Although this does not undermine the qualitative interpretation of these studies, and the preliminarily conclusion stated above (Question 4.2), the identification of the dose-response for neurodevelopmental effects based on mode of action is not possible. Further, given the urine and cord blood sampling frequency in the study there is a large degree of uncertainty in estimating absolute exposure-response relationships, as opposed to establishing relative exposure groups for evaluating associations. With respect to dose-response, critical durations of exposure, and windows of susceptibility are unknown. In 2008, the SAP cautioned against using the Columbia cohort data for deriving a point of departure due, in part, to only measuring biomarkers (3rd trimester maternal, cord blood, meconium) at one point in time, and because they cannot exclude possibility that the effects seen were due to chlorpyrifos in combination with other pesticides. In 2008, the SAP advised against using data from the epidemiology studies (including the Columbia Mothers and Newborn study which measured chlorpyrifos directly) for deriving a point of departure due to limitations of the

exposure assessment in these epidemiology studies for the purpose of risk assessment, *e.g.*, lack of repeated exposure estimates to ascertain more specifically the variability and periodicity of exposure over time (*i.e.*, predominant use of one-time exposure estimate).

Question 4.3

*a.* Due to the limitations of exposure assessment performed in the epidemiologic investigations for the purposes of quantitative risk assessment, the Agency has concluded that the epidemiologic data are not sufficient for deriving points of departure for quantitative risk assessment. The Agency proposes that AChE inhibition data from laboratory animals remain the most appropriate data to use for dose-response modeling and the derivation of points of departure. *Please comment on the scientific evidence that does and does not support this conclusion, as well as the strengths and limitations of the evidence.*

The Panel recognizes the limitations of estimating chlorpyrifos exposures based on the exposure measures collected in the three longitudinal children's cohort studies (*i.e*., the Columbia study, the Mt. Sinai study, and the CHAMACOS study). Consequently, the Panel largely concurs with EPA that the data generated from these studies alone are not adequate enough to obtain a point of departure (POD) for the purposes of quantitative risk assessment.

However, despite the limitations of the exposure assessment of these three cohort studies, the Panel recognizes the significance of these data, and advises the Agency to explore additional ways of using these studies, especially the data from the Columbia study, to inform the dose response assessment of chlorpyrifos. This recommendation is underscored by the Panel's concerns regarding the proposed use of dose-response data on AChE inhibition in laboratory animals to derive points of departure for the chlorpyrifos risk assessment. The Panel notes that multiple lines of evidence suggest chlorpyrifos can affect neurodevelopment at levels lower than those associated with AChE inhibition. These multiple lines of evidence include: 1) the collective findings of the three cohort studies, 2) *in vivo* animal neurodevelopmental studies summarized in the Draft Issue Paper that report differential expression of oxidative stress genes and altered serotonergic tone in rat brain associated with early life chlorpyrifos exposures at doses below which acetylcholinesterase inhibition was detected, and 3) several *in vitro* mechanistic studies reported in the Draft Issue Paper demonstrating interference with neurite and axon outgrowth, reduced axonal transport, and increased oxidative stress in a variety of cell types exposed to chlorpyrifos concentrations that do not or are not expected to inhibit AChE. As noted in the response to Charge Question 2.2, additional evidence comes from studies not included in the Draft Issue Paper, reporting effects of chlorpyrifos on nerve growth factors and mitochondrial morphology at concentrations below which acetylcholinesterase inhibition is expected.

The Panel recommends that the Agency consider developing a functional physiologically based pharmacokinetic **(**PBPK) model for chlorpyrifos for pregnancy and the prenatal lifestage. The PBPK model could be utilized for additional dose-response analyses to further characterization of the dose estimates in the epidemiology studies. This model

could also become important in the future if the Agency decides to transition from using AChE inhibition to another outcome.

The Panel suggests additional research that may answer the key question of whether chlorpyrifos induces neurodevelopmental effects in humans at doses that do not cause AChE inhibition. More specifically the Panel suggests conducting studies that test whether red blood cell or brain AChE inhibition occurs as a result of chlorpyrifos concentrations in cord blood being associated with neurodevelopmental effects.

Additional Panel concerns about the use of AChE inhibition dose-response data to protect against neurodevelopmental effects is based on the potential for AChE inhibition and adverse neurodevelopmental effects to be two separate events. AChE inhibition is the result of an acute exposure scenario and neurodevelopmental effects likely being caused by chronic low level exposure to chlorpyrifos *in utero*.

Lastly, the Panel cautions the Agency dose-response data for AChE inhibition by chlorpyrifos in the pregnant rat may not be predictive of AChE inhibition in the human fetus given known interspecies differences in CYP450 isoforms, substrate affinities, fetal expression levels, and degree of polymorphism.

Question 4.3

*b.* The Agency does, however, believe that the epidemiologic data are useful to informing other key aspects of the chlorpyrifos risk assessment including hazard characterization, exposure characterization, and quantitative uncertainty characterization and analysis. *Please suggest approaches/analyses for potentially using the epidemiology data in different parts of the chlorpyrifos risk assessment including those noted above.* (Note: Some of these may also be covered in Question 5.4 below.)

The Panel agrees that the epidemiologic data are useful to inform key aspects of the chlorpyrifos risk assessment including exposure characterization, hazard characterization, and quantitative uncertainty characterization and analysis.

In regard to the exposure characterization, the Panel notes that environmental monitoring and biomonitoring data in these epidemiology studies can contribute to the overall database on estimation of exposure, including (particularly) population variability. These data can also enable the Agency to characterize exposure levels over time among diverse populations including production workers, agricultural workers, individuals exposed via residential use, general population, *etc*.

With respect to toxicological hazard characterization, the Panel suggests that these data can serve as the key source of support for the identification of prenatal exposures to chlorpyrifos as a cause of neurodevelopmental effects in humans. These data have many strengths. First there are consistencies in the findings of neurodevelopmental effects across the three cohort studies. Second, the levels of chlorpyrifos exposure experienced in these cohorts are comparable and well-characterized, being based on biomonitoring

(blood and urine measurements of chlorpyrifos, metabolites, *etc*.) and environmental monitoring measures (*e.g*., personal air monitoring in the Columbia study) and having similar levels observed in data collected from other studies of the general U.S. population (*e.g*., National Health and Nutrition Examination Survey (NHANES), for similar time periods (*i.e*., pre- and post-cancellation of residential uses).

In reference to the epidemiology data being used to support the quantitative uncertainty characterization and analysis, the Panel agrees with the 2008 SAP suggestion that at a minimum the Agency should use available data from these studies to at least "bound" reference doses developed on the basis of animal data. Given the potential significance of the epidemiological findings, the Panel advises the Agency to consider the potential impact of factors of study design and interpretation to bound the dose-response relationship from the human studies. For example, it would be useful to consider systematically (and at least semi-quantitatively) the potential impact of exposure measurement error, outcome ascertainment, confounding variables and statistical analysis methodology on the reported dose-response analysis.

To increase the confidence in the selected point of departure, the Panel recommends a simple experimental protocol to determine whether chlorpyrifos levels measured in the cord blood in the Columbia study inhibit either red blood cells or brain AChE. The results of such an exercise could potentially contribute to the essential question of whether or not a causal association between chlorpyrifos exposure and neurodevelopmental effects in the absence of AChE inhibition is plausible for humans.

***Charge 5.0 Exposure Profile & Biomonitoring Research***

Question 5.1

a. Section 5 of the draft issue paper presents an overview of the principal chlorpyrifos biomarkers and a comparison of biomonitoring studies that measured urinary TCPy levels in a range of study populations involving both the general population and potentially vulnerable populations, including children, workers, and farm families. *Please comment on the degree to which the Agency identified the primary chlorpyrifos biomarkers of exposure, appropriately discussed the strengths and limitations of such biomarkers, and how the strengths and limitations affect the interpretation of the chlorpyrifos biomonitoring data.*

The Panel notes that the Agency was thorough in its coverage of the literature on biomarkers of chlorpyrifos exposure. The Panel recommends that chlorpyrifos in blood be the first choice for a biomarker, particularly because of its specificity, the availability of standard methods for measuring it, the relevance of its concentration levels, and the number of laboratories that are capable of conducting the measurements. However, the Panel acknowledges that this is a most challenging assay, and has been used in only a small percentage of published research. The second biomarker of choice is 3,5,6-trichloro-2-pyridinol (TCPy) followed by diethylthiophosphate/diethylphosphate (DETP/DEP), both measured in urine. These have roughly the same equivalence and

neither is equivalent to measuring chlorpyrifos directly in blood because of the frequent presence of these environmental degradates of the active ingredient. Total DAPs (as DMP and DEP) are not selective enough to be a useful biomarker for chlorpyrifos although DAPs may be more appropriate in a global risk calculation model when all AChE inhibiting chemicals are considered together when evaluating risk.

The Panel suggests that more importance should be afforded to the direct intake of TCPy which is mainly present in foods. A growing body of research developing since the 1990s, has established the significance of this factor.

The Panel also acknowledges the capability of measuring AChE and BuChE as biomarkers of exposure. However, inhibitions of these enzymes are even less specific than DAPs, although they are more indicative of potential health risk. Unfortunately, the ability to measure small changes in these enzymes differs widely among laboratories and among study designs.

The Panel also recommends including in future considerations the phase II conjugation products of chlorpyrifos (namely, glucuronidase and sulfonates). Quantifying the conjugative metabolism will ensure that levels of biomarkers are correctly interpreted with respect to biomonitoring data and for performing reverse dosimetry.

*b.* Section 5 of the draft issue paper compares biomonitoring findings from the three children's health cohorts with other major observational exposure studies in the United States. Based on comparison with NHANES 2001-2002, median TCPy levels in the CHAMACOS and Mount Sinai cohorts were slightly higher than in the general population. It should be noted that the exposures experienced by the CHAMACOS and Mount Sinai cohorts overlapped the start of the residential chlorpyrifos phase-out. By contrast, median TCPy levels in the Columbia cohort, for which sampling occurred when chlorpyrifos use should have rapidly declined due to the voluntary cancelation, were slightly lower than the levels measured by NHANES in the general population. *Please comment on the adequacy of the Agency's comparison for the purposes of evaluating chlorpyrifos exposure levels in the three children's health cohorts. Are there any additional biomonitoring studies that should be included in the Agency's comparison?*

The Panel concurs with EPA that the human studies discussed in this section are currently the best available, primarily because they are carefully designed and well executed. The Panel recommends the following additional biomonitoring studies listed in the ordered that they should be considered: 1) NHANES 1999-2004, 2) Barr *et al*. 2010, 3) Bradman *et al*. 2005 (because the families studied are likely to continue to see significant exposure which should be validated by the next round of NHANES data), 4) the Children's Pesticide Exposure Study (CPES) by Lu et. al. 2008 and Children's Post-Pesticide Application Exposure Study (CPPAES) studies, and 5) studies that are either currently in process or completed and will be published soon. This last group includes: 1) The Children's Pesticide Exposure Study that focuses on dietary intake of children and related pesticide exposures being conducted at the Harvard School of Public Health. 2) The

Children's Exposure to Environmental Pesticides study which is evaluating the utility of biomarkers of pesticide exposures, *e.g.*, DAPs and pesticide-specific markers of OP and pyrethroid exposures, and environmental levels measured in soil, house dust, and food being conducted by Emory University, and 3) the SAWASDEE cohort study, that is examining pesticide biomarker concentrations in pregnant mothers and similar markers in their newborn children, run by Emory University and Chiang Mai University in Northern Thailand. The Panel advises that in comparing the results among these different studies, it is important to verify that analytical results from the studies are directly analogous, especially in the methods used to control for the effect of small day-to-day variations in a laboratory's AChE results when trying to quantify small changes in an exposed population.

Question 5.2

In Section 5.0 of the draft issue paper, the Agency summarized the 2008 preliminary findings on the association between urinary TCPy levels and AChE/BuChE inhibition and discussed two recent studies involving manufacturing workers in the US and Egypt. *Please comment on the scientific quality of these studies and their findings. Please include a discussion of their strengths and limitations. Please comment on the strengths and limitations of the evidence from this research to show an association between TCPy and AChE/BuChE inhibition at exposure levels experienced by occupational populations.*

The Panel notes that both studies were in general well designed and implemented. They both have adequate power to demonstrate an association between TCPy and AChE or/and BuChE inhition at exposure levels experienced by occupational populations. In addition the cholinesterase data from the studies verify the ability of the PBPK model to predict that once chlorpyrifos is absorbed, it interacts first with BuChE and only starts to inhibit RBC AChE and AChE in the central nervous system after BuChE is more than 50% inhibited.

The Panel points out several primary weaknesses in both studies as it relates to using them within the weight of evidence paradigm. These include high levels of TCPy in pre-exposure samples indicative of prior exposure to chyorpyrifos or environmental degradates either from food or accumulated residues in the workplace. Neither study reported analyses that adjusted for levels in control groups.

The Panel recommends that the Agency separate the scenarios for occupational exposures, as reported in these two studies, from exposures to residential sources. The extrapolation of these data to the population as a whole is subject to criticism. The subjects in these two studies are adults and issues such as the "healthy worker effect" and the notion that low-level exposure and high-level exposures are likely to be detoxified by differing mechanisms make extrapolation difficult. Additionally, the Panel suggested that studies of agricultural workers and their families could provide a better avenue of investigation that compares "occupational-levels" exposure with other members of their families likely to see slightly "elevated" but lower levels of exposure, and to study the potential impact on the offspring in such cohorts either *in utero* or otherwise. The Panel

recommends that in the future, the Agency consider the quality of ChE measurements before pursuing further uses of these data in exposure and risk assessments.

Question 5.3

Several approaches ranging from qualitative to the most sophisticated PBPK/PD modeling approach were introduced as potential options for analyzing the chlorpyrifos biomonitoring data. *Please comment on the strengths and limitations of these approaches. In addition, please suggest, if appropriate, alternative approaches or analyses not identified by the Agency.*

There are a rising number of data-informed options for interpreting biomonitoring data. Choosing the adequate option relies on the extent of the data available, on the toxicokinetics of the relevant population subset, on the mode of action, and on the integration of these data. Integrating these data through a verified PBPK model has the potential to be the most informative approach while also being the most data intensive.

The Panel advises EPA that at the very least the chlorpyrifos biomonitoring data should be utilized as a means of "ground truthing" total external exposures under a variety of use conditions. Considering the availability of chlorpyrifos biomonitoring data on the general population, and as a basis to support its maximal consideration of public health, the Panel also recommends that the Agency seriously consider developing a "biomonitoring equivalent" at the same time the reference dose for chlorpyrifos is derived. The biomonitoring equivalent is defined as a calculated level of a biomarker associated with exposures consistent with health protective guidance values for the general population. The Panel also recommends that the Agency utilize a verified PBPK model which will provide a robust opportunity to integrate the considerable available data on external and internal exposures (*i.e.*, biomonitoring) to chlorpyrifos at different life stages under different conditions of exposure. With respect to a specific PB/PK model for the Agency to consider, the Panel recommends a sophisticated model such as the SimCYP pediatric model (SimCYP Company, Sheffield, UK) for children that is currently available.

Question 5.4

Characterization of chlorpyrifos exposure experienced by women in the Columbia cohort, particularly during the pre-cancellation period, remains an important uncertainty in using these data in quantitative risk assessment. Exposure levels in the range measured in the cord blood data from the epidemiology studies (pg/g plasma) are probably low enough that is unlikely that the cohort mothers were experiencing AChE inhibition at the time of delivery; however, the biomonitoring data were taken after birth and not necessarily associated in time with an application of chlorpyrifos. As such, the actual level of such exposure particularly during any critical window(s) of susceptibility is not known, and a better understanding of the range of possible exposures and the degree to which they may or may not have elicited inhibition of AChE, remains a key scientific question. *In light of Panel discussions of Questions 4.3 and 5.3, please suggest approaches and/or analyses which would inform the understanding of the degree to which exposure levels experienced by the Columbia cohort participants may or may not have been below doses*

*which result in 10% inhibition of AChE in the most sensitive lifestage. Please discuss the strengths and uncertainties associated with such analyses. Please include in your discussions approaches involving chlorpyrifos and its metabolites and also chlorpyrifos plus other AChE-inhibiting pesticides (propoxur, diazinon) which the cohort participants were exposed too.*

The Panel notes that it is important to realize that the short half-life of chlorpyrifos and its metabolites in the body calls into question any "spot data" that might be used. Large cross-sectional studies may capture some exposure but they do not put these exposures into context. Longitudinal investigations with frequent samplings are more likely to provide data that are more useful. Thus, the Panel recommends that a longitudinal study with measurement throughout the pregnancy (rather than a few samples in the last trimester) would fill many of the data gaps that currently exist for this group. Such a study is needed given the potential for neurodevelopmental effects on the fetus as well as the metabolic differences in pregnant women versus the workers from the 1984 study.

As discussed in the response to the previous question, the Panel again recommends that a more sophisticated PBPK model may provide better data particularly if the model is pertinent to the population being studied, *i.e.* pregnant women and small children.

Studies discussed in Question 5.1 provide data on the concentration of chlorpyrifos in various media (*i.e*. house dust, air and water) while market basket data exists on the concentration of chlorpyrifos on food. These data provide the main tools for developing an effective exposure assessment and a subsequent reconstruction of potential dose. Dose reconstruction can be used to evaluate the efficacy of the PBPK model since its prediction of excretion rates can potentially be validated with an accurate estimate of dose. This assessment of the PBPK model through reconstructed dose may bridge some of the data gaps in assessing risk by validating the PBPK model.

The Panel discusses the issue of mixtures of chlorpyrifos + Diazinon /chlorpyrifos + Propoxur or chlorpyrifos/Propoxur/Diazinon. The Panel recommends that the Agency address the following questions: "Do mixture components affect each other's half lives, distributions and clearance through metabolic competition?" "Are the net AChE effects of mixtures additive or multiplicative?, and "Do they share mechanistic pathways?"

Lastly, the Panel expresses concern over the Agency's focus on a 10% AChE activity reduction. They point out that to their knowledge there is no proposed mechanism whereby a 10% AChE activity reduction in pregnant women would be responsible for a cognitive defect or developmental delay in their offspring.

***Charge 6: Characterizing the range of potential risks.***

The 2009 NRC report, *Science and Decisions*, focused on improving the *technical analysis* through the development and use of scientific knowledge and information to promote more accurate characterizations of risk, and thus improving the *utility* of risk assessment for risk-management decisions. The NRC report also pointed out that regulatory risk assessment does not routinely approach public health and environmental problems by arraying a wide range of options for dealing with them. *In the case of chlorpyrifos, in light of the discussions of Questions 1-5, please provide guidance for assessing and presenting the range of plausible responses at given doses, and the effect of the overall uncertainty and variability around that range.*

With regard to characterizing the probable response at given doses, the Panel recommends that the Agency use the dose-response data to establish multiple points of departure. For instance, it would be informative for risk management purposes to fully characterize the nature of the risk above the reference dose. The Panel also recommends that the Agency maximizes its use of available data on dose response from the epidemiology studies as a basis to at least "bound" reference doses developed on the basis of points of departure from animal data. As advised by the Panel, options for dose-response analysis for acute effects should be considered independent from those based on long term exposures, *i.e.*, measures representing acute adverse neurological outcomes (ChE inhibition) commonly associated with occupational exposure versus those potentially related to lower level long term exposure in the general population, such as neurobehavioral disorders.

The Panel suggests that the Agency focus on the data of chlorpyrifos in the cord blood as a means for creating the point of departure for chronic exposures to chlorpyrifos based on the PBPK/PD model.

Another consideration that the Panel deems important is the degree to which epidemiological data on neurotoxicity is consistently used within EPA to establish points of departure for chemicals with the potential for neurotoxic outcomes (*e.g.*, mercury and lead). The Panel asks: "How is the evidence from epidemiological studies weighted in the assessment for these compounds, and how does this compare with what is proposed for chlorpyrifos?"

## DETAILED PANEL DELIBERATIONS AND RESPONSE TO CHARGE

It is well established that acetylcholinesterase (AChE) inhibition is the primary mode of action/adverse outcome pathway for organophosphorus chemicals (OPs) such as chlorpyrifos. In June 2011, consistent with the recommendations from the Scientific Advisory Panel (SAP) in 2008, the Agency performed a risk assessment utilizing AChE inhibition data in laboratory animals for deriving points of departure and for dose-response analysis as the Agency believes these data remain the most robust and most sensitive information available for regulatory risk assessment. However, newer lines of research on chlorpyrifos such as epidemiological studies in mothers and children, have posed the issue of whether AChE inhibition is the most sensitive health outcome, leading to questions about the chlorpyrifos risk assessment.

In order to determine the degree to which these recent studies are appropriate for incorporation into risk assessment (qualitatively and/or quantitatively), the Agency is taking a stepwise, objective and transparent approach to evaluate and interpret all the lines of scientific information related to the potential for adverse neurodevelopmental effects in infants and children as a result of their prenatal exposure to chlorpyrifos, as well as to characterize thoroughly the strengths and uncertainties associated with these studies. The issue paper entitled "Scientific Issues Concerning Health Effects of Chlorpyrifos" extends the Agency's September 2008 review of the available experimental toxicology and observational epidemiology data. This 2012 review incorporates experimental data available since the time of the last review relating to AChE inhibition and both cholinergic and non-cholinergic adverse outcomes, including neurodevelopmental studies on behavior and cognition effects. Similarly, the Agency also performed a more in-depth analysis of the epidemiologic studies from three major children's health cohort studies in the U.S., plausible hypotheses on modes of action/adverse outcome pathways (MOA/AOP) leading to neurodevelopmental outcomes, along with biomonitoring and physiologically-based pharmacokinetic/pharmacodynamic (PBPK/PD) modeling than was conducted in 2008. Overall, the Agency has updated and extended its evaluation of multiple lines of evidence informing the chlorpyrifos risk assessment.

As discussed the 2012 issue paper, two of the key scientific questions are: 1) the degree to which scientific data suggest that chlorpyrifos causes long-term neurodevelopmental effects from fetal or early life exposure and 2) the degree to which adverse effects can be attributed to doses lower than those which elicit 10% inhibition of AChE, *i.e.*, the dose levels previously used for regulatory decision making. The evaluation of these scientific questions requires integration numerous types of data, and consideration of the nature and degree of the uncertainties surrounding the data, including the extent to which alternative interpretations may be supported. This step is vital to robust risk characterization and uncertainty analysis. The 2011 preliminary risk assessment noted that a full weight of the evidence analysis that explicitly considers uncertainty and implications of experimental and epidemiologic lines of evidence using factors such as biological plausibility, strength, consistency, and dose-response and temporal concordance, will be conducted in the

future. A full weight of the evidence and full uncertainty characterization has not yet been conducted; the 2012 SAP is an important step toward this effort.

***Question 1.0 Mode of action/adverse outcome pathway: Acetylcholinesterase (AChE) inhibition***

Question 1.0

It is well established that AChE inhibition is the primary mode of action/adverse outcome pathway for OPs, like chlorpyrifos. Because AChE inhibition is the initiating event for this mode of action/adverse outcome pathway, using AChE inhibition as a regulatory endpoint is protective of downstream cholinergic effects. Moreover, historically, given the sensitivity of AChE inhibition data for OPs, these data have been considered to be protective of other potential toxicities and/or modes of action for OPs. In 2008, the Agency performed a comprehensive review of the available AChE data from multiple lifestages. This review has been supplemented with the newest studies. Consistent with the recommendations from the 2008 SAP, the Agency believes that AChE data remain the most robust dose-response data for deriving points of departure in *in vivo* experimental toxicology studies with laboratory animals. *Please comment on the Agency's preliminary conclusion that AChE data remain the most robust source of data for deriving points of departure for chlorpyrifos. Please include a discussion of the strengths and uncertainties of this preliminary conclusion.*

***Response***

The Panel agreed with the Agency's conclusion that the AChE data remain the most robust source of data for deriving points of departure for chlorpyrifos. This is based on the observation that all studies reporting neurobehavioral changes following *in vivo* prenatal or postnatal exposures to chlorpyrifos have been accompanied by AChE inhibition when measured at an appropriate time following administration of chlorpyrifos. Moreover many studies reporting persistent neurobehavioral changes used a potentially confounding vehicle (*e.g.*, DMSO). Most importantly, the experimental design for essentially all experimental studies evaluating neurodevelopmental effects of chlorpyrifos do not allow for the effective determination of a point of departure. There are candidates that may replace AChE as a more sensitive indicator but, at this juncture, these have not been fully validated and their alteration has not been determined to result in a well-defined, measurable neurotoxic outcome.

As in 2008, the Panel recommended that the Agency continue to use AChE data at the most sensitive lifestages for dose-response analysis and deriving points of departure. When looking at obstetric outcomes and pediatric exposures, life-stage levels of red blood cells (RBC) AChE activity in humans, which has been reported as significantly lower in fetal cord blood than in adults (de Peyster *et al.* ,1994), needs to be taken into account to eliminate potential uncertainties.

The Panel concurred with the 2008 Panel in expressed caution on the use of DMSO as a vehicle because of its intrinsic toxicity and potential influence on absorption. Again, uncertainty was expressed about potential interactions between DMSO and low doses of

chlorpyrifos and the effect of this interaction on the developing organism. In addition to the three papers cited by the 2008 SAP (FIFRA Scientific Advisory Panel, 2008b), more recent evidence is available to support the potential toxicity of DMSO. Hanslick *et al.* (2009) reported that following acute intraperitoneal injection of DMSO into 7 day-old mice, there was a significant increase in the number of apoptotic neurons at dosages as low as 0.3 ml/kg. An increased number of apoptotic neurons was also observed at 1 ml/kg which is the most frequent volume of DMSO administered in the cited studies using DMSO as a vehicle. Recent reports from the zebrafish literature suggest that DMSO has the capacity to directly induce neurobehavioral effects. Exposure to 0.05% DMSO induces anxiolytic behavior in adult zebrafish (Sackerman *et al.*, 2010) and exposure to 0.01% DMSO alters locomotor activity in larval zebrafish exposed embryonically (Chen *et al.*, 2011). Also, based on earlier studies observing that DMSO induces a stress protein response in zebrafish embryos (Hallare *et al.*, 2004; 2006), Turner *et al.*, (2012) reported that levels of DMSO as low as 25 µl/L (0.0025%) were sufficient to induce gene expression changes in embryonic zebrafish. While altered gene expression does not indicate a toxic response, it suggests disruption of homeostasis by low levels of this solvent. While the experimental studies reviewed in the White Paper all had controls with DMSO only, there is no way to rule out the potential for an interaction between DMSO and the OP. For example, Fossum *et al.* (2008) reported that 2% DMSO had no effect when microinjected into the periaquaductal gray region of rat brain, but it enhanced the potency of morphine when co-administered. In this case, if morphine was dissolved in 2% DMSO and the controls received DMSO only, the interpretation of the findings are confounded. It should be noted that the concentration of 100% DMSO is approximately 14 M. Because of the potential biological/cellular changes noted above, the lack of evaluation of potential interactions between DMSO and chlorpyrifos, and the well-known effects of DMSO on membrane permeability (Gurtovenko and Anwar, 2007), caution should be exercised in the use of data for quantitative risk assessment from *in vivo* (or *in vitro*) studies using DMSO as a solvent.

***Question 2.0 Mode(s) of action/adverse outcome pathway(s): Plausible pathways leading to potential neurodevelopmental outcomes***

Question 2.1

As discussed in Section 3.2.1, although there are numerous mechanistic studies in the scientific literature, the research on different hypotheses does not provide sufficient data to establish causal linkages among different levels of biological organization to show how effects lead to adversity. As such, a mode of action or adverse outcome pathway leading to effects on the developing brain cannot be established at this time. Moreover, although multiple biologically plausible hypotheses are being pursued by researchers, based on the current state of the science, no one pathway has sufficient data to be considered more credible than the others. *Please comment on the Agency's preliminary conclusion that although there are multiple biologically plausible hypotheses being evaluated by research scientists, the mechanistic experimental toxicology data do not yet support a coherent set of key events in a mode of action/adverse outcome pathway.*

***Response***
The Panel acknowledged the efforts EPA has taken to review all of the relevant data addressing the various cellular and mechanistic based studies on chlorpyrifos and relevant associated neurobiologically-based studies. Research scientists are examining multiple biologically plausible hypotheses regarding cellular mechanisms of chlorpyrifos neurotoxicity. Over the past approximately 15 years a number of studies have evaluated changes in neurite outgrowth, axonal transport, dendritic growth, and other cellular processes following chlorpyrifos or chlorpyrifos oxon exposure that could potentially disrupt the development of the nervous system. In no case, however, is there a defined, coherent set of events from alteration of any of these cellular functions to disrupted development of the nervous system sufficient to explain a variety of neurobehavioral changes. There is also limited evidence that these current research efforts are directed in such a manner to link the *in vitro* findings to a structural or functional change in the animal. The Panel agreed with the Agency that, based on the current state of the science, no single pathway has sufficient data to be considered more credible than the others with respect to a causal link between chlorpyrifos exposure and toxicological outcome.

As defined, the progression of events from molecular initiation to adverse outcome requires a logical sequence of changes in the mode of action/adverse outcome pathway. The Panel raised the issue that the example of domoic acid as a linear connection is likely to be a unique case that can provide components at each level of the pathway but may also generate a non-linear pattern. A linear pathway from mode of action to adverse outcome appears rare. Thus, the Panel agreed that while laudable, expectations of the existence of such a pathway may be artificially elevated and potentially unrealistic for risk assessment.

<u>Question 2.2</u>
Although a mode of action/adverse outcome pathway has not been established, qualitatively, the growing body of mechanistic studies does demonstrate that chlorpyrifos and/or its oxon are biologically active on a number of processes that affect the developing brain. Some mechanistic studies provide evidence of possible effects which are similarly sensitive or more sensitive than AChE inhibition (*e.g.*, neurite outgrowth, binding to muscarinic receptors, axonal transport; serotonergic nervous system development). Some of these comparisons must be considered with caution since the amount of change in the *in vitro* systems required to elicit an adverse effect *in vivo* is unknown. Moreover, extrapolation from *in vitro* perturbations to *in vivo* effects has not been established, which introduces additional uncertainties. *Given the doses/concentrations evaluated in the in vitro and in vivo mechanism studies, please comment on the degree to which these studies suggest that endpoints relevant to evaluating potential neurodevelopmental outcomes may or may not be more sensitive than AChE inhibition. Please include in your comments a discussion of the strengths and uncertainties. Please also include in your comments a discussion of the scientific understanding of dose-response relationships for biological perturbations and the magnitude, frequency and/or duration of such perturbations that can lead to adverse effects at higher levels of biological organization to support characterization of the likelihood of adverse outcomes in a human health risk assessment (as articulated in NRC 2007).*

***Response***
The Panel agreed with EPA that caution must be exercised in interpreting the *in-vivo* relevance of the changes observed across the various *in vitro* studies. There are a number of cellular processes such as, neurotransmitter receptor activation, and others that are currently under study; however, the studies lack the necessary data on the relevance of these changes observed *in vitro* and effects occurring *in vivo*. Much of this work is speculative because the underlying process being studied is assumed a process critical to brain development. *In vitro* work is best suited for testing hypotheses that can be further explored *in vivo*, but given their reductionist approach, they lack the experimental power and demonstrated predictive validity of an *in vivo* effect at this stage to be of scientific value for risk assessment considerations. The *in vitro* models that have been utilized to address the effects of chlorpyrifos and oxon such as neurite outgrowth, M2 acetylcholine receptor binding, mitochondrial morphology and axonal transport, as well as oxidative stress, may potentially provide information on non-cholinesterase related mechanisms.

The Panel discussed the limitations of the cellular *in vitro* systems to translate to *in vivo* systems as well as to identify a human health risk. They noted that the inherent complexity of the nervous system cannot be replicated in a cellular *in vitro* system. The majority of these studies were conducted to explore the potential for chemical exposure to induce a change in cell physiology and to further examine underlying mechanisms. Thus, while the reductionist approach provides information on possible mechanisms of action for chlorpyrifos, it does not translate to *in vivo* effects. The extrapolation of data from *in vitro* to *in vivo* is filled with uncertainty factors. Adding to this uncertainty is the possibility that although chlorpyrifos has a well-established molecular target (AChE) for cholinergic toxicity, it may not be the only toxicologically relevant target. In addition, due to the inherent complexity of the nervous system, which contains multiple regions and systems that are connected and interact with one another, effects on the molecular target could induce downstream effects that can result in a neurotoxic response that may not be directly attributable to the molecular target. This complexity and these connections cannot be replicated *in vitro*.

Issues of concern raised by the Panel were not specific to chlorpyrifos, but rather these concerns were directed toward the use of *in vitro* systems in general. Such concerns entailed the isolated nature of the cells in culture and the inability to address critical regulatory components of the *in vivo* environmental niche. For instance, the Panel noted that the *in vitro* model system can influence the effects observed in isolated cells, but the responses may be changed or absent when these same cells are co-exposed with other cells normally within their *in vivo* environment. The Panel also raised questions about how the isolated nature of the various model systems could deal with altered homeostasis or dose response differences that may occur as a function of differential recruitment of processes *in vivo*. In addition, the actual amount of compound or AChE inhibition at the target site is one critical factor toward determining if a toxicological event observed *in vitro* can occur *in vivo*. The Panel mentioned that that until such *in vivo* translation can be established, it is difficult to determine what level, magnitude, or duration of change is required within each model system to be indicative of a change that may occur *in vivo*.

The Panel also discussed the likelihood that using such models may not provide a linear dose-response relationship for chlorpyrifos or for AChE inhibition. Thus, providing a direct assessment of the relative sensitivity to AChE inhibition is difficult. Changes in neurite outgrowth, dendritic spine development, axonal transport and other cellular responses that have been seen with chlorpyrifos or chlorpyrifos oxon exposure *in vitro* have been reported to occur at levels of exposure below those necessary to inhibit AChE. These comparisons must be made with caution. For example, adding chlorpyrifos oxon directly to hippocampal neurons *in vitro* in tissue culture medium is unrealistic with regards to how the oxon would reach a neuron *in vivo*. Many detoxifying/binding proteins are not included in *in vitro* conditions, potentially modifying the interaction of the chemical with the cell. Similarly, if acetylcholinesterase in a disrupted tissue (*e.g.*, a homogenate) such as liver and is inhibited *in vitro* by chlorpyrifos oxon, its sensitivity is much higher than if that same enzyme is immunoprecipitated and then inhibited by chlorpyrifos oxon *in vitro* under similar conditions. Such effects of tissue components, in the relative potency of chlorpyrifos or chlorpyrifos oxon, make extrapolation of *in vitro* effects to *in vivo* settings difficult. In general, changes in neurodevelopmental endpoints require relatively higher exposures in *in vivo* models. Reviews of mechanistic/cellular studies with neurodevelopmental outcomes following chlorpyrifos exposure suggest that such responses may occur at dose levels at, near, or above those necessary to induce AChE inhibition. The development of a physiologically-based pharmacokinetic (PBPK) model that can estimate the *in vivo* dosage required to reach toxicant concentration at the target site that is similar to the *in vitro* concentration at which the effects were observed would be especially valuable. Such a model would assist in determining the plausibility that the effect observed *in vitro* also occurs during an *in vivo* exposure.

The Panel was in agreement that further mining of the published literature may provide significant information on how one might utilize the available data obtained for chlorpyrifos to generate plausible hypotheses for future evaluation. As an example, the analytical studies of the Lockridge group indicating that chlorpyrifos (chlorpyrifos) oxon can covalently modify key cytoskeletal proteins such as tubulin and motor proteins like kinesin provide information that can contribute to the interpretation of findings of alterations in neurite outgrowth and axonal transport, respectively. Such an integrated effort may allow for the design of specific targeted studies to test the hypothesis *in vivo* as an effort to obtain predictive validity (Jiang et al., 2010; Grigoryan et al., 2009).

The Panel considered items that might be added to the review. One topic is the effect of chlorpyrifos on neurotrophins (growth factors). Pope *et al.* (1995) provided the first evidence that OPs might alter the activity of growth factor-like (neurotrophic) molecules. Data suggesting that growth factors could be altered within the brain tissue, *in vivo*, has recently been provided in neonatal rats across low dose levels of chlorpyrifos. Alterations were observed in mRNA levels for specific members of the fibroblast growth factor (FGF) superfamily of neurotrophic factors (Slotkin *et al*., 2007). Further, early postnatal exposure to chlorpyrifos has also been associated with decreases in nerve growth factor (NGF) in the rat forebrain (Betancourt and Carr, 2004). These effects are not limited to the immature rodent in that adult exposure to chlorpyrifos can result in protracted alterations in NGF-related signaling proteins (*e.g.*, the high affinity nerve

growth factor receptor TrkA and its activated form, phospho-TrkA in the prefrontal cortex) (Terry *et al.*, 2007). Also, as a note, in the Middlemore-Risher *et al*. (2011) study, alterations in mitochondrial morphology and decreases in axonal transport were observed in primary cortical neurons exposed to chlorpyrifos and chlorpyrifos oxon at concentrations including ones that did not inhibit acetylcholinesterase.

The Panel raised the issue that oxon and protein adducts likely serve as a potentially important pathway for cellular/protein damage. The oxon has an incredibly rapid half life and despite a relatively high affinity for ChE, one would expect that they would occasionally bind non-ChE cellular components 3,5,6-trichloro-2-pyridinol (TCPy) and diethylthiophosphate (DETP) are conjugated by -O-sulfotransferases or gluruconosyltransferases; and as presented in public comments from Dow Chemicals, both sulfonylates and glucuronides are "equally prevalent" but relative affinities (Km) appear to be unknown. In translating from animal to human, it was considered of importance by the Panel that Uridine 5'-diphospho-glucuronosyltransferase enzymes (UGTs) develops late in children. In this case the sulfotransferase (SULTS) are usually present during gestation, despite the fact that the AChE adducting oxon would not be conjugated. One must consider that if glucuronidation is the rate-limiting pathway in children, then other metabolites may accumulate to toxic levels due to ontogenetic inadequacy of UGTs. This could result in a potential for error in biomarker analysis and generate errors in dosimetry estimations.

The Panel discussed the types of dose-response relationships that may be observed and allowed for non-linear patterns or no clear dose-response association to be observed. The Panel cautioned that when examining the dose-response relationship, one should not uncouple PD analyses from PK models too far given that PK differences can affect active site concentrations and hence, PD effects. To this end PK can significantly affect the magnitude and duration of an effect. It was noted that the p450s and PON1 were well integrated into the PB/PK/PD models and this was considered a major strength. There were, however a number of weaknesses discussed including the fact that there is a substantial lack of knowledge about the high capacity of phase II conjugation for chlorpyrifos in humans. For example, there are species differences in Phase II where UGTs and SULTS may be non-orthologous between humans and rodents; hence any extrapolation of animal data to humans must take this into consideration. The Panel noted that this was highlighted by the public presentation from Dr. Hattis who noted that the human data he subsequently used and modeled demonstrated lower clearance than seen in rat data. An issue was raised by the Panel that pregnancy can be considered as a specific state and that information is needed relative to how PK differs in a pregnancy scenario relative to exposure and toxicity. In addition to maternal influences such as metabolism by the liver, the placenta was also raised as a unique component requiring attention in such PB/PK/PD models for gestational exposure.

The Panel raised a concern that equivalency developmental stages between ages of rodents to human are not well defined with regards to cell type compositions, brain region, cellular architecture, and physiological or biochemical process. This is not a problem that the Agency needs to address but rather the Panel emphasized that specific

developmental periods in which perturbation may occur appear to be ill defined and may not translate between rodent and human species. The numerous *in vitro* mechanistic studies suggest that chlorpyrifos can alter numerous biological processes in normal brain development. However, these data do not permit translation to the *in vivo* situation nor do they provide a quantitative dose-response relationship that can be compared to AChE inhibition.

***Question 3.0 Neurodevelopmental data from laboratory animals***

Question 3.1

As discussed in Section 3.2.2, the experimental toxicology data in laboratory rodents show neurobehavioral effects following developmental exposure with changes in a number of neurological domains. In 2008, the SAP agreed to this preliminary conclusion, and the nine additional studies available since 2008 add further support. *Please comment on the degree to which these studies show changes in a number of neurological domains and support the qualitative conclusion that chlorpyrifos exposure during gestation and/or early post-natal period may result in long-term adverse effects on the developing nervous system. What evidence does and does not support this conclusion? Please also include in your comments a discussion of the strengths and uncertainties. Please also include in your comments a discussion of the scientific understanding of dose-response relationships for biological perturbations and the magnitude, frequency and/or duration of such perturbations that are can lead to adverse effects at higher levels of biological organization to support characterization of the likelihood of adverse outcomes in a human health risk assessment (as articulated in NRC 2007).*

***Response***

In order to address the first part of this charge question, the Panel critically reviewed these toxicology data in laboratory rodents. This review included a total of 21 developmental neurobehavioral effects studies which also entailed the nine studies published since the 2008 SAP review (These studies are identified in Appendix 3 of the Agency's Draft Issue Paper: Scientific Issues Concerning Health Effects of Chlorpyrifos.). Based upon their review, the Panel agreed with the 2008 SAP conclusions that developmental neurobehavioral experiments show adverse effects of chlorpyrifos exposure. However, the Panel cautioned that the existing neurobehavioral studies are limited and a number are under-powered and prone to Type I error (meaning the null hypotheses may have been falsely rejected) and therefore should be discounted in determining the weight of evidence for or against neurobehavioral effects from developmental exposure to chlorpyrifos. An additional concern was raised by the Panel in the inclusion of tests that have not been validated as to neurotoxicological significance. Such assessments included anxiety tests, depression tests, or social interactions. The Panel concluded that, in the current state, such outcomes should be regarded as exploratory, and hypothesis-generating, rather than evidence of toxicity. The Panel considered that the lack of observable effects below dosages equal to or exceeding 1.0 mg/kg/day in some studies could be due to the possibility that the effects at lower dosages might not be the same as those observed at higher dosages. In addition, low dose exposure may lack a sufficient level of response to elicit a physiological compensation

response that would occur with higher dose levels. In this scenario, the toxic effects of the higher and lower dosages may manifest themselves in different ways. If the response is through the same physiological target, the expected dose-response curve may be altered (such as a U-shaped or inverted U-shaped curve). In contrast, the response of the lower dosages may be quite different from that of the higher dosages and require a different behavioral paradigm.

Many of the 21 studies reviewed included 2 or 3 dose levels of chlorpyrifos (ranging from 0.2 in one case up to as high as 10 mg/kg in another but more commonly from 1-7 mg/kg), but dose-effect outcomes were, more often than not, not observed. Some of the most consistent effects are from the Slotkin-Levin experiments (see below) where radial arm maze deficits and several other effects have been replicated but the studies are seldom designed with three doses levels and even when two dose levels were included the findings were not dose-dependent in most cases. The Panel agreed with the Agency that the lack of specificity of direction of the neurobehavioral findings is problematic. A statistically significant change in isolated markers of certain behaviors may not be supported by other studies under similar dosing paradigms thereby raising concern regarding the biological significance of the observed change. The Panel questioned the Agency's interpretation that a change in either direction or a specific behavior is necessarily indicative of an adverse effect. Rather, such discrepancies may suggest methodological error, problems in study execution, or a predilection toward searching data for positive rather than negative findings. Dose-response and attention to methodological issues should play a role in evaluating the weight of evidence within and across the different studies. The Panel agreed that the overall evidence across these studies is persuasive, indicating that there are enduring effects from chlorpyrifos exposure at 1.0 mg/kg or above on the CNS. Future neurodevelopmental studies need to focus on levels below 1.0 mg/kg/day and to expand the studies to identify the correct testing paradigm to detect these effects and possibly identify a neural substrate. The Panel also considered the possibility that additional negative data exists at lower dose levels but is not available in the published literature. The Panel suggested that cross-laboratory or collaborative studies could provide systematic comparison of the effects of chlorpyrifos on neurodevelopmental domains using unified exposure periods, dosing, age of testing, and methods, combined with urinary analysis of chlorpyrifos' metabolites, and accurate assessments of AChE inhibition.

In evaluating the inconsistencies among these studies, the Panel suggested that the Agency should consider factors such an the distinct ontogeny of the various brain regions, cellular components, and neurotransmitter systems in the fetal/gestational exposures and that the structural and functional maturation of each system is unique and thus may be at different stages during the age of exposure. In addition, the redundancy and compensatory capability of each system should be considered as the level of insult may be required to reach a substantial level before it manifests as a neurobehavioral change.

The Panel's detailed review of each these developmental neurobehavioral effects studies is provide in Appendix A of this report. However, the following provides a collective summary of the Panel's observations and conclusions regarding the studies.

Among the 21 reviewed articles (which include more than 21 experiments), many effects are reported at chlorpyrifos doses ranging from at 1.0 to 7.0 mg/kg (and in one study 10 mg/kg). However, these are dose levels known to significantly inhibit cholinesterase in RBC, therefore, based on these 21 studies, cholinesterase inhibition is an adequate threshold as no credible evidence of neurobehavioral effects below 1.0 mg/kg were found.

Three studies tested doses <1.0 mg/kg chlorpyrifos. Two studies used 0.3 mg/kg and one used 0.2 mg/kg (Jett *et al*., 2001;Braquenier *et al*., 2010;Maurissen *et al*., 2000). Of these, two found no effects at these lower doses (Maurissen *et al.*, 2000;Braquenier *et al.*, 2010). Only one study found effects at 0.3 mg/kg (Jett *et al*., 2001), however, this study contains serious methodological flaws which are of sufficient magnitude to cast serious doubt on the credibility of the findings. While the data from Jett *et al*. (2001) raise the possibility of neurobehavioral effects at 0.3 mg/kg/d, these data require replication in a study that is properly designed, adequately powered, and appropriately analyzed. Until such time as new data at such lower doses become available, it is concluded that no dose <1.0 mg/kg in any neurodevelopmental behavioral studies shows evidence of adverse effects (or of any effects, even including those outcome measures of indeterminate/unknown toxicological significance).

In addition, effects of chlorpyrifos at 1.0 mg/kg are difficult to interpret because of methodological limitations, inconsistencies, and variation in study design, sometimes lack of control for litter effects, oversampling issues, behavioral methods used, and lack of dose-response findings.

At doses exceeding 1.0 mg/kg, the data show somewhat more consistency, but even here, dose-response experiments are the exception. A 5.0 mg/kg of chlorpyrifos, reduced body weight is sometimes seen, and at doses above 5.0 mg/kg increased mortality may occur along with other evidence of toxicity. Given this, it is a significant gap in the literature that more dose-response studies are not available in the range downward toward 0.2 mg/kg and extending up to and including doses previously tested of 1.0-2.0 mg/kg in order to determine what, if any, dose-effect curve occurs in this range for neurobehavioral effects.

It appears that prenatal and prenatal-neonatal exposures are more sensitive than neonatal exposure alone on neurobehavioral outcomes. This implies that prenatal exposure may be the exposure period contributing to this observation, but unfortunately, most of the pre- and neonatal studies are not entirely informative because the neonatal exposure was to the dam rather than directly to the progeny. This makes it unclear what the exposure to the offspring actually was or whether it was at similar levels to those reaching the embryo and fetus. More studies, especially dose-response studies, in the lower dose ranges with exposure from implantation to the end of major neurogenesis (approximately P20) are

needed, again with doses below 1.0 mg/kg and with concomitant measurement of maternal, fetal, and neonatal cholinesterase activity.

Many of the existing studies expose for only a narrow interval during gestation or the neonatal period. Prenatal exposures should be from E6-20 to 21 for rats, and E6-18 or 19 in mice in order to span most of early brain development (equivalent to human first and part of second trimester). And for neonatal treatment, exposures should be from shortly after birth to approximately P20 (equivalent to the latter half of second and all of third trimester equivalent brain development comparable to that for humans). If the critical period or most sensitive period is within this range, then such comprehensive exposure should cover the entire span of CNS development that represents the species being modeled, *i.e.*, human beings.

In the prenatal studies, the use of timed-pregnant females shipped from breeders is problematic for behavioral studies because maternal stress, even if regarded as equivalent across dams assigned to the treated and control groups, introduces a variable that has the potential to interact with the independent variable. If maternal stress were to interact with chlorpyrifos, it would confound the outcome and make a result difficult to interpret (which is exactly what is found in many of the reviewed studies). Since no one has tested for this, it is currently impossible to rule it out.

Many studies use diurnal and some nocturnal testing. If additional dose-response studies are undertaken, this factor should be held constant so that results can be better compared.

Question 3.2

The dose-response data in the *in vivo* experimental neurodevelopmental toxicity studies are not amenable to empirical dose-response modeling as many studies use only one or two doses, and in some cases the lower dose, but not higher dose level, produced significant effects. Many studies report effects at a dose of 1 mg/kg/d-- a dose that produces some amount of brain ChE inhibition when given directly to the pups post-natally, but may or may not alter fetal brain ChE activity when given to the dams gestationally. One study (Braquenier *et al.*, 2010) using lower doses, administered to the dam on GD15-LD14, reported a NOEL of 0.2 mg/kg/d. Comparing the NOEL of 0.2 mg/kg/d to a repeated dosing AChE inhibition $BMDL_{10}$ of 0.03 mg/kg/d suggests that AChE inhibition is a sensitive and protective endpoint.

*a. Please comment on the scientific quality and robustness of the animal neurodevelopmental toxicity studies.*

***Response***

The quality of the studies in this category varies, but there are some of high quality. Overall, the studies by Slotkin (Dam *et al*., 2000; Icenogle *et al*., 2004) and Levin (Levin *et al*. 2001 and 2002), those by Carr *et al.* (Carr *et al.* 2001 and Johnson *et al.*, 2009), the study by Maurissen (Maurissen et all, 2000), and several of those from the Ricerri and Venerosi group (Ricerri *et al.*, 2003 and 2006 and Venerosi *et al.*, 2006, 2008, and 2010), are among the better ones because they generally used multiple doses, had adequate sample sizes, controlled for litter effects, used sound behavioral methods, and used

appropriate statistical methods to analyze the data. Because each of these studies found long-term neurobehavioral effects, these data may be regarded as robust within the limits of what was tested. These more persuasive findings only occurred at doses of 1.0 mg/kg of chlorpyrifos and above. This data set is, therefore, moot concerning effects at <1.0 mg/kg of chlorpyrifos. Among these studies, however, several concerns remain. In the Slotkin and Levin experiments, the use of commercially supplied timed-pregnant rats is a concern as is the use of Zivic-Miller (ZM) Sprague-Dawley rats. The ZM rat is known to be different from other Sprague-Dawley rats on some behavioral tests where it often performs as an outlier. For studies with regulatory implications, it is preferable to use a mainstream rat strain such as Sprague-Dawley from Charles River or Harlan where much more is known about their behavioral characteristics and they do not perform at the extremes of the distribution. Also, the concern about the RAM method as noted in Panel review in Appendix A should be taken into account with regard to robustness. Notwithstanding these caveats, the weight of evidence from the neurobehavioral studies is that there are too many long-term effects for them all to be attributable to Type I errors, hence, it is more likely than not that there are significant long-term adverse effects and in this the Panel concurs with the conclusions of the 2008 SAP findings and the EPA White Paper background document.

c. *Please comment on the degree to which studies that measured AChE inhibition and those that measured neurodevelopmental outcomes can be integrated to evaluate whether points of departure based on 10% AChE inhibition provide more sensitive endpoints than endpoints measured in the experimental neurodevelopmental studies (as reviewed in Section 3.2.2). Please include in your comments a consideration of the strengths and uncertainties associated with this assessment.*

***Response***

Data in the available studies, including the nine additional studies reported since 2008, provide qualitative (emphasis on qualitative) support for the effect of chlorpyrifos exposure during gestation and/or early post-natal period and long-term adverse effect on the developing nervous system. Several of these studies examined AChE activity in the brain after oral and/or subcutaneous chlorpyrifos exposures during postnatal periods, and inhibition of AChE within one day of exposure was observed in these studies at doses as low as 1 mg/kg/day. Since AChE inhibition recovers quickly the data are insufficiently refined to allow for a linkage between the mode of action and the neurodevelopmental effects (acute vs. chronic, respectively).

Since the mode of action of these effects is not established and cannot be presumed *a priori* to be related to AChE inhibition, these studies do not exclude the possibility that other mechanisms may be involved, especially long-term effects where functional characteristics may be unmasked at later life-stages due to neuroplasticity. A few studies have reported AChE inhibition when a dose of 1 mg/kg/d was administered directly to the pup postnatally (Dam, *et al*., 2000; Johnson, *et al*., 2009; Ricceri, *et al*., 2003). However, none of the neurobehavioral studies described in the Panel's review tested for fetal AChE inhibition when 1 mg/kg/d was given during gestation. A companion study to Maurissen

*et al.* (2000) reported no cholinesterase inhibition in samples taken from fetuses 4 h after dosing the dam when 1 mg/kg/d had been administered daily since E6 (Mattsson, *et al.*, 2000). Qiao *et al*. (2002) also reported no brain AChE inhibition in fetuses 24 h after the last dose of 1 mg/kg/d to the dam on E17-20. No other time points or days were assessed in either study. These results suggest, but do not confirm, that the fetus would not experience AChE inhibition at 1 mg/kg/d to the dam, further suggesting that the behavioral effects reported in those studies were not due to AChE inhibition.

The studies published since 2008 demonstrate alterations in a number of neurodevelopmental and biochemical outcomes. The amount of AChE inhibition required to elicit the various endpoints was however inconsistent and varied, because of differences in study designs, analysis of different endpoints, and how long animals were followed-up. Many of the studies measured AChE inhibition 24 hours or longer after dosing, which can underestimate the amount of AChE inhibition. Furthermore, since the neurodevelopmental effects may be independent of AChE inhibition, it needs to be considered whether AChE inhibition represents a critical marker for derivation of points of departure when considering chronic studies.

Finally, the Panel notes that there has been little consideration of the relationship to genetic variability on experimental outcomes, the exception being paraoxonase 1 (PON1). Recovery of AChE activity is linked to changes in AChE gene expression. It has been previously reported that molecular and behavioral effects may be attributable to alternative splicing of the AChE gene. Within the brain there are 2 variants: AChE-S (synaptic) and AChE-R (read-through splice variant) mRNA. Under normal conditions variant AChE-S dominates; however, under stress conditions, such as OP exposure (chlorpyrifos has yet to be studied), the transcription of AChE-R increases. Following stressful events, AChE-R increases to a level that is no longer adaptive and result in varied physiological changes. Of the neurobehavioral effects reported in the reviewed experiments that assessed AChE inhibition, no studies were identified that showed effects on behavior at low levels of AChE inhibition, including at 1.0 mg/kg of chlorpyrifos. Doses below 1.0 mg/kg/day chlorpyrifos did not show convincing evidence of neurobehavioral effect; hence, no extrapolation to lower doses in terms of AChE inhibition is possible from the data reviewed herein.

***Question 4.0 Epidemiology Regarding Children's Health***

Question 4.1

Section 4.0 and Appendices 5 and 6 provide the Agency's review of the available epidemiology studies from the Columbia Mothers and Newborn study, the Mt. Sinai Child Development study, and the Center for Health Assessment of Mothers and Children of Salinas Valley (CHAMACOS) study. Consistent with the 2008 SAP recommendations, the Agency has considered information offered from each of the three cohort investigations; however EPA acknowledges the primacy of the Columbia cohort data for the purposes of informing risk assessment because researchers measured chlorpyrifos parent compound directly in this study. *Please comment on the sufficiency, clarity, and quality of the Agency's epidemiology review as contained in Section 4.0 and*

*Appendices 5 and 6 of the draft issue paper with respect to identifying the major strengths and limitations of each study.*

***Response***

The Panel believes that the epidemiology section of the draft issue paper is very well-written, clear, accurate and fairly complete. The Panel commends the Agency staff on the thorough review of the epidemiology literature, for putting their epidemiology review in the context of the modified Bradford Hill criteria, as recommended by the 2010 SAP, and for reviewing the potential for selection and information biases in each of the studies. In particular, the Panel commends the Agency staff for the tremendous amount of work and thoughtfulness that went into Appendices 5 and 6. The Panel believes that the epidemiology review appropriately concludes that the studies show some consistent associations relating exposure measures to abnormal reflexes in the newborn (using the Brazelton Neonatal Behavioral Assessment Scale), pervasive development disorder at 24 or 36 months, mental development at 7-9 years, and attention and behavior problems at 3 and 5 years of age , in addition to less consistent results for reduced mental and psychomotor development (measured by Bayley scores) at 12 and 24 months. Inconsistent results were found for associations between exposure and measures of fetal growth.

The Agency's epidemiology review provided an excellent description of the strengths and limitations of the studies conducted to examine the relation of chlorpyrifos to children's growth and neurodevelopment. Epidemiologic studies such as these require that large numbers of mothers and infants/children be followed longitudinally for an extended period with extensive data collection at regular intervals to ascertain exposure measures, potential confounders, and health outcomes, all of which can change over time. These studies are logistically difficult to implement and require great commitment by the researchers and a potentially large burden in terms of time and effort on the part of the study participants over a lengthy period of time, often with little or no specific benefit or return to themselves.

The Agency's review adequately summarizes the challenges and scientific contributions of each of three studies: one conducted in an inner city sample of African American or Dominican initially pregnant women and their infants/children by Columbia University investigators; one conducted by Mt. Sinai investigators in a predominately Hispanic and African American sample in New York City, and one conducted in the Salinas Valley in California led by University of California Berkeley investigators. As noted in the review, all three studies had the significant strength of being longitudinal, prospective designs, the most effective design for establishing the temporal sequence in relating exposure to health outcomes, specifically in these studies relating exposure measures obtained prenatally and/or at delivery to outcomes measured at six months up to nine years of age.

Regarding exposure assessment, only the study by the Columbia investigators measured chlorpyrifos parent compound in cord blood, and conducted an exposure validation study and looked at correlations of cord blood measures with mothers blood and meconium, and used the cord blood as the measure of exposure in relation to scores on standard

neurodevelopmental test batteries for children at age 7 years (Rauh *et al.*, 2011). The Panel agrees with the Agency that because this study included the most specific exposure measure, particular attention should be focused on its results. The Panel disagrees with some of the public comments claiming that because the observed associations with adverse neurodevelopmental outcomes occurred at levels below those required for AChE inhibition, the results should be discounted. Instead, the Panel believes that these findings are derived from a well-designed and conducted study and thus suggest that the mechanism may be other than that of AChE inhibition. Further, the very large effect observed in this study for the relationship of exposure to attention deficits, reduced birth weight prior to the voluntary withdrawal of chlorpyrifos, reduced mental and psychomotor development, and reduced Intelligence Quotient (IQ) are unlikely to be due to important misclassification of or bias in assessing exposure or to uncontrolled confounding. [The Panel notes that in Appendix 5, some of the directional signs of effect, *i.e.*, betas, and their 95% confidence intervals (CIs), are missing or incorrect so that all of them should be double-checked and corrected.] Finally, the review might additionally note that the timing of exposure may be important, although the critical time window of exposure is not known with certainty.

Although, the other two studies only used metabolites as markers of chlorpyrifos exposure, all three studies used TCPy, a metabolite specific to chlorpyrifos, in relation to at least some of the outcomes assessed. (Even though TCPy is a better measure than dialkylphosphates (DAPs), the other indirect biomarker, the TCPy measure has limitations as noted by the review. The potential misclassification of exposure should be more explicitly stated. The value of using TCPy as a biomarker also hinges on the mode of action, which is not established with certainty.) As the epidemiology review appropriately notes, all three studies were strengthened in design by using biomarkers of exposure instead of relying on self-reports of exposure, which would be likely to result in much greater misclassification of exposure. All three studies also used similar standard, validated measures of non-verbal and general intelligence, behavior and home environment, which enhance the quality of the outcome assessments and the ability to assess the consistency of the findings regarding these outcomes across studies. In addition, as appropriately noted in the Agency's review, all three studies collected extensive data on potential confounding variables, used appropriate multivariate statistical techniques to control for confounding effects of socioeconomic factors, lifestyle and behavioral factors as well as additional environmental exposures, and conducted sensitivity analyses to determine if assumptions made about missing data, for example, were appropriate.

While agreeing that chlorpyrifos could have played a role in the neurodevelopmental outcomes observed in the Columbia cohort, some panel members raised concern about associating the observed deficits in neurodevelopmental outcomes in children with a single chemical, given that this was a multi-chemical exposure spanning a multi-year period that encompassed an important period of sequential developmental processes necessary for brain maturation. Rauh *et al*. (2011) reported that decreased working memory and full-scale IQ in 7 year-olds were statistically significantly associated with prenatal chlorpyrifos exposure. In an earlier examination of the same cohort, Perera *et al*.

(2009) reported an association between a decrease in full-scale IQ and verbal IQ in 5 year-olds with prenatal polycyclic aromatic hydrocarbons (PAH) exposure rather than chlorpyrifos, thus, raising an issue of the shift in chemical exposure association with increase in age. In each of these analyses, statistical modeling showed that the exposures were independently associated with IQ, and no significant interaction was observed with the other chemical. While this is a statistically sound approach to determine independent responses, panel members noted that it is very difficult to identify the independent physiological effects of a single chemical in this type of multi-chemical exposure scenario. This identification is further complicated by limitations in exposure assessment with respect to on-going and post-natal exposures and the potential for chemical interactions during the exposure period. In addition, developmental progression of the children and the level of skills examined by the tests employed may have been confounding factors. Maturation of the brain is a critically timed sequence of events with each subsequent event dependent upon the successful completion of the previous one. Thus, appropriate brain function at age 7 is dependent on completion of maturation processes that occur at earlier ages. Panel members noted that, while this statistical approach could be used in studies examining the exact same endpoint at a single age, this brain maturation process would need to be taken into consideration prior to determining that at 5 years of age the cognitive deficit was due to one exposure and at 7 years of age it was due to a different chemical. The ever-changing aspect of any developmental study is further demonstrated in the assessments of this cohort of children at earlier ages. At 36 months of age, the deficits in the Bayley Mental Development Index scores were associated with exposures to prenatal chlorpyrifos (Lovasi *et al*., 2011), prenatal phthalates (Whyatt *et al*., 2012), prenatal PAHs (Perera *et al*., 2006), and prenatal piperonyl butoxide (Horton *et al*., 2011). Thus, panel members cautioned about identifying any one specific chemical as the main one associated with the cognitive deficits observed at 7 years of age in the Columbia cohort.

One additional concern is that in general, the sample sizes of the three studies were only moderately large, ranging from just over 100 to slightly under 500, depending on which subset of data from mothers and children were analyzed. The more recent papers had fewer participants, ranging from just under 200 to just over 300. The epidemiology review correctly notes that the modest samples sizes were a limitation in having sufficient statistical power to detect as statistically significant possible modest relations of exposure to outcomes or interactions with other variables. Thus, modest sample sizes were one of the most important limitations of these studies, which is reflected in the wide confidence intervals for some of the effect estimates and the use of moderate (*e.g.* 1 standard deviation) or large (*e.g.*, 10-fold) increases in exposure measures (which did not seem to be mentioned in the Agency's epidemiology review) to see statistically significant effects, *e.g.*, in IQ (Rauh *et al.* 2011; Bouchard *et al.* 2011). However, some evidence of interaction with paraoxonase 1 (PON 1) genotype and/or phenotype was provided in some of the studies (Berkowitz *et al*. 2004, Engel et al 2007; Engel *et al*. 2011; Harley *et al*. 2011), and some examinations of interactions with other exposures were presented in the studies and summarized in the review. Future examination of potential epigenetic effects might also be informative. The Panel also recommended that investigators of the

three studies consider possible pooling of samples and data to enhance the ability to investigate effect modification and possible roles for other agents.

Two other items that might be added to the review are: 1) replacing "null" and "positive (ns)" with point estimates and 95% CIs for effect estimates (to the extent possible, realizing that quantiles or betas might have to be used) to Table 10 on page 59; and 2) noting in the text that other interactions (*e.g.*, with sex of the child, gestational age at measurement of exposure, length of breastfeeding, use of alcohol, *etc*.) were not consistently described in the three studies, and in most cases sample sizes were inadequate to have sufficient statistical power to detect meaningful effect modification as statistically significant. Providing the point estimates and 95% CIs for effects in Table 10 will permit assessment of the magnitude, variability, and direction of the effects, which are more important in assessing consistency than statistical significance. The second point concerning interaction is important because it means that the potential for stronger associations (larger effects) in subgroups with potentially enhanced susceptibility could not be or were not adequately examined or reported. In addition, the Agency's epidemiology review mentions that the restriction of some of the study samples by race/ethnicity and/or to low risk pregnancies (*e.g.*, nonsmokers, women without comorbidities) reduced the potential for confounding (which was a plus) but also reduced the generalizability of the results. However, the review perhaps did not sufficiently emphasize that this limitation also meant that modification of effect by race/ethnicity or other risk factors could not be examined with these study sample restrictions, and the sample sizes in general were inadequate to examine interactions with such factors. Thus, differential effects for subgroups with other risk factors or characteristics could not be determined.

The Agency's epidemiology review also examined the potential for misclassification and bias in each of the studies and mentions the likelihood that any such misclassification and/or biases that operated were non-differential and thus likely to result in an under-estimation of effect. For example, chlorpyrifos exposures, particularly when the parent compound was not measured, could have been misclassified, especially because some analyses indicated greater within-person than between-person variability in exposure measures. However, this was unlikely to be differential with respect to the neurodevelopmental outcomes measured and thus would likely have resulted in bias to the null or under-estimation of effect measures. Similarly, although not explicitly stated in the publications, it was unlikely that those who were assessing outcomes using standardized measurement instruments knew the exposure levels of the participants, which could have biased their assessments. Thus, again, misclassification of outcomes could have occurred but were likely non-differential with respect to exposure levels and thus were likely to have resulted in bias to the null or under-estimation of effects.

The Panel also felt they should respond to the issue of multiple comparisons that was raised in the public comments. The Panel feels it is important to note that all the comparisons made in the three studies were hypothesis-driven and dealt with related outcomes, rather than reflecting "fishing expeditions" that would have been likely to result in significant findings by chance due to multiple comparisons. The Panel thus

believes that the multiple comparisons issue is not an important concern regarding the findings of the three studies over the years.

The Agency's epidemiology review reflects the authors' views from the three studies that among the statistically significant effects seen, most appeared to have a linear relation with exposure with no evidence of a threshold. However, upon examination of some of the graphs and other results presented in some of the papers, it would appear that this point requires some further data and examination. For example, the graphs in the Columbia study seem to suggest no threshold for the effect on working memory but do suggest a threshold for the full-scale IQ (Rauh *et al*. 2011). In the California sample, the graphs presented in the most recent paper (Bouchard *et al*. 2011) suggest a drop in IQ beginning generally with the second quintile of exposure level (depending on which outcome is examined) and seem not to worsen greatly in higher quintiles of exposure levels, which is also suggestive of a threshold effect. The graphs presented in the recent publication from the Mt. Sinai study (Engel *et al*. 2011) seem to indicate no threshold when using tertiles, but the confidence intervals were quite wide. Thus, the Agency's epidemiology review appropriately notes that, due to the modest sample sizes, statistical power may have been inadequate to detect departures from linearity with log transformed exposures or outcomes. It might also be mentioned that modest sample sizes limited statistical power to assess dose-response adequately, which is one of the key postulates promulgated by Bradford Hill, so that such attempts sometimes resulted in wide confidence intervals around effect measures in each quantile, making adequate assessment of dose-response difficult, and that different quantiles were used across the studies, making direct comparisons difficult. Additional analyses of dose-response in both animal and human data and particularly at lower levels of exposure would be very helpful in informing inferences from the epidemiologic studies.

In conjunction with the modified Bradford Hill criteria, the epidemiology review states that a biologically plausible role for chlorpyrifos in relation to adverse neurodevelopmental outcomes is believed to involve inhibition of AChE. While this is a reasonable assumption, the recent papers from the epidemiologic studies noted that noncholinergic mechanisms may play roles in the associations of exposure with the neurodevelopmental outcome measures (Bouchard *et al.* 2011; Rauh *et al*. 2011). In addition, the observed effect modification also suggests other mechanisms, including oxidative stress and lipid peroxidation. The mode of action is discussed in more detail elsewhere in the draft issue paper, but these additional potential mechanisms might be added in the epidemiology review. As noted above, just because the significant effects are observed at exposure levels below which (acetyl cholinesterase) AChE inhibition occurs does not mean that the observed associations are not real, but rather that the mechanism(s) in humans may be other than by AChE inhibition. Further mechanistic work needs to be done to clarify this issue.

In summary, the epidemiology review contained in the draft issue paper is very clearly written, accurate and generally provides a very thorough review in the context of the modified Bradford Hill criteria. As noted above, a few additions would enhance the completeness of the review.

Question 4.2

Similar to the initial conclusions from 2008, the Agency has preliminarily concluded that, qualitatively, chlorpyrifos likely played a role in the neurodevelopmental outcomes reported in the epidemiologic studies, and that information available since 2008, including both new etiologic investigations as well as epidemiologic methods papers, strengthens this conclusion. *Please comment on the Agency's preliminary, qualitative conclusion that chlorpyrifos likely played a role in the neurodevelopmental outcomes observed in the epidemiologic studies. Please include in your comments a discussion of the strengths and uncertainties associated with this preliminary conclusion.*

***Response***

Overall, the Panel reiterates the 2008 SAP's conclusion and the Agency's concurrence with the statement that chlorpyrifos likely plays a role in neurodevelopmental outcomes in the three cohort studies. The qualitative conclusion of the epidemiology review seems well-justified. The Panel agrees with the Agency that although exposures to other AChE-inhibiting compounds cannot be ruled out as contributing to neurodevelopmental outcomes, the potential combination and/or additive effects of these compounds do not rule out the role of chlorpyrifos. However, it should be noted that it cannot be stated that chlorpyrifos is the sole contributor to the observed outcomes.

The conclusion is enhanced by the strengths of the three studies reviewed, specifically:

- the longitudinal designs which permitted clear indications of the temporal relation of chlorpyrifos exposure to adverse neurodevelopmental outcomes;
- the inclusion of biomarkers of exposure as well as self reported exposure;
- the relative consistency of findings in different populations but using similar standardized exposure and outcome measures;
- the strength of the associations found;
- the use of objective measures of exposure and standardized, validated measures of outcomes;
- the control of multiple confounding variables including other environmental exposures and other pesticides;
- the suggestion of a dose-response effect;
- minimization in bias in assessing outcomes and exposures and the likelihood that biases and misclassification of exposures and outcomes resulted in a bias to the null, *i.e.*, under-estimation of effect; and
- attempts to investigate genetic and phenotypic effect modification and dose-response effects.

The conclusion is further supported by the following details of strength and consistency of association and a crude exposure response relationship. Some of these details were previously presented in the 2008 SAP report, but they are reiterated here to include all epidemiologic evidence in one place. It should be noted that studies published since 2008 have continued to show associations between neurodevelopmental outcomes and potential exposure to chlorpyrifos and have strengthened the available epidemiologic

evidence. Recent analyses have looked at neurodevelopmental outcomes in older children and addressed some of the issues of confounding by socioeconomic status, other pesticides, and issues of exposure measurement validation.

Strength of association: This criterion focuses on the Columbia cohort because this cohort specifically measured chlorpyrifos directly from cord blood and therefore has the most robust exposure measurement. Although the results from the other cohorts are useful even if these studies were negated due to non-specific exposure measurement, the Columbia cohort provides a number of strong associations. The effects as described below are seen as early as fetal growth and continue through early childhood with recent evidence of neurodevelopmental effects until age seven.

a) Fetal growth: Statistically significant deficits of birth weight of 186 grams when comparing high exposure to lowest quartile of exposure and decreases of 43 grams in birth weight per log increase in chlorpyrifos in cord blood (Whyatt *et al.*, 2004).

b) Infant neurodevelopment: Statistically significant deficits of 6.5 points on Bayley Psychomotor Development Index (PDI) at 3 years of age when comparing high to low exposure groups (Rauh *et al*., 2006). Notably these decrements in PDI persist even after adjustment for group and individual level socioeconomic variables (Lovasi *et al*., 2010).

c) Increased odds of mental delay (OR=2.4; 95% CI: 1.1-5.1) and psychomotor delay (OR=4.9; 95% CI: 1.8-13.7) at age three when comparing high to low exposure groups (Rauh *et al.*, 2006). When controlling for diazinon and propoxur exposures, chlorpyrifos still showed significant increased odds of mental (OR=3.2; 95% CI: 1.3-8.2) and psychomotor delay (OR=7.9; 95% CI: 2.1-29.1) (Appendix 4, Whyatt & Rauh, 2011 unpublished)

d) Attention problems: Extremely large odds ratios for attention disorders (OR=11.26; 95% CI: 1.79-70.99), ADHD (OR=6.50; 95% CI: 1.09-38.69), and PDD (OR=5.39; 95% CI: 1.21-24.11) were seen when comparing high to low chlorpyrifos exposure groups (Rauh *et al.*, 2006). The magnitude of these results as so large that they are unlikely to be affected by residual confounding although limited sample sizes resulted in imprecise estimates.

e) Intelligence measures: Statistically significant decreases of 1.4% in full scale IQ and 2.8% in working memory among seven-year olds for each standard deviation increase in chlorpyrifos exposure (Rauh *et al.*, 2011). These results persist even when performing sensitivity analyses including only those with detectable chlorpyrifos levels. In addition, no evidence was provided of mediation by child behavior on the measure of working memory instrument.

Consistency of association: This criterion outlines the results from the Berkeley and Mt. Sinai cohorts which were consistent with or supportive of the conclusions

of the Columbia cohort. It should be noted that the Berkeley and Mt. Sinai cohorts did not replicate the effects on fetal growth that were seen in the Columbia cohort. Although the cohorts had similar composition and study design, it should be noted that the Berkeley and Mt. Sinai cohorts used non-specific measures of general organophosphate exposure (TCPy and DAPs). However, the internal validity across cohorts gives confidence in the consistency of the results for the neurodevelopmental outcomes. It should also be noted that neurodevelopmental effects are seen in both of these cohorts beginning at neonatal development and extending to early childhood.

a) Neonatal neurodevelopment: Increased abnormal reflexes in neonates were significantly associated with maternal and urinary DAPs in both the Berkeley and Mt. Sinai cohorts (Young *et al.*, 2005; Engel *et al*., 2007).

b) Infant neurodevelopment: In the Mt. Sinai cohort, prenatal DAP was significantly associated with deficits in Bayley mental development index (MDI) at 12 months among blacks and *Hispa*nics. This as*socia*tion was enhanced among children with maternal carriers of PON1 QR/RR, *i.e*. fast metabolizers (Engel *et al.*, 2011). In the Berkeley cohort significant decreases in MDI at 24 months were associated with increased prenatal and infant urinary DAP measures (Eskenazi *et al.* 2007). Examination by PON1 status also showed evidence of poorer MDI scores at 2 years among those children with the PON1-108T allele (Eskenazi *et al., 20*10).

c) Attention problems: In the Berkeley cohort, total urinary prenatal and postnatal DAP measures were associated with significantly increased odds of PDD at 2 (Eskenazi *et al.*, 2007). In addition, prenatal DAP was associated with ADHD and Child Behavior Checklist attention problems at 5 years. Child concentrations of diethylphosphate (DEP) were also adversely associated with a composite measure of attention (Marks *et al.*, 2010).

d) Intelligence measures: In the Berkeley cohort, a significant deficit of 7 points in full scale IQ was seen among seven year olds when comparing the highest quintile of maternal DAP to the lowest level (Bouchard *et al.*, 2011). In the Mt. Sinai cohort, there were slight but not significant decrement in full scale IQ, perceptual reasoning and working memory associated with prenatal maternal urinary DEP in 6 to 9 year olds. Increased p*renata*l maternal urinary DAP was also associated with decreases in perceptual reasoning in maternal QQ carriers. This association showed a monotonic trend (Engel *et al.*, 2011).

e) Crude exposure response relationship: This was demonstrated in the pre-post residential cancellation analyses in the Columbia cohort in the outcomes of birth weight, birth length, and three year MDI and PDI sco*res (W*hyatt *et al.*, 2004; Rauh *et al*, 2006). In addition a significant reduction in cord blood chlorpyrifos and maternal personal air samples was seen when comparing pre and post cancellation levels (Whyatt *et al.*, 2004). The effectiveness of a

prevention measure can often be shown when reductions in effect can be measured subsequent to a reduction in exposure. This was the case in the natural 'experiment' that occurred during the course of the Columbia cohort. Although the study was not designed to test an exposure-response relationship, decreases in both outcomes and exposure following the residential ban argue for a crude dose-response relationship.

The following uncertainties should be noted:

- Relatively modest sample sizes which limited the statistical power to classify some meaningful differences as statistically significant and to examine the effect of modification by race/ethnicity and other characteristics.
- Relatively moderate to large exposure differences needed to see significant effects, likely due to the modest sample sizes used.
- Exposure at one point in prenatal time with no additional information regarding postnatal exposures.
- Lack of clarity regarding a linear dose-response instead of a potential threshold effect.
- Use of a single or average sample for exposure. Although Whyatt *et al.* (2009) noted moderate but significant correlations between meconium and cord and maternal blood and average urine TCPy, the representativeness of a single point exposure is still unclear. Time-varying exposures or the ability to define cumulative exposures would be preferable.
- Lack of specificity of a critical window of effect and the potential for misclassification of individual exposure measures.
- External generalizability of the cohorts given their unique racial/ethnic and socioeconomic characteristics. However, it should be noted that their exposures were within the range of those seen in NHANES.
- Questions about biologic plausibility due to lack of clarity on mechanism of action, particularly at the low exposure levels seen in the cohorts and the limited and mixed results of animal studies showing neurodevelopmental effects.

One panel member suggested that before the Agency could conclude that chlorpyrifos is likely to play a role in the neurodevelopmental outcomes observed in epidemiologic studies, particularly in the Columbia study, additional analyses need to be conducted.

In order to eliminate the possible causes of neurodevelopmental effects by other pesticides in the Columbia study, it is suggested that EPA should repeat the pre-post residential cancellation analysis done for chlorpyrifos using other pesticide measurements, such as malathion diacid (MDA), a specific metabolite of malathion. The outcomes from those additional analyses will either confirm or reject EPA's preliminary conclusion that chlorpyrifos is likely to play a role in the neurodevelopmental outcomes.

While one panelist agreed with the overall statement, the Panelist also endorsed changes in the phrasing from "chlorpyrifos likely played a role ..." to "chlorpyrifos may [or could] have played a role ." That Panelist noted that TCPy has some serious limitations as a

quantitative indicator of exposure to chlorpyrifos due primarily to its common occurrence in foods. In addition, triethylphosphate (TEP) has some similar limitations particularly within the Berkeley cohort because the usage rate of diazinon in Monterrey County is at least10 times more than the use rate of chlorpyrifos; and diazinon also produces TEP. This ratio presents a dilemma between the characterization of this cohort as farm laborers and the attribution of their higher levels of urinary TCPy to exposure to chlorpyrifos.

In conclusion although the three studies were not comparable in all regards, more similarities than discrepancies were found across them. The Panel concludes that the additional literature since the 2008 SAP continues to support and strengthens the evidence for the conclusion that chlorpyrifos plays a likely role in the adverse effects in child neurodevelopment.

Question 4.3

As discussed in Question 2.0, a mode of action/adverse outcome pathway has not yet been fully elucidated for the potential neurodevelopmental outcomes as a result of prenatal chlorpyrifos exposure. Although this does not undermine the qualitative interpretation of these studies, and the preliminarily conclusion stated above (Question 4.2), the identification of the dose-response for neurodevelopmental effects based on mode of action is not possible. Further, given the urine and cord blood sampling frequency in the study there is a large degree of uncertainty in estimating absolute exposure-response relationships, as opposed to establishing relative exposure groups for evaluating associations. With respect to dose-response, critical durations of exposure, and windows of susceptibility are unknown. In 2008, the SAP cautioned against using the Columbia cohort data for deriving a point of departure due, in part, to only measuring biomarkers (3rd trimester maternal, cord blood, meconium) at one point in time, and because they cannot exclude possibility that the effects seen were due to chlorpyrifos in combination with other pesticides. In 2008, the SAP advised against using data from the epidemiology studies (including the Columbia Mothers and Newborn study which measured chlorpyrifos directly) for deriving a point of departure due to limitations of the exposure assessment in these epidemiology studies for the purpose of risk assessment, *e.g.,* lack of repeated exposure estimates to ascertain more specifically the variability and periodicity of exposure over time (*i.e.,* predominant use of one-time exposure estimate).

*a.* Due to the limitations of exposure assessment performed in the epidemiologic investigations for the purposes of quantitative risk assessment, the Agency has concluded that the epidemiologic data are not sufficient for deriving points of departure for quantitative risk assessment. The Agency proposes that AChE inhibition data from laboratory animals remain the most appropriate data to use for dose-response modeling and the derivation of points of departure. *Please comment on the scientific evidence that does and does not support this conclusion, as well as the strengths and limitations of the evidence.*

***Response***

The Panel acknowledged the limitations in the three longitudinal children's cohort studies of estimating chlorpyrifos exposures (*i.e.,* the Columbia study, the Mt. Sinai study, and

the CHAMACOS study), based on the exposure measures collected, and was in general agreement that the data from these studies alone were not sufficient to derive a point of departure (POD) for purposes of quantitative risk assessment. As a panel member noted, these three epidemiologic studies were primarily focused on assessing health outcomes associated with a variety of environmental factors, and were not designed to conduct a quantitative exposure assessment for chlorpyrifos. In addition, the use by the three studies of different exposure matrices (urine, maternal blood, cord blood, and meconium) and different targeted analytes (TCPy, DAPs, and chlorpyrifos) makes the effort of deriving a definitive POD based on those data alone impossible.

Despite the exposure assessment limitations noted for these three epidemiology studies, the Panel recognized the value of these data and urged the Agency to find ways to use the epidemiology studies, and in particular, the data from the Columbia study, to inform the dose-response assessment of chlorpyrifos. Only the Columbia study provided data on measurements of chlorpyrifos in cord blood coupled with neurodevelopmental measurements. As noted by the Panel, if one assumes that cord blood measurements reflect exposure levels during the critical prenatal period for induction of neurodevelopmental effects, then in theory, these would be the ideal data from which to derive the POD for chlorpyrifos in humans. Specific Panel suggestions included using the Columbia data "as an exercise" to derive a POD for neurodevelopmental effects in infants, and analyzing the data from each of the cohorts to put some bounds on the range of chlorpyrifos doses associated with the observed neurodevelopmental effects.

The Panel also recognized the value in developing a functional PBPK model for chlorpyrifos for pregnancy and the prenatal lifestage. Such a model could be used to further characterize the dose estimates in the epidemiology studies, for additional dose-response analyses. Such a PBPK model will become even more important in the event that the Agency might, at some point in the future, decide to move from using AChE inhibition to another outcome. In particular, such a tool could not only relate a dose of chlorpyrifos to a non-AChE outcome but it could also link a dose to the chlorpyrifos or/and chlorpyrifos oxon concentration *in vitro* to a non-AChE target-site *in vivo*.

The Panel expressed concerns regarding the Agency's proposal to use the dose-response data on AChE inhibition in laboratory animals to derive points of departure for the chlorpyrifos risk assessment, and referred to multiple lines of evidence suggesting that adverse neurodevelopmental effects may be attributed to chlorpyrifos doses lower than those that elicit a 10% inhibition of AChE.

This evidence comes from the epidemiological data derived from the three longitudinal children's cohort studies *i.e.*, the Columbia study, the Mt. Sinai study, and the CHAMACOS study. A number of findings of neurodevelopmental outcomes associated with chlorpyrifos are consistent across these three cohorts. For example, there is a consistent association between chlorpyrifos exposure and deficits in mental development at age 7 as ascertained by decrements in full-scale IQ and Working memory using the Wechsler Intelligence Scale for Children (WISC-IV) (Engel *et al.*, 2011; Bouchard *et al.*,

2011; Rauh *et al*., 2011). (See responses to Charge Questions 4.1 and 4.2 for more detailed discussion and assessment of the findings from these studies.)
There are limitations to the exposure assessment in these three cohorts. The Columbia study has the most direct measure of exposure to chlorpyrifos, measuring the compound in cord and maternal blood at time of delivery (Rauh *et al*., 2011). This study also has 48-hr personal air measurements of chlorpyrifos for pregnant women, air chlorpyrifos measurements (stationary samples) collected during the last 8 weeks of pregnancy, urinary metabolite data (TCPy) during the last trimester (up to 4 measurements for some participants) and at delivery for mom and baby, and TCPy in meconium (Whyatt *et al*., 2007; 2009). In an exposure validation study conducted by the Columbia researchers, the levels of TCPy in meconium and maternal urine correlated with cord blood chlorpyrifos levels (Whyatt *et al*., 2009). This suggests that cord blood levels can be used as a representative measure of exposure. Overall, the estimates of chlorpyrifos exposure in the Columbia cohort (based on measured levels of maternal urinary TCPy) were slightly lower, but generally comparable with the levels of urinary TCPy measured in adults in the general U.S. population at that time, based on the NHANES data for 1999-2000 and 2001-2002 (CDC, 2009). The estimates of chlorpyrifos exposure (based on measured levels of maternal urinary TCPy) in the Mt. Sinai (Berkowitz *et al*., 2003) and CHAMACOS (Eskenazi *et al*., 2007; Castorina *et al*., 2010) cohorts were slightly higher, but generally also comparable to the 1999-2000 and 2001-2002 NHANES data for the U.S. adult population.

The Panel suggested that while there are no data on AChE inhibition in either the Columbia study participants (*e.g.*, Rauh *et al.*, 2006; Whyatt *et al*. 2007; 2009; Rauh *et al*., 2011)or the NHANES participants (CDC, 2009), the measured levels of chlorpyrifos exposure are not anticipated to produce AChE inhibition. Specifically, as noted in the Draft issue paper, neurodevelopmental effects seen in the Columbia cohort were associated with cord blood chlorpyrifos levels > 6.17 pg/g (Rauh *et al*., 2006). Based on AChE inhibition studies in adult men dosed with chlorpyrifos (Nolan, 1984), in which AChE inhibition was associated with peak blood levels of 0.01-0.03 µg/ml (more than $10^4$ more), blood levels of 6.17 pg/g are unlikely to elicit AChE inhibition.

Additional evidence suggesting that adverse neurodevelopmental effects may be attributed to chlorpyrifos doses lower than those that elicit a 10% inhibition of AChE comes from the *in vivo* animal neurodevelopmental studies.

As discussed in response to Charge Question 3, the Panel concluded there are only 3 animal neurobehavioral studies that evaluated doses below 1 mg/kg and also assessed AChE inhibition—2 found no effects at doses below 1 mg/kg (Maurissen *et al*., 2000; Braquenier *et al*., 2010), and one reported effects at 0.3 mg/kg, but had serious methodological flaws (Jett *et al*., 2001). In addition, as discussed in the Agency's Draft Issue Paper and the SAP public meeting presentation entitled "Adverse Outcome pathway: Data for Chlorpyrifos at Varying Levels of Biological Organization", there are another three *in vivo* neurodevelopmental studies conducted in rats that report effects at doses below those at which acetylcholinesterase inhibition was detected (Ray *et al*., 2010;

Aldridge *et al*., 2004; Aldridge *et al*., 2005). These additional three studies are briefly summarized below.

The study, of Ray *et al* 2010, reported differential expression of oxidative stress genes in rat pup forebrain 24 hours after administration of chlorpyrifos, in the absence of AChE inhibition. Briefly, chlorpyrifos was administered via gavage to 7 day old rat pups at doses of 0, 0.1, 0.5, 1, or 2 mg/kg, and forebrain gene expression and AChE activity assessed after 24 hours. Gene expression changes, including differential expression of genes associated with oxidative stress, were observed at all doses, while inhibition of brain AChE was observed only at the highest dose tested (2 mg/kg). As noted on p. 30 of the Agency's Draft Issue Paper, it has been suggested that oxidative stress can result in dysregulation of signaling pathways controlling neuroprogenitor cell function.

The studies of Aldridge and colleagues (Aldridge *et al*., 2004; 2005) reported a number of molecular, biochemical, and functional changes associated with altered serotonergic tone in the brains of adult rats exposed prenatally to chlorpyrifos at doses shown in separate studies employing the same (Qiao *et al*., 2002) or similar (Mattsson *et al*., 2000) experimental designs to not result in fetal AChE inhibition. Briefly, in the Aldridge studies rats were administered chlorpyrifos (0, 1, or 5 mg/kg in DMSO) by subcutaneous injection on gestation days 17-20, and assessed in adulthood for a number of brain parameters on postnatal day 60. Developmental chlorpyrifos exposure at the 1 mg/kg dose level was associated with increases in serotonin receptors (5-$HT_{1A}$ and 5-$HT_2$) (Aldridge *et al*., 2004; 2005), increases in serotonin reuptake receptors (Aldridge *et al*., 2004; 2005), increased serotonin turnover (Aldridge *et al*., 2005), and changes in the adenylate cyclase response to serotonin in the cerebral cortex and mid-brain (Aldridge *et al*., 2004; 2005). While neither of the Aldridge studies measured AChE, another study from this group reported that subcutaneous injection of 1 mg/kg chlorpyrifos on gestation days 17-20 had no significant effect on fetal rat AChE levels, which were measured 24 hours after the last administered dose (Qiao *et al*., 2002). The Panel noted that the measurement of AChE activity 24 hours after the last chlorpyrifos dose, and the use of DMSO as the vehicle, raises some concerns about the validity of the Qiao *et al*. 2002 findings regarding fetal AChE activity. These concerns are tempered somewhat by the study of Mattsson *et al*. (2000), in which pregnant rats were dosed with 0, 0.3, 1, or 5 mg/kg chlorpyrifos (in corn oil) from gestation day 6 through gestation day 20, after which fetal rat AChE activity was assessed 4 hours post-gavage. No inhibition of fetal AChE activity was observed at either the 0.3 or 1 mg/kg dose levels (Mattsson *et al.*, 2000).

Evidence that adverse neurodevelopmental effects may be attributed to chlorpyrifos doses lower than those which elicit a 10% inhibition of AChE also comes from the several *in vitro* mechanistic studies that have been summarized in the Draft Issue Paper, demonstrating a variety of effects at the molecular and cellular level, including interference with neurite and axon outgrowth (Das and Barone, 1999; Howard *et al.et al.*, 2005; Yang *et al.et al.*, 2008), reduced axonal transport (Middlemore-Risher *et al.*, 2011), and increased oxidative stress (Crumpton *et al.*, 2000; Qiao *et al.*, 2005; Giodano et al ., 2007; Saulsbury *et al.*, 2009). Briefly, the study of Das and Barone (1999) in PC12 cells

shows that chlorpyrifos interferes with neurite outgrowth at concentrations that do not inhibit AChE, and the studies of Howard *et al.* (2005) in rat sympathetic neurons and Yang *et al.* (2008) in dorsal root ganglion sensory neurons show that chlorpyrifos decreases axonal outgrowth at concentrations that do not inhibit AChE. The studies of Middlemore-Risher *et al.* (2011) show that incubation of rat cortical neurons with chlorpyrifos or chlorpyrifos oxon reduces axonal transport of mitochondria at concentrations that do not inhibit AChE. The studies of Crumpton *et al.* (2000), Qiao *et al.* (2005), Giodano et al. (2007) and Saulsbury *et al.* (2009) demonstrate that exposures of a variety of cell types (*i.e.*, primary cerebellar granule cells, oligodendrocyte progenitor cells, PC12 cells) to chlorpyrifos at concentrations thought to be so low as not to inhibit AChE result in increased levels of reactive oxygen species and oxidative damage (measured as lipid peroxidation).

As mentioned in the response to Charge Question 2.2, there are additional effects that should be included in the EPA review, namely, the effects of chlorpyrifos on nerve growth factors (Pope *et al.*, 1995; Slotkin *et al.*, 2007; Betancourt and Carr, 2004; Terry *et al.*, 2007) and mitochondrial morphology (Middlemore-Risher *et al.*, 2011). Many of these effects have been observed in the absence of AChE inhibition, or at concentrations below which acetylcholinesterase inhibition would be predicted.

In summary, these lines of evidence suggest that chlorpyrifos can affect neurodevelopment at levels lower than those associated with AChE inhibition, and that the use of AChE inhibition data may not be the most appropriate for dose-response modeling and derivation of a point of departure for assessment of the neurodevelopmental risks of chlorpyrifos.

The Panel suggested additional research that could answer the critical question of whether chlorpyrifos induces neurodevelopmental effects in humans at doses that do not cause AChE inhibition. This suggestion was to test whether the chlorpyrifos levels measured in cord blood that were associated with neurodevelopmental effects in the Columbia study would result in either red blood cell or brain AChE inhibition. This study could be easily performed by EPA researchers, or by others.

Additional concern about the use of AChE inhibition dose-response data to protect against neurodevelopmental effects was based on the potential for the outcomes of AChE inhibition and adverse neurodevelopmental effects to be two separate observations, in which the former is the result of an acute exposure scenario and the latter is likely to be caused by chronic low level exposure to chlorpyrifos *in utero*. All 3 cohort studies report neurodevelopmental outcomes associated with maternal or *in utero* chlorpyrifos exposure measures, which are considered to be representative of chronic exposures during the prenatal period. None of these studies assessed AChE inhibition or other acute responses to recent chlorpyrifos exposures.

Additional questions and concerns about the use of the rodent AChE inhibition dose-response data were raised. The AChE inhibition study that serves as the basis for selecting 0.03 mg/kg/day as the POD (BMDL10) for a benchmark response of 10%

AChE inhibition is Maurissen et al (2000). In this repeat dosing study pregnant dams received daily doses of chlorpyrifos by oral gavage from gestation day (GD) 6 to 20, and red blood cell AChE inhibition was measured 4-5 hours after the last dose of chlorpyrifos was administered. One question raised by the Panel is whether the time of AChE assessment was optimal to detect the peak inhibition effect. The Draft Issue paper provides no information on how the time of AChE assessment in this study was justified by the study investigators, although a general statement on p. 17 indicates that the peak inhibitory effect on AChE activity is typically within one to several hours after dosing.

A second question broached by the Panel regarding Maurissen et al (2000) was whether inhibition of AChE had reached steady state in this study. The Panel noted that a similar BMDL10 for RBC AChE inhibition of 0.044 mg/kg/day was obtained from a companion study (Mattsson et al , 2000) that dosed dams for a longer period of time, *i.e*., from GD6 to lactation day 10. This comparison suggests that steady state inhibition of RBC AChE likely had been reached in the Maurissen et al (2000) study.

A more important question is whether the dose-response for AChE inhibition in the pregnant rat is predictive of AChE inhibition in the human fetus. The Panel cautions the Agency on using pregnant rodent and rodent neonatal/juvenile data as the basis for deriving a point of departure for quantitative calculation of dose-response and risk assessment in human pregnancy and human children for the following reasons:
The AChE inhibition is caused by an oxon of chlorpyrifos that is produced metabolically by CYP450 (P450) metabolism. The isoforms involved include P450 1A2, a 2B isoform, 3A4, 2C9 and 2C19 (there may be others). This presents the following problems with extrapolation from rodents to humans:

- Several of these P450s are highly polymorphic in humans, which will cause considerable variation in human responses.
- The polymorphisms existing in humans may be different from those in rodents.
- Since rodents have different homologues and orthologues, metabolic activation rates and extents may differ between rodents and humans based on differing enzyme affinities for chlorpyrifos.
- Several of these P450s are not active (or only active at very low levels) in the human fetal liver and arise in months-to-years after birth, yet their corresponding rodent P450s are commonly present in the fetal rodent liver.

A positive suggestion in this respect is that much of the ontogeny work in humans and rodents has already been performed and ontogenetic differences are known. For human pediatric CYP ontogenies, the Panel recommended that the Agency explore the work of Professors Ron Hines and J. Steven Leeder to determine qualitative and quantitative differences ((e.g., de Wildt et al, 1999; Pearce et al, 2001; Koukouritaki et al , 2004; Nong et al , 2006; Blake et al, 2007, Hines, 2007; Hines, 2008; Stevens et al, 2008)

Question b. The Agency does, however, believe that the epidemiologic data are useful to informing other key aspects of the chlorpyrifos risk assessment including hazard characterization, exposure characterization, and quantitative uncertainty characterization and analysis. *Please suggest approaches/analyses for potentially using the epidemiology*

*data in different parts of the chlorpyrifos risk assessment including those noted above.* (Note: Some of these may also be covered in Question 5.4 below.)

***Response***

The framework for integrative analysis to evaluate multiple lines of evidence in the context of understanding the AOP/MOA proposed by the Agency is extremely helpful as a basis for framing thoughts on the weight of evidence and the integration of increasingly varied types of information, including epidemiological data.

In relation to the specific use of the epidemiological data to inform key aspects of the chlorpyrifos risk assessment, this is likely best expanded beyond the scope included in the question – – *i.e.*, "hazard characterization, exposure characterization, and quantitative uncertainty characterization and analysis," since the epidemiological data are also informative in the context of dose-response analysis.

Although the panel was not explicitly charged with making a FQPA safety factor determination, one panel member suggested that the epidemiologic data, which represent a significant portion of the evidence base demonstrating increased sensitivity of early lifestages to the neurodevelopmental effects of chlorpyrifos, be used in selecting the Food Quality Protection Act (FQPA) factor to be applied in the risk assessment. The Panel recognizes that it is constituted as a technical advisory body, not a panel intended to provide policy advice. However, the choice to apply particular FQPA safety factors in the EPA's risk assessment involves both policy and science. The FQPA safety factor recommendation is based on the scientific evidence provided to the panel. As discussed in detail in the responses to Charge Questions 4.1, 4.2, and 4.3.a, the strengths of the three longitudinal children's cohort studies, the consistency of associations of chlorpyrifos with neurodevelopmental outcomes across these studies, and the large effect measures observed for serious long-term neurological effects (*e.g.*, attention problems), coupled with data indicating that chlorpyrifos exposures in these cohorts were generally comparable with those of the general U.S population and unlikely to be associated with AChE inhibition, all suggest that in the event that the Agency continues to use dose-response data for AChE inhibition to derive a point of departure, a FQPA factor of 10=fold is recommended to protect sensitive early lifestages,

<u>Exposure Characterization:</u>

Environmental monitoring and biomonitoring data in the epidemiological studies contribute to the overall database on estimation of exposure, including (particularly) population variability and (to some degree) inter-individual variability in the study populations. They also provide insight into more generalizable observations on temporal trends in exposure of the general population – *e.g.*, following the impact of withdrawal of domestic (nonagricultural) uses of chlorpyrifos.

The biomonitoring and environmental monitoring data from the three children's cohort studies should be used, then, along with exposure information from other studies and sources, to characterize the levels of exposure to chlorpyrifos experienced in different

populations (production workers, agricultural workers, individuals exposed via residential use, general population, *etc.*), and in similar populations over time (*e.g.*, before and after cancellation of residential uses).

Data available from the epidemiological studies also provide unique opportunity to investigate the relationship between environmental levels and results of biomonitoring (*e.g.*, dose reconstruction as described by public commenter, Dr. Dale Hattis) since for some of the studies, both types of data (including air monitoring in the Columbia study) are available.

To some degree, the epidemiological studies can also provide sources of data to consider the suitability of the various biomarkers as measures of short and/or long term exposure to chlorpyrifos.

Toxicological Hazard Characterization:

The epidemiological data contribute to an evolving database on potential toxicological hazards to humans. They have contributed and continue to contribute to hypothesis generation for targeted investigations of developmental neurotoxicity in animal studies. To (limited) degree, they also confirm expectations concerning potentially susceptible subgroups based on mode of action – *i.e.*, the PON 1 genotype. They also provide some information on the extent of impact of other factors, which in combination with chlorpyrifos, may have an impact on the observed effects.

These studies represent the key datasets that support the identification of chlorpyrifos prenatal exposures as causing neurodevelopmental effects in humans. Important elements to discuss in their evaluation include i) consistency in the findings of neurodevelopmental effects across these three studies, and ii) comparison of the levels of chlorpyrifos exposure experienced in these cohorts based on biomonitoring (blood and urine measurements of chlorpyrifos, metabolites, *etc.*) and environmental monitoring measures (*e.g.*, personal air monitoring in the Columbia study) with data collected in other studies of the general U.S. population (*e.g.*, NHANES), for similar time periods (*i.e.*, pre- and post-cancellation of residential uses).

Quantitative uncertainty characterization and analysis:

It seems important to address this aspect in the context of dose-response analysis, given particularly, that EPA has concluded that the current epidemiological database strengthens the 2008 SAP conclusion that "chlorpyrifos likely plays a role" in observed adverse effects on child neurodevelopment (specifically those reported by Columbia University). There is also a need to address the consistent epidemiologic findings of significant, long-term neurodevelopmental effects across the three cohorts at levels within the same range as those in the general population since this would seem to suggest that these effects occurred at exposures below those associated with AChE inhibition.

As a minimum, then, it seems important to maximally utilize available data on dose-response from these studies to at least "bound" reference doses developed on the basis of animal data (Given that this was also recommended by the 2008 SAP, prioritization of this work seems critical.). However, the scientific weight given to the different measures of dose and of response necessarily needs to take into consideration that most of the effort in the epidemiological studies has been directed to the assessment of outcome rather than exposure. In addition, the use of different exposure matrices (urine, maternal blood, cord blood, and meconium) and the difference in the targeted analytes (TCPy, DAPs, and chlorpyrifos) complicates derivation of the POD based on epidemiological data, uncertainties which need to be assessed in dose-response evaluation and risk characterization.

In addition, given the potential significance of the observations in the epidemiological studies, it is also clearly desirable to consider at least semi-quantitatively the potential impact of factors of study design and interpretation that bound the dose-response relationship from the human studies. It would be helpful, for example, to consider systematically (and at least semi-quantitatively) the potential impact on the reported dose-response analysis of exposure measurement error, outcome ascertainment, confounding variables and statistical analysis.

For example, in relation to limitations of data on exposure in the epidemiological studies, a Panelist noted that despite a fairly high portion of the samples whose results were below the limit of detection or quantification for whatever was being analyzed, little use was made of techniques to integrate non-quantified samples into the statistical test. [One of the studies utilized a method described by Richardson and Ciampi (2003).] Various methods were reviewed by the July 2010 SAP that can be applied to either normally or lognormally distributed data that include a significant (even a majority) of non-detectable sample Specifically, the use of "probability plots" was described that can yield an estimate of the geometric mean of the distribution [GM], the geometric standard deviation [GSD], and corresponding percentiles. Various aspects of the technique are described in publications such as Cunane (1978), Haas and Scheff (1990), Travis and Land (1990), Helsel (1990), Hattis and Burmaster (1994), and Hattis *et al.* (1999). Another method called the "maximum likelihood estimate" is not recommended for data sets with a large number of measurable values (Cohen, 1961; Perkins *et al.*, 1990).]

As a basis to increase the confidence in the selected point of departure, a relatively simple experimental protocol to determine whether chlorpyrifos levels measured in the cord blood in the Columbia study inhibit either red blood cells or brain AChE inhibition would be helpful. This seems to be an important priority, given that human data (*e.g.*, coupling of chlorpyrifos measurements in cord blood with neurodevelopmental measurements from the Columbia study) would typically be preferred in estimating dose-response relationships (and particularly for potentially susceptible age groups, such as infants).

The outcomes from the above exercise should contribute to consideration of the critical question of whether or not "a causal association between chlorpyrifos and neurodevelopmental effects in the absence of AChE inhibition is plausible for humans."

Given that AChE inhibition results from acute exposure and adverse neurodevelopmental effects are likely to be caused by chronic low levels of chlorpyrifos, it is important to verify whether or not maintaining long-term exposure to levels below those likely to cause AChE inhibition is likely to be sufficiently protective to prevent neurodevelopmental effects.

With regard to quantitative uncertainty characterization, the results and the uncertainties associated with the dose-response analysis of the neurodevelopmental epidemiology findings should be taken into consideration, along with uncertainties in the dose-response assessment for acetylcholinesterase inhibition, data gaps and database uncertainties regarding whether neurodevelopmental effects or acetylcholinesterase inhibition is the most sensitive endpoint in humans, and uncertainties associated with pharmacokinetic differences due to lifestage and genetic polymorphisms in metabolic enzymes.

***Question 5.0 Exposure Profile & Biomonitoring Research***

Question 5.1

c. Section 5 of the draft issue paper presents an overview of the principal chlorpyrifos biomarkers and a comparison of biomonitoring studies that measured urinary TCPy levels in a range of study populations involving both the general population and potentially vulnerable populations, including children, workers, and farm families. *Please comment on the degree to which the Agency identified the primary chlorpyrifos biomarkers of exposure, appropriately discussed the strengths and limitations of such biomarkers, and how the strengths and limitations affect the interpretation of the chlorpyrifos biomonitoring data.*

***Response***

The draft paper was thorough in its coverage of the literature on chlorpyrifos and its biomarkers of exposure. Considering the availability of standard methods, the specificity of the biomarker, the number of laboratories capable of making the measurement, and the relevant concentration levels, the first choice for a biomarker would be chlorpyrifos in blood. The Panel recognizes that this is the most difficult assay and represents only a small percentage of the literature, but it s deemed to be the highest priority because of its specificity.

The next biomarker of choice is TCPy, then DETP/DEP in urine. These have roughly the same equivalence and neither is close to the validity of measuring chlorpyrifos directly in blood because they are both present in the environment as degradates of the active ingredient. Total DAPs (as DMP and DEP) are not selective enough to be a useful biomarker for chlorpyrifos although it may be more appropriate in a global risk calculation model because all AChE inhibiting chemicals should be considered together when evaluating risk. The Panel recognizes the inability of urinary TCPy to distinguish between exposure of chlorpyrifos, chlorpyrifos-methyl, trichlopyr as well as direct exposure to TCPy (a chlorpyrifos degradatein the open environment). However, TCPy is more selective than any of the DAPs and currently is the most selective of the urinary metabolites but questions remain about its efficacy because the Panel believes there could

be significant contribution from environmental and dietary TCPy as a chlorpyrifos degradate.

More emphasis should be placed on the direct intake of the environmental degradate TCPy, mainly present in foods. As early as the late 1990s, the Ryan group (See MacIntosh, *et al.*, 1999) had identified an anomaly in that the amount of TCPy found in urine was substantially greater than the measured likely intake of chlorpyrifos. This work has continued with the papers by Morgan, *et al.*, Wilson, *et al.*, and Lu, *et al*. 2005, 2008, indicating the presence of degradates in foods. Radford *et. al*., 2012 have continued this work on the kinetics of this process. While this work has not yet been published, and the other works are mentioned in the Issues Paper, insufficient emphasis has been placed on the presence of TCPy in food or other exposure media (dust, air particulate, *etc*.) putting into question the utility of urinary TCPy as a useful measure of exposure to the parent compound. During the discussion it was pointed out that Lu et. al., 2005 found that roughly 30% of the TCPy measured in urine could be coming from TCPy directly, present in foods.

The Panel also recognizes the ability to measure AChE and BuChE as biomarkers of exposure, but they are even less specific than DAPs. They are however more indicative of potential health risk and are more than just a biomarker of exposure. Unfortunately the ability to measure these enzymes is likely to vary widely from lab to lab and method to method as they are difficult to calibrate. Changes in cholinesterase activity after an exposure should probably be evaluated more within a laboratory (especially via the use of an unexposed control group) than across laboratories or from study to study.

From its earliest years, measuring AChE has been subject to unresolved inter-day variability (Gage, 1967). For instance, Grob and Harvey (1958) could measure AChE in replicate samples on one day with a standard deviation of ±3% but only to within ±5% on separate days following storage of hemolyzed RBC. The literature dating as far back as Gage (1967) has recommended that in order to measure small changes within an exposed group's cholinesterase activity, researchers should collect blood from an unexposed group of controls, measure the cholinesterase activity in their blood at the same time as the exposed group, and apply a correction factor based on the daily change in the mean of the measured activity in the unexposed group of controls. Yager *et al.* (1976) collected 10 blood samples from 10 unexposed people over five weeks and found that the measured intra-individual coefficient of variation for RBC enzyme activity of ±10% could be reduced to ±6% by controlling the day-to-day component of the variance (*i.e.*, accounting for a shift in the average laboratory results from one day to the next). They also found that plasma activity is more variable between individuals but less variable day-to-day. Similar findings for plasma ChE were reported by Trundle and Marcial (1988) and Brock and Brock (1990).

In an occupational (or other repeated exposure) dose-response study, it is generally cost-effective to adjust the blood ChE results of each member of the exposed or "test" group for the change in the laboratory's reported mean blood ChE of an unexposed or "control" group analyzed at the same time (typically the same day as the post-exposure group or in

the same batch if they were stored). This adjustment has traditionally been made in proportion to the change in the mean AChE of the unexposed controls; however, this form of adjustment could also be applied to plasma ChE except using ChE values. The more accurate fraction of inhibited enzyme [ΔAChE] would be calculated for each subject using this adjusted activity. The final variance of the group would decrease in proportion to the square-root of the number of subjects within the study.

In the future the phase II conjugation products of chlorpyrifos (namely, glucuronides and sulfonates) should be considered. Quantifying conjugative metabolism will ensure that levels of biomarkers are correctly interpreted with respect to biomonitoring data and for performing reverse dosimetry. Even though the AChE adducting oxon is not conjugated, the TCPy and DETP metabolites are extensively biotransformed by the glucuronosyl transferases and sulfotransferases, although the precise isoform pathways are not yet known. Therefore, particularly in the fetus and child, if glucuronidation or sulfonation are saturated and/or ontogenetically deficient, then TCPy and DETP may accumulate. This would almost certainly cause error in biomarker analysis through overestimates of exposure. Moreover, accumulation of these metabolites may present the opportunity for direct metabolite toxicity. Panelists noted there has not been significant effort to look at either the glucuronide or sulfate metabolites possibly because these metabolites have only recently been evaluated both from a physiological and analytical perspective.

The oxon is believed to be the most toxic of the metabolites of chlorpyrifos and is not an environmental degredate. While the oxon does not exist long in the blood, a method to directly measure the oxon in blood is likely to be available in the near future. As the most toxic form and an exclusive measure of exposure to chlorpyrifos, the chlorpyrifos-oxon may be the most predictive biomarker of risk, once a method is published.

When evaluating any of the biomarkers in blood, the EPA will need to consider that some of these biomarkers will differ in concentrations between cord blood and maternal blood as they will have different lipophilicity.

Measuring multiple metabolites simultaneously and then taking ratios of metabolites such as TCPy/DETP represents an untested route to provide greater discrimination between exposure to chlorpyrifos and its degradation product TCPy. However, the Panel could find no direct studies on the stability of these two degradates in the environment. By using this ratio and assuming that the ratio of degradates-to-active in the environment is different from the 1:1 ratio that results from metabolism, it may even be possible to do a source apportionment and separate exposure to the degredate from exposure to the active ingredient.

<u>Other considerations:</u>

The Agency suggested in their public presentation that meconium could be used as a biomarker of fetal chlorpyrifos exposure throughout pregnancy. chlorpyrifos in meconium represent the unmetabolized pesticide. The metabolized form oxon or TCpy may have arrived at the fetus in that form rather than having been metabolized by the

fetus itself. Although there should be chlorpyrifos or metabolites in amniotic fluid if it is found in meconium, no studies exist as to the residence time, flow or amount. It is conceivable (if not totally likely) that chlorpyrifos and metabolites are sent directly into the fetal blood (across the placenta) and that meconium picks it up from sloughed cells. More importantly, the utility of meconium as a cumulative biomarker is uncertain. More specifically, the Panel suggests that this is not (currently) a good idea for the following reasons:

1) No studies of chlorpyrifos in amniotic fluid have been performed. The ratio of aminotic fluid chlorpyrifos to metabolite may add evidence that the fetus is actually metabolizing the chlorpyrifos (if they are developed enough to metabolize the chlorpyrifos).

2) The diffusion and/or transport of chlorpyrifos across the placenta (in either direction) is unknown, but since it is rather fat soluble, equilibration with maternal serum might be postulated. This does not seem to be the case with at least one umbilical: maternal serum study presented at this meeting (Yan, 2010).

3) Metabolism across the placenta is unknown. Are the metabolites passed or only the chlorpyrifos, possibly the oxon?

4) The contribution of umbilical tissue, including any adducting of cord tissue for example by the oxon is unknown. However, since umbilical tissue is so well perfused, it may be expected to be a target for oxon binding. Umbilical cord tissue consists of a polymatrix of Wharton's jelly, which is made up of mucopolysaccharides. (Kliman, 1998). Based on these characteristics, the very fat-soluble nature of chlorpyrifos and its relatively fat soluble metabolites, the umbilical cord would not be expected to function as a good reservoir of the parent compound or metabolites, but may be a target for oxon binding and deregulation of pregnancy homeostasis.

5) Meconium, being composed mostly of intestinal epithelia, lanugo, mucus, amniotic fluid, bile, and water, is reasonably hydrophilic and thus should also be considered a poor matrix (reservoir) for chlorpyrifos and other fat-soluble xenobiotics to accumulate. It may be marginally better for the metabolites TCPy and DETP, but these molecules would also, in addition to diffusion, have some net flow in the paracellular pathway. Thus, they may over represent the exposure to chlorpyrifos (as described above).

6) Human placental studies of chlorpyrifos metabolism and transport have not, to the best current knowledge been published, and this is a limitation of data available to the Agency. Several important points are already known from the illicit drug literature and should be considered when attempting the same type of monitoring for chlorpyrifos (or other xenobiotics). For example, antipyrine (an amphetamine derivative) is used in placental perfusion experiments as a marker of pure diffusive transport with effectively no barrier (Schneider et. al. 1972). In contrast, cocaine and cotinine show differing and, slightly less fat-soluble profiles. For example, cocaine is transferred across the placenta

at only 80% the rate of antipyrine (Schenker et. al.,1993) and some studies have suggested that the placenta acts as a depot for cocaine accumulation preventing transfer to the fetus (Simone et.al. 1994). Additionally, previous studies have indicated that while nicotine (again highly fat soluble) is transferred into the fetal compartment up to 5 times the concentration in the maternal blood, cotinine concentrations in the fetal compartment were considerably lower than corresponding maternal serum levels (Luck et. al. 1985). Again, it has been suggested that cotinine adducts the placenta, preventing equilibration of concentrations between maternal and fetal systems. These studies support the need for greater consideration of the trans-placental characteristics of chlorpyrifos, and since placental characteristics change drastically by term (the placental barrier becomes increasingly "leaky" after ~36 weeks), placental studies need to consider each trimester. In the first trimester, the placenta is perfused only after ~8 weeks; prior to 8 weeks only active transport or diffusion across the placenta can occur because villi are being blocked. Analgous studies for chlorpyrifos are recommended before extrapolating fetal exposure and may be included as part of a longitudinal study in pregnancy. Such a longitudinal study may present additional problems.

The real question for the Agency is almost certainly not related to the fetal load of chlorpyrifos or its metabolites at birth or even at discrete pregnancy time points. The exposure information (fetal load) needs to be correlated to a time in fetal development when the fetus is susceptible to effects of chlorpyrifos, perhaps during critical points of neurodevelopment. Unless the time of these exposures can be definitively correlated with specific adverse health effects, then consensus on how to relate fetal effects to a biomarker concentration is unlikely. Rather, the Agency seems to be seeking to quantify the amount of maternal chlorpyrifos ingestionor/exposure is subsequently experienced by the fetus. In general, the half-lives of chlorpyrifos and its metabolites are rather short. This means that even in the case of TCPy, which has the longest systemic residence time, the terminal half-life (*i.e.*,. complete clearance of TCPy from the fetal compartment) would occur within several days. Therefore, unless the pregnant woman is exposed to chlorpyrifos either chronically or acutely exposed but within a few days of testing, quantifying the chlorpyrifos exposure of the fetus would be difficult. It would require collection of samples from pre-term as well as [full or near full term fetal tissues or sampling directly from placentas (such as chorionic villus sampling), amniotic fluid (amniocentesis), or umbilical blood. A longitudinal study would almost certainly be needed to determine exposure over pregnancy, which may not be cumulative but pulsatile.

These points highlight the uncertainty of using meconium as a measure of exposure over the course of pregnancy at this time. Essentially, production of meconium is from fetal swallowing of amniotic fluid as well as some sloughing of intestinal epithelia, and meconium should not be thought of as a matrix into which chlorpyrifos or its metabolites may accumulate by simple diffusion through fetal tissues.

*d.* Section 5 of the draft issue paper compares biomonitoring findings from the three children's health cohorts with other major observational exposure studies in the United States. Based on comparison with NHANES 2001-2002, median TCPy levels in the

CHAMACOS and Mount Sinai cohorts were slightly higher than in the general population. It should be noted that the exposures experienced by the CHAMACOS and Mount Sinai cohorts overlapped the start of the residential chlorpyrifos phase-out. By contrast, median TCPy levels in the Columbia cohort, for which sampling occurred when chlorpyrifos use should have rapidly declined due to the voluntary cancelation, were slightly lower than the levels measured by NHANES in the general population. *Please comment on the adequacy of the Agency's comparison for the purposes of evaluating chlorpyrifos exposure levels in the three children's health cohorts. Are there any additional biomonitoring studies that should included in the Agency's comparison?*

***Response***

The human studies discussed in this section are the best available. They were carefully designed and well implemented. They do, however, look at specific types of exposure: agriculturally based exposure and exposures in city dwelling units likely treated for insects on a regular basis. Further, they span a range of times from when chlorpyrifos use was ubiquitous through the phase-out of indoor uses of the insecticide. Because of this, there are "inconsistencies" in the data that are indicative of changes in use patterns. Current use in indoor settings is dominated by pyrethroids rather than chlorpyrifos. Agricultural settings are still likely to see large exposures to chlorpyrifos (although apparently not in the county surrounding Salinas, CA). There appears to be inconsistent recognition of this change, especially in light of comparisons with "group norms" via, for example, the NHANES studies. It would be to no one's surprise if the 1990-2000 NHANES data indicate higher exposures to chlorpyrifos in residential settings than the later data. Among these three studies the Panel believes the Columbia study has a particular importance because it has data collected before and after the indoor use "ban," and the results reflect the pathway from exposure to biomarker concentration and health outcome.

The Panel recommended the following order in which the studies should be considered. They believe that the next NHANES data set may be the most important as it is likely to reflect the decrease in exposure caused by the voluntary removal of chlorpyrifos from the home market. If the levels progress in a manner similar to those predicted by the decrease demonstrated in the Columbia study, the risk from chlorpyrifos might also decline as rapidly. Even if this is true, chlorpyrifos as a model compound for a risk paradigm that includes epidemiological, dose reconstruction, PBPK modeling, and exposure dosimetry, requires a much broader consideration of studies. The Agency seemed to concentrate on studies that include a reported health outcome, and the Panel wonders why these studies were the principal focus as many studies provide data on exposure and dose. For example to be protective the agency should consider the National Human Exposure Assessment Survey, (NHEXAS-Az, summarized in Egeghy et al. (2011) study many of the participants from Arizona were exposed through agricultural application of chlorpyrifos and this represents the highest non manufacturing level of exposure and may continue to represent direct or indirect agricultural exposure.

Many of the studies listed in the draft paper but not directly discussed should be considered when estimating dose and subsequently risk. The NJ studies where cord

blood measurements were used as the principal sample type are important because that is likely to be the desirable biomarker and used more frequently in future studies. Farm-workers studies are important because their families are likely to be one of the remaining populations that continue to see significant exposure, again an expectation to be validated by the next round of NHANES data. They should however probably be considered primarily in relation to farm workers' families. The Children's Pesticide Exposure Study (CPES) by Lu, 2009 and Children's Post-Pesticide Application Exposure Study (CPPAES) studies are important because they provide data on multiple exposure vehicles/media and will be especially useful in dose reconstruction. Dose reconstruction will be paramount in validating PBPK models using media (dust, food, air particulate) measured concentrations and estimated exposure levels, to be subsequently discussed sections 5.3. Among the current studies those that look at both the urinary concentrations and the media where the exposure is likely to occur, will provide the best models for closing the knowledge gap between exposure and dose; and studies where urine was collected within one half-life after a fresh exposure may provide the most useful information.

Although not ready for this report, studies now underway that are longitudinal in design will afford a better understanding of actual exposure profiles when compared to cross-sectional approaches. Due to the short biological half-lives of the metabolites of chlorpyrifos in the body, a spot check of a relatively small number of people may not be enough to represent the exposure of a vulnerable population at key time periods. Only a longitudinal investigation can get at these important data.

Several new studies of interest to this group have been completed and will be published in the near future that are. The Children's Pesticide Exposure Study, led by Dr. Alex Lu of Harvard School of Public Health, focuses on dietary intake of children and related pesticide exposures. The Children's Exposure to Environmental Pesticides, led by Dr. P. Barry Ryan of Emory University, evaluated the utility of biomarkers of pesticide exposures, *e.g.*, DAPs and pesticide-specific markers of OP and pyrethroid exposures, and environmental levels measured in soil, house dust, and food. The target population is children ages 3-6. The SAWASDEE study, led by Drs. Dana Boyd Barr and Ryan, and Dr. Tippawan Prapamontol of Chiang Mai University in Northern Thailand, examined pesticide biomarker concentrations in pregnant mothers, and similar markers in their newborn children. Multiple measurements in both urine and serum have been made throughout pregnancy giving a better longitudinal picture of exposure. Several smaller investigations are underway designed to evaluate the direct intake of pesticide degradates and to evaluate the kinetics of the degradation process in environmental media, including food.

In comparing the results from study to study, it is important confirm that analytical results are directly comparable. Data quality of some studies has been called into question due to apparent changes in limits of detection associated with two analytical methods developed by Center for Disease Control (CDC) used to evaluate serum chlorpyrifos concentrations. The questions arose due to a misunderstanding of the methods. There is an apparent 20-fold difference in the limits of detection (LOD)

between the two methods. This can be accounted for in three ways. First, the "newer" method uses a sample size one-half as large as the "older" method, and injects one-half as large an aliquot thereby accounting for a factor of four difference in LOD. Second, although both methods are multi-contaminant, the newer method spans a much larger range of analyte polarities. In order to obtain adequate recoveries for some of the less polar compounds, there is some sacrifice in sensitivity toward more polar compounds, such as chlorpyrifos.

Third, the newer method was developed with the expectation that higher concentrations would be evident in the samples analyzed, hence precluding the need for a lower limit of detection; a listed limit of detection of 10 ppb was adequate for the purposes of the study. Attribution to the new, higher limit of detection to samples analyzed by the older, more sensitive method, is therefore not warranted. The value for the LOD determined for the earlier method should be viewed as appropriate for the samples analyzed by that methods and deemed useful for presentation in any other work.

In addition to analytical differences these studies (Columbia, CHAMACOS and Mount Sinai) are all cross-sectional in design with some repeated measurements during the pregnancy period. Because of the cross-sectional design coupled with the short biological half-life of chlorpyrifos, the spot urine measurement would be highly affected by daily chlorpyrifos exposure, as well as the timing of sample collection. It would be great if all three epidemiologic studies were using the identical sampling protocol so the outcome measurements could be compared across the board they weren't. Generally speaking, it should not be surprising either to see the similarities in the CHAMACOS and Mount Sinai cohorts during the period they overlapped the Columbia study, before the residential chlorpyrifos phase-out. It is likely that dietary exposure to chlorpyrifos in these two cohorts may represent a portion of the overall exposure. However, the Columbia cohorts may differ from the CHAMACOS and Mount Sinai cohorts because the Columbia study reported the reduction of chlorpyrifos in the indoor air after the phase-out.

Question 5.2

In Section 5.0 of the draft issue paper, the Agency summarized the 2008 preliminary findings on the association between urinary TCPy levels and AChE/BuChE inhibition and discussed two recent studies involving manufacturing workers in the US and Egypt. *Please comment on the scientific quality of these studies and their findings. Please include a discussion of their strengths and limitations. Please comment on the strengths and limitations of the evidence from this research to show an association between TCPy and AChE/BuChE inhibition at exposure levels experienced by occupational populations.*

***Response***

Both of the occupational exposure studies were observational in nature. Garabrant *et al.* (2009) involved 53 workers manufacturing chlorpyrifos in Michigan, while Farahat *et al.* (2001) involved 38 field workers applying chlorpyrifos onto cotton plants in Egypt. Both of these studies contain data that have multiple sources of imprecision (as will be detailed below), but they both included enough participants that their overall results match PBPK

model predictions quite well. In many ways both studies were well designed and implemented. Both studies had sufficient power to show an association between TCPy and AChE or/and BuChE inhibition at exposure levels experienced by occupational populations. In fact, the PBPK model and cholinesterase data confirm that chlorpyrifos once absorbed interacts first with BuChE and only starts to inhibit RBC AChE and AChE in the central nerve system after BuChE is more than 50% inhibited.

Perhaps the most unique feature of the Farahat study was the extremely high levels of TCPy found in urine from these field workers after applying chlorpyrifos to the target cotton fields. For example, the mean post-exposure values of urinary TCPy were about 25× more than the TCPy from the manufacturers reported by Garabrant *et al*. and over 1000× more than those in the women and children cohorts discussed in Section 4. On the one hand, the Panel pointed out that this contrast made this study less relevant to our discussion. On the other hand, a major strength of the study is that not only were the qualitative patterns of both BuChE and AChE activities when paired to urinary TPCy from the same individuals qualitatively similar to the patterns predicted using the PBPK model described by Timchalk *et al*. (2002) and used by Garabrant *et al*. (2009), but the "inflection points" within the paired data closely match those predicted by the PBPK model. This correspondence between the measured and predicted TCPy excretions and cholinesterase inhibitions is strong evidence for the robustness of the PBPK model over a wide range of exposures.

The Panel pointed out five weaknesses within the Farahat study for use within the weight of evidence. First, virtually all of these field workers had high levels of TCPy in their pre-exposure urine samples. These background concentrations (with sub-group means ranging from 10 to 2000 µg TCPy/g creatinine) were up to three orders of magnitude higher than the levels in the women and children epi cohorts. The source of this background is unknown but seems likely to have been due in large part to these workers' prior use of chlorpyrifos outside of the jobs being studied and possible contributions from TCPy on chlorpyrifos treated food and from TCPy or/and chlorpyrifos within homes treated with chlorpyrifos. Second, the urine samples were collected from morning voids that a study by Lu *et al*. (2006) found to be less reliable than evening voids. Nonetheless, these high background TCPy levels jumped about 30× after the applications began. Thirdly, the cholinesterase values were measured by the battery-powered kit based on the Ellman method. Prior publications (including one by the same researchers who participated in the Farahat study) concluded that cholinesterase activities measured by those field test kits are not as reproducible either from kit-to-kit or as a function of temperature as those using more robust clinical methods (Oliveira *et al*, 2002; Hofmann *et al*., 2008). Fourth, the study was not designed to analyze blood samples from an unexposed control group concurrent with blood from their field workers; the importance of such control was discussed in the Panel's response to Charge Question # 5.1. This deficiency further weakens the precision of their cholinesterase results which was offset somewhat by having 38 participants. Lastly, for reasons not stated, the authors chose to report (and plot) individual cholinesterase activities rather than inhibitions in comparison to individual baseline values. Thus, the reader is led to believe that the ratio of an activity of 2 U/g Hgb for the individual with the lowest AChE and the highest TCPy to

about 25 U/g Hgb for the cluster of individuals with the highest AChE and lowest TCPy measurements (in their Figure 3) represents an inhibition of almost 90%. While some found this degree of inhibition incredulous, Grob *et al.* (1947) and Grob and Harvey (1958) showed that a sequence of small oral doses of an OP (DFP) delivered over three to five days can cause someone's AChE to be reduced down to about 1% of their normal level or to be 99% inhibited) but still not cause symptoms if delivered slowly enough. But of course, a fractional ΔAChE inhibition of 30-50% in one day can cause acute symptoms (*e.g.*, Gage, 1967; Reigart and Roberts, 1999). Thus, the idea that any particular level of ΔAChE either is or is not clinically important depends on more than just its numeric value. Another troubling observation in Farahat *et al*. (2011) is the persistent elevated TCPy measurements in some of the workers and the persistent depressed RBC AChE 14 days post-application; perhaps these lingering effects are linked to the high preexposure levels or the inhibition may have "aged." With these caveats, not only does the pattern of paired levels of AChE activity and concentrations of TCPy in urine qualitatively match the pattern predicted by the Timchalk PBPK model, but also the value of the mean of four measured AChE inflection points at 3161 µg TCPy/g creatinine quantitatively matches the inflection point predicted for AChE by that model.

The study reported by Garabrant *et al.* (2009) has some broadly similar and some different weaknesses for use within the weight of evidence. One different weakness is the greater potential for a proportion of the chlorpyrifos employees' urinary TCPy to have come from doses of residues of TCPy that might have accumulated within the manufacturing workplace (Burns *et al*., 2006). This study added urine collection to an on-going occupational health monitoring program that involved monthly blood samples that were analyzed for cholinesterase via a proprietary system (*Vitros* by Johnson & Johnson) with which the Panel was not familiar. The time at which the pre-exposure cholinesterase was measured was not stated but could have been some years earlier. Despite the study having a "referent group," there is no indication that the cholinesterase results for the chlorpyrifos workers were adjusted for variations in the results of blood samples from an unexposed control group (again see CQ#5.1). The three urine samples per person collected in this study were also collected in the morning (first voids in this case); however, an additional source of uncertainty was introduced into the results of Garabrant *et al*. because the blood and urine samples were collected between 5 and 14 days apart. The authors concluded that conducting paired analyses using only the 48% of the urine results that were collected within 7 days of a blood sample was optimum; however, this interval spans several half-lives for TCPy within the human body. The range of ΔAChEs reported in this study slightly exceeded ±20% but, as predicted by the PBPK model, showed no correlation with TCPy. Only the BuChE inhibition could be attributed to chlorpyrifos exposures. Indeed, the inflection point for ΔBuChE found by this study (110 µg TCPy/g creatine)) not only matched that found by Farahat *et al.* (114 µg TCPy/g creatinine) but also matched that predicted by the PBPK model.

In the responses to Charge Questions 4.2 and 4.3, the Panel suggested that the Agency should separate scenarios for occupational exposures, as reported in these two studies, from exposures from environmental sources. Indeed, one panel member suggested that data from Farahat *et al.* (2011) should not be considered for any further uses. It should

be noted that the subjects in these two studies were adults. Although participants in Farahat's study were as young as 15 and roughly 25% of participants in Garabrant's study were females, none are directly comparable to newborn infants. Even the extrapolation of any working population to the population as a whole is subject to criticism. Such criticisms include the "healthy worker effect" and the idea that low-level exposure and high-level exposures are likely to be detoxified by differing mechanisms. Studies of agricultural workers and their families could offer a better avenue of investigation that compares "occupational-levels" exposure with other members of their families likely see slightly "elevated" but lower levels of exposure, and to study the potential impact on the offspring in such cohorts exposed either *in utero* or otherwise. In the future, the Agency should take into account the quality of ChE measurements prior to further uses in the exposure and risk assessments.

Question 5.3
Several approaches ranging from qualitative to the most sophisticated PBPK/PD modeling approach were introduced as potential options for analyzing the chlorpyrifos biomonitoring data. *Please comment on the strengths and limitations of these approaches. In addition, please suggest, if appropriate, alternative approaches or analyses not identified by the Agency.*

***Response***
The increasingly data-informed options for interpreting biomonitoring data presented by the Agency range from qualitative (non-comparative, looking at trends or comparative, taking into consideration controlled human studies data where ACN inhibition has been measured) to semi quantitative approaches (estimating biomarker levels associated with regulatory exposure guidelines or estimating exposures from biomarker levels using reverse dosimetry or a PBPK model).

Presentation of a number of options in an increasingly data-informed construct of this nature has potential to maximize the use of biomonitoring data for different applications (accounting internally for more factors contributing to variability in exposure than do external estimate), taking into account (relative) uncertainty depending on: availability and specific nature of biomonitoring data, and the intent of use (*i.e.*, what degree of uncertainty is acceptable for the intended purpose; what population; and what application?).

The selection of appropriate options is necessarily dependent on the extent of the data available on toxicokinetics relevant to the population subset and mode of action, and their integration, with a verified PBPK model having the potential to be the most informative, but being the most data intensive. In relation to intended application, for example, if the objective is media specific assessment or management, dose reconstruction (reverse dosimetry) from biomonitoring data is required.

As a minimum, currently, the biomonitoring data on chlorpyrifos should be helpful in "ground truthing" total external exposure estimates under various use conditions, which

are necessarily based on many more assumptions such as activity patterns and intakes and concentrations in various media.

Given the availability of biomonitoring data on chlorpyrifos in the general population, and as a basis to encourage its maximal consideration in a public health risk context, the Agency is also encouraged to seriously consider the development of a value akin to a "biomonitoring equivalent" concurrently with the derivation of a reference dose for chlorpyrifos. (A biomonitoring equivalent (BE), is a calculated level of a biomarker associated with exposures consistent with health protective guidance values for the general population). This BE would provide a valuable addition for interpretation of population biomonitoring data with limited additional effort, drawing efficiently on the existing process for review and consultation for the regulatory assessment (*i.e.*, BEs are based on similar considerations as the reference dose but incorporating toxicokinetic translation to internal doses).

Clearly, a verified PBPK model provides the most robust opportunity to integrate the considerable available data on external and internal exposure (*i.e.*, biomonitoring) to chlorpyrifos at different life stages under different conditions of exposure.

As indicated in the response to Q. 5.4a), prediction of excretion by the PBPK model can potentially be validated or verified with an accurate estimate of dose, through dose reconstruction based on data from the epidemiological studies on the concentration of chlorpyrifos in media such as house dust, air and water combined with market basket data on the concentration of chlorpyrifos on food. This would permit the effective prediction of exposure at the critical windows of maximum effect (*i.e.* AChE suppression) with measured urine concentrations. However, it's somewhat unclear currently based on input at the meeting from Dr. Bartels of Dow Chemical and Agency staff whether or not the developed PBPK model is life-stage specific. In the interest of addressing this need, the following recommendations are offered: If an adult PK or PBPK model is used, simple allometric scaling (3/4 power) or scaling based on Wang's modification of the Dubois and Dubois equation (Wang *et al.*, 1992) can be useful, relatively accurate and robust for extrapolating to children (Anderson *et al.*, 2009; Anderson, 2010). This is a simple way to improve prediction for pediatric populations. Moreover, plasma proteins differ drastically in infants (and in pregnant women); since chlorpyrifos is so highly protein bound, this should be taken into account, but may be less important for TCPy or DETP.

A sophisticated PB/PK model for children is also available that allows for flexible inputs (*i.e.*, SimCYP pediatric (SimCYP Company, Sheffield, UK). Although building a pregnancy PK or PBPK model is challenging and ambitious, it was extremely gratifying to see Dr. Hattis' progress on development of a multi-compartment model where the fetal compartment (including the fetus, amniotic sac/fluid and placenta) is separately considered. While it is acknowledged that this will affect outcomes from Dr. Hattis' current oral exposure model (but less so the inhalational) by altering first pass, it's important to recognize that at term, the placenta is perfused to ~600 mL/minute of maternal blood (*i.e.*, the equivalent of the entire mother's blood supply passes through the placenta in about 8 minutes) and has an average surface area of 11 m2. Moreover it

expresses significant CYPs, UGTs and SULTs that have been implicated in chlorpyrifos metabolism (Benirschke *et al.*, 2006). These considerations are relevant to the importance of the feto-placental unit as a separate compartment which is both well perfused and metabolic.

Additionally, while passage from maternal blood, to placenta and fetal blood may be bidirectional, distribution into amniotic fluid is uncertain; it would be helpful, then, to confirm whether or not placental effects might be negligible, retaining the placenta as part of the "liver metabolism." Based on similar scenarios for bisphenol A (BPA), this is not at all certain.

In response to a request from the Agency, it was clarified that the uncertainties in any PBPK model cannot be estimated at this time, since working model parameters [for the Agency's assessment] are not yet defined. In response to a further request from the Agency, Panelists suggest that a more expeditious path to attaining reasonable estimates of fetal exposure would be to generate an equation or algorithm that describes the relationship between maternal serum levels and cord blood levels of chlorpyrifos or its metabolites. Although less certain than the output of a verified PBPK model, this would enable basic dose reconstruction that can then be validated or verified by comparison to parameters in urine and blood reported in epidemiological studies. This may also be a starting point for assessing fetal exposure by defining "flow" and for continuing to build a more sophisticated model. The main limitation in using an equation describing the maternal: fetal ratio (and hence the flow) of chlorpyrifos and/or its metabolites is their short systemic residence time (*i.e*. the blood may only reflect exposures up to a few days prior to blood sampling). As a result, this method will not necessarily reflect cumulative exposure or acute exposures in earlier prenatal periods.

It was also noted that reported relationships between chlorpyrifos in maternal and cord blood warrant reconsideration. In particular, the ratio of 1.05 between the mean values 3.9 pf chlorpyrifos / g maternal blood to 3.7 pf chlorpyrifos / g cord blood in Table 2 of Whyatt *et al.* 2005 differs widely from the ratio of 1.49 between the mean value of 5.96 pg chlorpyrifos/g maternal blood derived from the regression equation Cord = 1.03 Maternal 0.76 given in Whyatt *et al.* 2004 and 4.0 pg chlorpyrifos/g cord blood in Table 1 of that publication. Only the former ratio was referred to in discussions of current PBPK models. Independent of whether or not these ratios represent the same populations, the broad range of this relationship needs to be defined.

Question 5.4

Characterization of chlorpyrifos exposure experienced by women in the Columbia cohort, particularly during the pre-cancellation period, remains an important uncertainty in using these data in quantitative risk assessment. Exposure levels in the range measured in the cord blood data from the epidemiology studies (pg/g plasma) are probably low enough that is unlikely that the cohort mothers were experiencing AChE inhibition at the time of delivery; however, the biomonitoring data were taken after birth and not necessarily associated in time with an application of chlorpyrifos. As such, the actual level of such exposure particularly during any critical window(s) of susceptibility is not known, and a

better understanding of the range of possible exposures and the degree to which they may or may not have elicited inhibition of AChE, remains a key scientific question. *In light of Panel discussions of Questions 4.3 and 5.3, please suggest approaches and/or analyses which would inform the understanding of the degree to which exposure levels experienced by the Columbia cohort participants may or may not have been below doses which result in 10% inhibition of AChE in the most sensitive lifestage. Please discuss the strengths and uncertainties associated with such analyses. Please include in your discussions approaches involving chlorpyrifos and its metabolites and also chlorpyrifos plus other AChE-inhibiting pesticides (propoxur, diazinon) which the cohort participants were exposed too.*

***Response***

It is important to realize that the short half-life of chlorpyrifos and its metabolites in the body calls into question any" spot data" that might be used. Large cross-sectional investigations may "catch" some exposure, but do not put them in context. Only longitudinal investigations, with frequent sampling are likely to give results that are of real use.

What is called for in estimating the peak dose is prediction of the dose-response curve that would correspond to the vulnerable populations that were exposed. Understanding the limitations of the data available, a PBPK model having the potential to estimate dose given a fixed time since exposure, may provide some information. Additional information that would still be required for a reasonable estimation of maximum dose includes, whether the exposure/dose was steady state or bolus and approximately how long after the bolus exposure was the sample collected. With a very simple one compartment model and a time after exposure a reasonable estimate of the maximum dose can be calculated as well as whether the AChE inhibition threshold was reached. A more sophisticated PBPK model may provide even better data assuming that the PBPK model is applicable to the population being studied, specifically to pregnant women and small children.

Previous Panels have noted the decided lack of a realistic PBPK model for chlorpyrifos for all populations. An effective PBPK model that is applicable to target groups such as pregnant women and infants/small children should be used for these vulnerable populations. An effective commercial version has already been identified for infants/small children (SimCYP pediatric from the SimCYP Company, Sheffield, UK) and should be used for a more comprehensive risk assessment model. Children are potential targets for any developmental issues related to exposure and while there are effective PBPK models for children, they have yet to be discussed here. Utilizing PBPK models designed for the individual and unique demographic (e.g. children and pregnant women) means more than the adjustment of body mass within the model designed for adult males.

In assessing both exposure and dose, a significant data gap exists for the population as a whole but especially for pregnant women which should be addressed with a longitudinal study. A single dose PK study like Clement (1984) provides the foundation for like

populations (adult males) but does not address steady state (or the approximation that is our real world exposures) or populations with different metabolic conditions such as pregnant women or children. As discussed in Section 5.2, progress on this front has already been made as a compartment model with the fetus as a compartment currently exists.

A longitudinal study throughout the pregnancy rather than a few samples in the last trimester would fill many of the data gaps that currently exist for this group. The potential for neurodevelopmental affects on the fetus as well as the metabolic differences in pregnant women versus the workers from the 1984 study, necessitate such a study. Placental tissue might provide more information on the metabolism and the delivered dose to the fetus as the concentration of chlorpyrifos going into the fetus cannot be measured directly from the cord blood or from the difference between cord blood and maternal blood. The tissue concentration may provide information on the chlorpyrifos stores. This information will be vital in creating an effective PBPK model for pregnant women. For any PBPK model used in a comprehensive risk assessment, validation would add confidence to the predictions derived from its use.

Many of the studies discussed in Q 5.1 provide data on the concentration of chlorpyrifos in the media such as house dust, air and water while market basket data exists on the concentration of chlorpyrifos on food. These are the primary tools for generating an effective exposure assessment and a subsequent reconstruction of potential dose. Dose reconstruction can be used to evaluate the efficacy of the PBPK model since its prediction of excretion rates can potentially be validated with an accurate estimate of dose. This assessment of the PBPK model through reconstructed dose may bridge some of the data gaps in assessing risk by validating the PBPK model. A validated model allows for effective prediction of exposure at the critical windows of maximum effect (AChE suppression) with measured urine concentrations. More data exists on chlorpyrifos than other pesticides in the environment, and this may be the best opportunity for utilizing exposure data to evaluate a PBPK model. It is noted however that is there is a significant difference between the predicted urine or blood concentrations that both the PBPK modelers and those that produce the exposure estimate, will point to the other for using "bad" assumptions. In this case both models should be reevaluated for the assumptions used.

The effects mixtures of chlorpyrifos + Diazinon /chlorpyrifos + Propoxur or chlorpyrifos/Propoxur/Diazinon have not previously been considered. Like from all mixtures both constructive and destructive interference can occur. Questions will have to be addressed; do they affect each other's half lives and distributions and clearance through metabolic competition (Coughli et. al. 2012)? Are their net AChE effects additive or multiplicative? Do they share mechanistic pathways? To address these questions, the Panel recommends further studies described in 5.3 to improve estimates of effects when mixtures of xenobiotics are used compared to single agents. In particular PK parameters such as distribution, half-life and clearance/elimination can be altered if admixtures of chemicals interfere with the absorption or metabolism of another component of the admixture. Using currently available data, other than improving in

silico PK or PBPK approaches (again described above in 5.3), the Panel is not sure there is more that the Agency can do to reanalyze or transform the available data into more meaningful studies. However, any estimation of effect should have an additive dose effect as a minimum and perhaps greater protective factors until mechanistic studies can be done.

Other considerations

A further criticism is in the focus on 10% AChE activity reduction. While certainly a benchmark, the fact that no mechanism has been proposed that would tie such a reduction to any specific outcome begs the question; what is the role of the 10% reduction of AChE in predicting negative health outcomes. The Panel noted that to their knowledge there is no proposed mechanism whereby a 10% reduction in AChE activity in a pregnant woman, even at a specific point in pregnancy, is responsible for cognitive deficit or neurodevelopmental delay of the fetus? The current proposed mechanisms focus on correlation; the deficit in AChE in the mother is assumed to be associated with some other activity, e.g., transport of parent chlorpyrifos (or TCPy for that matter) across the placenta and the nascent blood-brain barrier in the developing fetus? Since no one knows whether this occurs, the utility of the measurement of maternal AChE reduction is unknown. AS is often the case, "more research is needed."

Some on the Panel feel that the 10% figure is merely a marker of some level of exposure. This level may differ in its impact depending on the association of the AChE inhibition with the parent pesticide concentration in the serum. If the Panel assumes that each OP produces exactly the same level of AChE inhibition on a, say, molar basis, does that imply that there is an identical effect of each? Focusing again on chlorpyrifos, is the parent, the oxon, or some other metabolite that is responsible for some of the effects seen in the Columbia study? Only with a better understanding of exposure to chlorpyrifos at various gestational ages will the Agency be able to determine what exposures are causing the effects. The mouse studies do not seem to help all that much.

***Question 6: Characterizing the range of potential risks.***

The 2009 NRC report, *Science and Decisions*, focused on improving the *technical analysis* through the development and use of scientific knowledge and information to promote more accurate characterizations of risk, and thus improving the *utility* of risk assessment for risk-management decisions. The NRC report also pointed out that regulatory risk assessment does not routinely approach public health and environmental problems by arraying a wide range of options for dealing with them. *In the case of chlorpyrifos, in light of the discussions of Questions 1-5, please provide guidance for assessing and presenting the range of plausible responses at given doses, and the effect of the overall uncertainty and variability around that range.*

**Response**
Part of the value of the framework for integrative analysis to evaluate multiple lines of evidence in the context of the AOP/MOA is to enable us to draw inference on the weight of evidence from the totality of the data.

In characterizing the range of plausible responses at given doses, it seems important to draw maximally on the dose-response data, beyond a single or several points of departure. For example, for risk management purposes, characterization to the extent possible, of the nature of potential risks above the reference dose would be informative. Presentation, then, of an array of points of departure for various endpoints for different types of effects bounded by their relative uncertainty, would more meaningfully characterize that value selected for the Reference Dose in the context of the range of effects reported in the broader database. It should also promote reliance on more certain rather than the most conservative data. As a minimum, it would be helpful in communicating the relative degree of protection provided by the selected point of departure.

As indicated, in response to previous questions, the maximal use of the available dose-response data from the epidemiological studies is recommended as a basis to at least, "bound" reference doses developed on the basis of points of departure from animal data. To the extent possible, this step should take into account at least the semi-quantitative bounding of the dose-response relationship from human studies based on the impact of identified uncertainties. This would perhaps clarify the basis for (the seeming) conclusion that the uncertainties associated with the exposure–response relationship in the epidemiological studies are greater than those associated with the POD derived on the basis of the animal data (*i.e.*, the basis for relying on the latter for dose-response analysis).

Similarly, options for dose-response analysis for acute effects should be considered separately from those based on long term exposures – *i.e.*, measures representing acute adverse neurological outcomes (ChE inhibition) commonly associated with occupational exposure versus those potentially related to long term exposure in the general population, such as neurobehavioral disorders. This separation would underscore the significant variation in the range of exposures in the population associated with these different types of effects, as reflected in reported TCPy levels measured in the three birth cohort studies and the recent occupational studies. Reconciliation of variability and uncertainty for these different options will likely require additional, focused study (see response to Question 4.3b).

It is further suggested that the Agency focus on the data of chlorpyrifos levels in the cord blood samples as the base to develop the POD for chronic exposures to chlorpyrifos based on a PBPK/PD model. This preliminary work would not only identify priorities for the acquisition of additional data, but would also reduce overall uncertainty and variability.

It seems important, also, to consider comparability within the Agency across compounds for which epidemiological data on neurotoxicity have served as the basis of points of departure – *e.g.*, mercury and lead. How does the weight of evidence from epidemiological studies for these compounds compare with that for chlorpyrifos? For example, in the background paper, it is stated that: "There are a number of known developmentally neurotoxic chemicals with well established relationships between exposure and neurological disorders in humans for which a definitive mode of action has not been established: for example, lead, methyl mercury and ethanol." While documentation of an MOA is, then, not a prerequisite for basing points of departure on human epidemiological data, the nature of the weight of evidence that distinguishes chlorpyrifos from these cases, as a basis for reliance on animal rather than epidemiological data to characterize the point of departure, is unclear.

In relation to the databases of studies which underlie considerations related to weight of evidence including consistency, specificity and biological plausibility, it would also be extremely helpful to have *a priori* criteria (to be presented initially) as the basis for evaluation of the individual studies on, for example, neurodevelopmental effects in animals and humans. While it is recognized that these criteria cannot be prescriptive, an upfront discussion of the factors taken into account in judging the adequacy of individual studies and hence, the weighting of their contribution within the weight of evidence, would be valuable.

The most susceptible lifestage(s), populations that would be expected to be more vulnerable to the effects of the chemical, and the effects of background exposures on these risks need also to be addressed in risk characterization.

Background exposures would include:

- exposures to other sources of chlorpyrifos
- exposures to other chemicals that affect
    - key steps thought to be involved in chlorpyrifos' neurodevelopmental adverse outcome pathway, *e.g.*, exposures to other chemicals competing with chlorpyrifos for metabolism by the same enzyme system (see response to Charge Question 5.4).
    - the same apical neurodevelopmental endpoints, *e.g.*, decrements in working memory or full scale IQ.

In assessing the range of plausible responses at a given dose of chlorpyrifos, variability in response within the population of concern should be taken into account, and uncertainties in estimation of the response should be addressed.

Sources of variability in response to chlorpyrifos include:

- differences in biological susceptibility, such as differences in lifestage, health and disease status, and genetics (*i.e.*, polymorphisms in Phase I and Phase II metabolism)
- differences in background exposures

Significant uncertainty in the draft chlorpyrifos risk assessment, arises from multiple sources. These include uncertainties:

- estimation of chlorpyrifos exposures in the children's cohort studies,
- whether protecting against AChE inhibition is protective against neurodevelopmental effects,
- whether the dose-response data for AChE inhibition in the most susceptible animal model, pregnant rats, can be used to derive a dose at which AChE inhibition would not occur in humans exposed prenatally,
- fully characterizing the neurodevelopmental effects of chlorpyrifos,
- translation of in vitro concentration–response relationships for neurotoxicity to *in vivo* dose-response relationships,
- dose-response for neurodevelopmental effects,
- identifying the neurodevelopmental adverse outcome pathway(s) of chlorpyrifos,

## References


Abou-Donia MB, Khan WA, Dechkovskaia AM, Goldstein LB, Bullman SL, Abdel-Rahman A (2006), In utero exposure to nicotine and chlorpyrifos alone, and in combination produces persistent sensorimotor deficits and Purkinje neuron loss in the cerebellum of adult offspring rats. *Arch Toxicol, 80,* 620-631.

Aldridge, J. E., Meyer, A., Seidler, F. J., & Slotkin, T. A. (2005). Alterations in central nervous system serotonergic and dopaminergic synaptic activity in adulthood after prenatal or neonatal chlorpyrifos exposure. *Environ Health Perspect, 113*(8), 1027-1031.

Aldridge JE, Meyer A, Seidler FJ, Slotkin TA (2005). Alterations in central nervous system serotonergic and dopaminergic synaptic activity in adulthood after prenatal or neonatal chlorpyrifos exposure. *Environ Health Perspect 113,* 1027-1031.

Aldridge, J. E., Seidler, F. J., & Slotkin, T. A. (2004). Developmental exposure to chlorpyrifos elicits sex-selective alterations of serotonergic synaptic function in adulthood: critical periods and regional selectivity for effects on the serotonin transporter, receptor subtypes, and cell signaling. *Environ. Health Perspect., 112*(2), 148-155.

Anderson, B.J. and Holford, N.H.G. (2009) Mechanistic basis of using body size and maturation to predict clearance in humans. *Drug Metab Pharmacokinet*, 24(1), 25-36.

Anderson, G.D. (2010) Developmental pharmacokinetics. *Semin Pediatr Neurol*, 17(4), 208-213.

Benirschke K, Kaufmann, P., Baergen, R.N. .*Pathology of the Human Placenta*. 5th ed: Springer; 2006.

Betancourt A.M., Carr R.L. (2004). The effect of chlorpyrifos and chlorpyrifos-oxon on brain cholinesterase, muscarinic receptor binding, and neurotrophin levels in rats following early postnatal exposure. *Toxicol Sci*. 2004 Jan;77(1), 63-71.

Berkowitz GS, Wetmur JG, Birman-Deych E, Obel J, Lapinski RH, Godbold JH, Holzman IR, Wolff MS. (2004). In utero pesticide exposure, maternal paraoxonase activity, and head circumference. *Environ. Health Perspect*.; 112 (3) 88-91.

Billauer-Haimovitch H, Slotkin TA, Dotan S, Langford R, Pinkas A, Yanai J (2009) Reversal of chlorpyrifos neurobehavioral teratogenicity in mice by nicotine administration and neural stem cell transplantation. *Behav. Brain Res.* 205, 499-504.

Blake MJ, Gaedigk A, Pearce RE, Bomgaars LR, Christensen ML, Stowe C, James LP, Wilson JT, Kearns GL, Leeder JS. (2007). Ontogeny of dextromethorphan O- and N-demethylation in the first year of life. *Clin. Pharmacol. Ther*. Apr;81(4):510-6.

Bouchard MF, Chevrier J, Harley KG, Kogut K, Vedar M, Calderon N, Trujillo C, Johnson C, Bradman A, Barr DB, Eskenazi B (2011). Prenatal exposure to organophosphate pesticides and IQ in 7-year-old children. *Environ. Health Perspect.* 119(8), 1189-1195.

Braquenier JB, Quertemont E, Tirelli E, Plumier JC (2010). Anxiety in adult female mice following perinatal exposure to chlorpyrifos. *Neurotoxicol Teratol*. 32, 234-239.

Brock A, 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, *Scand J Clin Lab Invest*. 50 (4), 401-8

Burns, C.J., D. Garabrant, J.W. Albers, S. Berent, B. Giordani, S. Haidar, R.Garrison, R.J. Richardson (2006). Chlorpyrifos Exposure and Biological Monitoring among Manufacturing Workers. *Occupational and Environmental Medicine,* 63(3) 218-220

Carr RL, Chambers HW, Guarisco JA, Richardson JR, Tang J, Chambers JE (2001). Effects of repeated oral postnatal exposure to chlorpyrifos on open-field behavior in juvenile rats. *Toxicol Sci.* 59, 260-267.

Castorina, R., Bradman, A., Fenster, L., Barr, D. B., Bravo, R., Vedar, M. G., *et al.* (2010). Comparison of current-use pesticide and other toxicant urinary metabolite levels among pregnant women in the CHAMACOS cohort and NHANES. *Environ Health Perspect, 118*(6), 856-863.

Centers for Disease Control (CDC). (2009). *Fourth National Report on Human Exposure to EnvironmentalChemicals.* CDC, Department of Health and Human Services.

Chen, T.H., Wangb, Y. H., & Wu, Y. H. (2011). Developmental exposures to ethanol or dimethylsulfoxide at low concentrations alter locomotor activity in larval zebrafish: implications
for behavioral toxicity bioassays. *Aquat Toxicol.* 102(3-4), 162-166

Cohen, A. Clifford, "Tables for Maximum Likelihood Estimates: Singly Truncated and Singly
Censored Samples", *Technometrics*, 3(4):535-541, 1961.

Cunnane, C. 1978. "Unbiased Plotting Positions--A Review," *J. Hydrol*. 37:205-222.

Crumpton, T. L., Seidler, F. J., & Slotkin, T. A. (2000). Is oxidative stress involved in the developmental neurotoxicity of chlorpyrifos? *Brain Res Dev Brain Res, 121*(2), 189-195.

Cunnane, C. (1978). "Unbiased Plotting Positions--A Review," *J. Hydrol*. 37, 205-222.

Dam K, Seidler FJ, Slotkin TA (2000). Chlorpyrifos exposure during a critical neonatal period elicits gender-selective deficits in the development of coordination skills and locomotor activity. Brain Res Dev Brain Res 121:179-187.

Das, K. P., & Barone, S., Jr. (1999). Neuronal differentiation in PC12 cells is inhibited by chlorpyrifos and its metabolites: is acetylcholinesterase inhibition the site of action? *Toxicol Appl Pharmacol, 160*(3), 217-230.

de Peyster, A., Willis, W. O., & Liebhaber, M. (1994). Cholinesterase activity in pregnant women and newborns. *J Toxicol Clin Toxicol*, 32(6), 683-696.

de Wildt SN, Kearns GL, Leeder JS, van den Anker JN.Cytochrome P450 3A: ontogeny and drug disposition.*Clin Pharmacokinet*. 1999 Dec;37(6):485-505.

Egeghy P., Cohen H., Tulve N., Melnyk L., Morgan, M., Fortmann, R., Sheldon L. (2011). Review of Pesticide Urinary Biomarker Measurements from Selected US EPA Children's Observational Exposure Studies *Int. J. Environ. Res. Public Health* 2011 (*8*), 1727-1754

Engel SM, Wetmur J, Chen J, Zhu C, Barr DB, Canfield RL, Wolff MS (2011). Prenatal exposure to organophosphates, paraoxonase 1, and cognitive development in childhood. Environ Health Perspect 119(8), 1182-1188.

Engel SM, Berkowitz GS, Barr DB, Teitelbaum SL, Siskind J, Meisel SJ, Wetmur JG, Wolff MS. (2007). Prenatal organophosphate metabolite and organochlorine levels and performance on the Brazelton Neonatal Behavioral Assessment Scale in a multiethnic pregnancy cohort. Amer J Epidemiol, 165, 1397-1404.

Engel SM *et al.* (2007). Prenatal organophosphate metabolite and organochlorine levels and performance on the Brazelton neonatal behavioral assessment scale in a multiethnic pregnancy cohort. *American Journal of Epidemiology*.; 165: 1397-1404.

Eskenazi B *et al.* (2007). Organophosphate pesticide exposure and neurodevelopment in young Mexican-American children. *Environmental Health Perspectives*; 115, 792-798.

Eskenazi B *et al.* (2010). PON1 and neurodevelopment in children from the CHAMACOS study exposed to organophosphate pesticides in utero. *Environmental Health Perspectives*.; 118: 1775-1781.

Eskenazi, B., Marks, A. R., Bradman, A., Harley, K., Barr, D. B., Johnson, C., *et al.* (2007). Organophosphate pesticide exposure and neurodevelopment in young Mexican-American children. *Environ Health Perspect, 115*(5), 792-798.

Fossum, E. N., Lisowski, M. J., Macey, T. A., Ingram, S. L., & Morgan, M. M. (2008). Microinjection of the vehicle dimethyl sulfoxide (DMSO) into the periaqueductal gray modulates morphine antinociception. *Brain Res*. 1204, 53-58.

Farahat, F. M., Ellison, C. A., Bonner, M. R., McGarrigle, B. P., Crane, A. L., Fenske, R. A., *et al.* (2011). Biomarkers of chlorpyrifos exposure and effect in Egyptian cotton field workers. Environ Health Perspect, 119(6), 801-806.

Giordano, G., Afsharinejad, Z., Guizzetti, M., Vitalone, A., Kavanagh, T. J., & Costa, L. G. (2007). Organophosphorus insecticides chlorpyrifos and diazinon and oxidative stress in neuronal cells in a genetic model of glutathione deficiency. *Toxicol Appl Pharmacol, 219*(2-3), 181-189.

Gurtovenko, A. A.& Anwar, J. (2007). Modulating the structure and properties of cell membranes: the molecular mechanism of action of dimethyl sulfoxide. *J Phys Chem B.* 111, 10453-10460.

Gage, J.C.: (1967).The Significance of Blood Cholinesterase Activity Measurements. *Residue Reviews* 18, 159-173

Grob, D., J.L Lilienthal, A.M. Harvey, and B.F. Jones (1947).The Administration of Di-opropyl fluorophosphate (DFP) in Man. *John Hopkins Hosp. Bull*. 81:217-292

Gage, J.C.: (1967).The Significance of Blood Cholinesterase Activity Measurements. *Residue Reviews* 18:159-173

Grob, D., and J.C. Harvey: (1958). Effects in Man of the Anticholinesterase Compound Sarin (Isopropyl Methyl Phosphonofluoridate). *J. Clinical Investigation*. 37, 350-368

Grigoryan H, Li B, Anderson EK, Xue W, Nachon F, Lockridge O, Schopfer LM. (2009). Covalent binding of the organophosphorus agent FP-biotin to tyrosine in eight proteins that have no active site serine. Chem Biol Interact. Aug 14;180(3):492-8.

Garabrant, D. H., Aylward, L. L., Berent, S., Chen, Q., Timchalk, C., Burns, C. J., *et al.* (2009). Cholinesterase inhibition in chlorpyrifos workers: Characterization of biomarkers of exposure and response in relation to urinary TCPy. *J Expo Sci Environ Epidemiol, 19*(7), 634-642.

Grob, D., J.L Lilienthal, A.M. Harvey, and B.F. Jones: (1947). The Administration of Di-isopropyl fluorophosphate (DFP) in Man. *John Hopkins Hosp. Bull*. 81:217-292

Grob, D., and J.C. Harvey: (1958). Effects in Man of the Anticholinesterase Compound Sarin (Isopropyl Methyl Phosphonofluoridate). *J. Clinical Investigation*. 37, 350-368

Harley KG, Huen K, Schall RA, Holland NT, Bradman A, Barr DB, Eskenazi B. (2011). Association of organophosphate pesticide exposure and paraoxonase with birth outcome in Mexican-American women. *PLoS One*; 6,1-10.

Haviland JA, Butz DE, Porter WP (2010). Long-term sex selective hormonal and behavior alterations in mice exposed to low doses of chlorpyrifos in utero. *Reprod Toxicol* 29, 74-79.

Hallare, A. V., Kohler, H. R., & Triebskorn, R. (2004). Developmental toxicity and stress protein responses in zebrafish embryos after exposure to diclofenac and its solvent, DMSO. *Chemosphere*. 56(7), 659-666.

Hallare, A., Nagel, K., Kohler, H. R, & Triebskorn, R. (2006). Comparative embryotoxicity and proteotoxicity of three carrier solvents to zebrafish (*Danio rerio*) embryos. *Ecotoxicol Environ Saf.* 63(3), 378-388.

Hanslick, J. L, Lau, K., Noguchi, K. K., Olney, J. W., Zorumski, C. F., Mennerick, S., Farber, N. B. (2009). Dimethyl sulfoxide (DMSO) produces widespread apoptosis in the developing central nervous system. *Neurobiol Dis*, 34(1), 1-10.

Haas C N, Scheff PA. (1990). Estimation of averages in truncated samples. *Environ Sci Technol* 24(6):912-919.

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

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

Helsel DR. (1990). Less than obvious: Statistical treatment of data below the detection limit. *Environ Sci Technol* 24(12) 1766-1774.

Hofmann, J.N., A. Carden, R.A. Fenske, H.E. Ruark, M.C. Keifer. (2008).Evaluation of a Clinic-based Cholinesterase Test Kit for the Washington State Cholinesterase Monitoring Program. *American Journal of Industrial Medicine*, 51(7) 532-538,

Howard, A. S., Bucelli, R., Jett, D. A., Bruun, D., Yang, D., & Lein, P. J. (2005). Chlorpyrifos exerts opposing effects on axonal and dendritic growth in primary neuronal cultures. *Toxicol Appl Pharmacol, 207*(2), 112-124.

Hines RN., (2007) Ontogeny of human hepatic cytochromes P450. *J Biochem Mol Toxicol*.;21(4):169-75.

Hines RN. Stevens JC, Marsh SA, Zaya MJ, Regina KJ, Divakaran K, Le M, (2008). Developmental changes in human liver CYP2D6 expression. *Drug Metab Dispos*. Aug;36(8):1587-93.

Hines RN. (2008). The ontogeny of drug metabolism enzymes and implications for adverse drug events. *Pharmacol Ther*. May; 118(2), 250-67.

Hines RN, .Koukouritaki SB, Manro JR, Marsh SA, Stevens JC, Rettie AE, McCarver DG (2004) Developmental expression of human hepatic CYP2C9 and CYP2C19. *J Pharmacol Exp Ther*. Mar;308(3):965-74.
Kliman, HJ, (1998). *The Umbilical Cord*. In: Ernst Knobil JN, (ed). The Encyclopedia of Reproduction. Elsevier Press; Maryland Heights, MO.

Jett DA, Navoa RV, Beckles RA, McLemore GL (2001). Cognitive function and cholinergic neurochemistry in weanling rats exposed to chlorpyrifos. *Toxicol Appl Pharmacol* 174:89-98.

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. Toxicol Sci. May;115(1):183-93.

Johnson FO, Chambers JE, Nail CA, Givaruangsawat S, Carr RL (2009). Developmental chlorpyrifos and methyl parathion exposure alters radial-arm maze performance in juvenile and adult rats. *Toxicol Sci* 109, 132-142.

Icenogle LM, Christopher NC, Blackwelder WP, Caldwell DP, Qiao D, Seidler FJ, Slotkin TA, Levin ED (2004). Behavioral alterations in adolescent and adult rats caused by a brief subtoxic exposure to chlorpyrifos during neurulation. *Neurotoxicol Teratol* 26, 95-101.

Laviola G, Adriani W, Gaudino C, Marino R, Keller F (2006) Paradoxical effects of prenatal acetylcholinesterase blockade on neuro-behavioral development and drug-induced stereotypies in reeler mutant mice. *Psychopharmacology* (Berl) 187, 331-344.

Levin ED, Addy N, Baruah A, Elias A, Christopher NC, Seidler FJ, Slotkin TA (2002). Prenatal chlorpyrifos exposure in rats causes persistent behavioral alterations. *Neurotoxicol Teratol*. 24, 733-741.

Levin ED, Addy N, Nakajima A, Christopher NC, Seidler FJ, Slotkin TA (2001). Persistent behavioral consequences of neonatal chlorpyrifos exposure in rats. *Brain Res Dev Brain Res* 130, 83-89.

Lu, C., K. Toepel, R. Irish, R.A. Fenske, D.B. Barr, and R. Bravo. (2006).Organic Diets Significantly Lower Children's Dietary Exposure to Organophosphorus Pesticides. *Environ Health Perspect* 114(2):260–263

Luck W, Nau H, Hansen R, Steldinger R. (1985). Extent of nicotine and cotinine transfer to the human fetus, placenta and amniotic fluid of smoking mothers. *Dev Pharmacol Ther*. 8(6), 384-95.

Lu, C., K. Toepel, R. Irish, R.A. Fenske, D.B. Barr, and R. Bravo. (2006) Organic Diets Significantly Lower Children's Dietary Exposure to Organophosphorus Pesticides. *Environ Health Perspect* 114(2):260–263.

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. *J Toxicol Environ Healt*h.;68:209–227.

Lovasi GS *et al.* Chlorpyrifos exposure and urban residential environment characteristics as determinants of early childhood neurodevelopment. *American Journal of Public Health*. (2010).; 101: 63-70.

Mattsson, J. L., Maurissen, J. P., Nolan, R. J., & Brzak, K. A. (2000). Lack of differential sensitivity to cholinesterase inhibition in fetuses and neonates compared to dams treated perinatally with chlorpyrifos. *Toxicol Sci, 53*(2), 438-446.

Middlemore-Risher, M. L., Adam, B. L., Lambert, N. A., & Terry, A. V., Jr. (2011). Effects of chlorpyrifos and chlorpyrifos-oxon on the dynamics and movement of mitochondria in rat cortical neurons. *J Pharmacol Exp Ther, 339*(2), 341-349.

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.

MacIntosh DL, Needham LL, Hammerstrom KA, and Ryan PB. (1999). A longitudinal investigation of selected pesticide metabolites in urine. J. Exp. Anal. Environ. Epidem. 9(5), 494-501.

Morgan, M. K., L. Sheldon, C. Croghan, P. A. Jones, G. L. Robertson, J. C. Chuang, N. K. Wilson, and C. Lyu. (2004).Exposure of preschool children to chlorpyrifos and its degradation product 3, 4, 6 Trichloro-2-pyridinol in their everyday environments *Journanl of exposure analysis and enviromental epidemiology* Nature Publishing Group, .London, Uk, , 1-13,

Maurissen JP, Hoberman AM, Garman RH, Hanley TR, Jr. (2000). Lack of selective developmental neurotoxicity in rat pups from dams treated by gavage with chlorpyrifos. *Toxicol Sci* 57, 250-263.

Marks AR *et al.* (2010). Organophosphate pesticide exposure and attention in young Mexican-American children: the CHAMACOS study. *Environmental Health Perspectives*.; 118: 1768-1774.

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, McCarver DG, Hines RN, Krishnan K. (2006). Modeling interchild differences in pharmacokinetics on the basis of subject-specific data on physiology and hepatic CYP2E1 levels: a case study with toluene. *Toxicol Appl Pharmacol*. Jul 1; 214(1):78-87.

Oliveira G.H., J.D. Henderson, B.W. Wilson. Cholinesterase Measurements with an Automated Kit. *Am J Ind Med Supp* 2:49-53, 2002.

Pope, C. N. C. J , and Chakraborti, T. K. (1995). Organophosphorus-sensitive cholinergic receptors. Possible role in modulation of antichohnesterasemduced toxicity. In *Enzymes of the Cholinesterase Family* (A S. Balasuberamian, B. T. Doctor, P. Taylor, and D. M. Quinn, Eds.), pp. 305-312. Plenum, New York.

Perkins, Jimmy L., Cutter, Gary N., and Cleveland, Michael S., "Estimating the Mean, Variance, and Confidence Limits from Censored (< Limit of Detection), Lognormally Distributed Exposure Data". *American Industrial Hygiene Association Journal*, 51(8):416-419, 1990.

Pearce RE, Gotschall RR, Kearns GL, Leeder JS. (2001). Cytochrome P450 Involvement in the biotransformation of cisapride and racemic norcisapride in vitro: differential activity of individual human CYP3A isoforms *Drug Metab Dispos*. Dec;29(12):1548-54.

Qiao, D., Seidler, F. J., Padilla, S., & Slotkin, T. A. (2002). Developmental neurotoxicity of chlorpyrifos: what is the vulnerable period? *Environ Health Perspect, 110*(11), 1097-1103.

Qiao, D., Seidler, F. J., & Slotkin, T. A. (2005). Oxidative mechanisms contributing to the developmental neurotoxicity of nicotine and chlorpyrifos. *Toxicol Appl Pharmacol, 206*(1), 17-26.

Rauh V, Arunajadai S, Horton M, Perera F, Hoepner L, Barr DB, Whyatt R (2011). Seven-year neurodevelopmental scores and prenatal exposure to chlorpyrifos, a common agricultural pesticide. *Environ. Health. Perspect.* 119(8):1196-1201.

Rauh, V. A., Garfinkel, R., Perera, F. P., Andrews, H. F., Hoepner, L., Barr, D. B., *et al.* (2006). Impact of prenatal chlorpyrifos exposure on neurodevelopment in the first 3 years of life among inner-city children. *Pediatrics, 118*(6), e1845-1859.

Ray, A., Liu, J., Ayoubi, P., & Pope, C. (2010). Dose-related gene expression changes in forebrain following acute, low-level chlorpyrifos exposure in neonatal rats. *Toxicol Appl Pharmacol, 248*(2), 144-155.

Ricceri L, Markina N, Valanzano A, Fortuna S, Cometa MF, Meneguz A, Calamandrei G (2003). Developmental exposure to chlorpyrifos alters reactivity to environmental and social cues in adolescent mice. *Toxicol Appl Pharmacol* 191:189-201.

Ricceri L, Venerosi A, Capone F, Cometa MF, Lorenzini P, Fortuna S, Calamandrei G (2006). Developmental neurotoxicity of organophosphorous pesticides: fetal and neonatal exposure to chlorpyrifos alters sex-specific behaviors at adulthood in mice. *Toxicol Sci* 93:105-113.

Rauh VA *et al.* (2006) Impact of prenatal chlorpyrifos exposure on neurodevelopment in the first 3 years of life among inner-city children. *Pediatrics.* ; 18 1835-1859.

Rauh VA *et al.* (2011) Seven-year neurodevelopmental scores and prenatal exposure to chlorpyrifos, a common agricultural pesticide. *Environmental Health Perspectives*.; 119: 1196-1201

Reigart, J.R. and J.R. Roberts: *Recognition and Management of Pesticide Poisonings*, 5th Ed., (EPA #735-R-98-003), U.S. Environmental Protection Agency, Washington DC (1999).

Rauh V, Arunajadai S, Horton M, Perera F, Hoepner L, Barr DB, Whyatt R. Seve-year neurodevelopmental scores and prenatal exposure to chlorpyrifos, a common agricultural pesticide. Environ Health Persp 2011;119:1196-1201.

Richardson, D.B. and A. Ciampi. (2003). Effects of Exposure Measurement Error When an Exposure Variable Is Constrained by a Lower Limit. *American Journal of Epidemiology*. 157(4):355-363

Sackerman, J., Donegan, J. J., Cunningham, C. S., Nguyen, N. N., Lawless, K., Long, A., Benno, R. H., & Gould, G. G. (2010). Zebrafish behavior in novel environments: Effects of acute exposure to anxiolytic compounds and choice of danio rerio line. *Int J Comp Psychol*, 23(1), 43-61.

Samantha A. Radford, Parinya Panuwet, Dana Boyd Barr, and P. Barry Ryan. *Insecticide Degradation in Food and Beverages: Implications for Risk Assessment.* Annual Meeting of the Southern Section of the Association of Analytical Chemists. Atlanta, GA. May 2012.

Saulsbury, M. D., Heyliger, S. O., Wang, K., & Johnson, D. J. (2009). Chlorpyrifos induces oxidative stress in oligodendrocyte progenitor cells. *Toxicology*, *259*(1-2), 1-9.

Schneider H Panigel M, Dancis J. Transfer across the perfused human placenta of antipyrine, sodium and leucine. *Am J Obstet Gynecol*. 1972; Nov 15;114(6):822-8.

Schenker S, Yang Y, Johnson RF, Downing JW, Schenken RS, Henderson GI, King TS. The transfer of cocaine and its metabolites across the term human placenta. *Clin Pharmacol Ther*. 1993; March:53(3):329-39.

Simone C, Derewlany L, Oskamp M, Knie B, Koren G. Transfer of cocaine and benzoylecgonine across the perfused human placental cotyledon. Am J Obstet Gynecol. 1994; May;170(5 Pt 1):1404-10.

Slotkin, T. A., MacKillop, E. A., Ryde, I. T., & Seidler, F. J. (2007). Ameliorating the developmental neurotoxicity of chlorpyrifos: a mechanisms-based approach in PC12 cells. *Environ Health Perspect, 115*(9), 1306-1313.
Terry, A. V., Jr., Gearhart, D. A., Beck, W. D., Jr., Truan, J. N., Middlemore, M. L., Williamson, L. N., *et al.* (2007). Chronic, intermittent exposure to chlorpyrifos in rats: protracted effects on axonal transport, neurotrophin receptors, cholinergic markers, and information processing. *J Pharmacol Exp Ther, 322*(3), 1117-1128.

Travis CC, Land, ML. (1990). Estimating the mean of data sets with nondetectable values. *Environ Sci Technol* 24(7):961-962.

Trundle D, Marcial G. (1988). Detection of cholinesterase inhibition. The significance of cholinesterase measurements. *Ann Clin Lab Sci*. 18(5):345-52

Turgeman G, Pinkas A, Slotkin TA, Tfilin M, Langford R, Yanai J (2011). Reversal of chlorpyrifos neurobehavioral teratogenicity in mice by allographic transplantation of adult subventricular zone-derived neural stem cells. *J Neurosci Res* 89:1185-1193.

Turner, C., Sawle, A., Fenske, M., & Cossins, A. (2012). Implications of the solvent vehicles dimethylformamide and dimethylsulfoxide for establishing transcriptomic endpoints in the zebrafish embryo toxicity test. *Environ Toxicol Chem.* 31(3), 593-604.

Venerosi A, Calamandrei G, Ricceri L (2006). A social recognition test for female mice reveals behavioral effects of developmental chlorpyrifos exposure. *Neurotoxicol Teratol* 28:466-471.

Venerosi A, Cutuli D, Colonnello V, Cardona D, Ricceri L, Calamandrei G (2008). Neonatal exposure to chlorpyrifos affects maternal responses and maternal aggression of female mice in adulthood. *Neurotoxicol Teratol* 30:468-474.

Venerosi A, Ricceri L, Rungi A, Sanghez V, Calamandrei G (2010). Gestational exposure to the organophosphate chlorpyrifos alters social-emotional behaviour and impairs responsiveness to the serotonin transporter inhibitor fluvoxamine in mice. *Psychopharmacology (Berl)* 208, 99-107.

Whyatt, R. M., Garfinkel, R., Hoepner, L. A., Andrews, H., Holmes, D., Williams, M. K., *et al.* (2009). A biomarker validation study of prenatal chlorpyrifos exposure within an inner-city cohort during pregnancy. *Environ Health Perspect, 117*(4), 559-567.

Whyatt, R. M., Garfinkel, R., Hoepner, L. A., Holmes, D., Borjas, M., Williams, M. K., *et al.* (2007). Within- and between-home variability in indoor-air insecticide levels during

pregnancy among an inner-city cohort from New York City. *Environ Health Perspect, 115*(3), 383-389.

Whyatt RM *et al.* (2004). Prenatal insecticide exposures and birth weight and length among an urban minority cohort. Environmental Health Perspectives.; 112: 1125-1132.

Wang, Y., Moss, J. and Thisted, R. (1992) Predictors of body surface area. *J Clin Anesth*, **4**(1):4-10.

Whyatt RM, Garfinkel R, Hoepner LA, Andrews H, Holmes D, Williams MK, Reyes A, Diaz D, Perera FP, Camann DE, Barr DB. (2009). A biomarker validation study of prenatal **CPF** exposure within an inner-city cohort during pregnancy. *Environ Health Perspect*. Apr;117(4):559-67. Epub 2008 Dec 5.

Yang, D., Howard, A., Bruun, D., Ajua-Alemanj, M., Pickart, C., & Lein, P. J. (2008). Chlorpyrifos and chlorpyrifos-oxon inhibit axonal growth by interfering with the morphogenic activity of acetylcholinesterase. *Toxicol Appl Pharmacol, 228*(1), 32-41.

Yager, J., H. McLean, M. Hudes, and R. Spear: (1976). J. Occup. Med. 18(4):242-244. Components of Variability in Blood Cholinesterase Assay Results.

.

Young JG *et al.* Association between in utero organophosphate pesticide exposure and abnormal reflexes in neonates. NeruoToxicology. (2005); 26: 199-209.

## Appendixes

### *Appendix A to Question 3.1- Evaluation of Individual Studies*

The strengths and weaknesses of the studies are outlined as are the principal findings and study designs that are also summarized in a different way in the EPA Draft Issue Paper in Appendix 3. In general, the studies fall into two major groups: (a) those that have adequate group sizes controlled for litter effects by sampling either only 1 offspring/sex/litter or over-sampled only slightly by testing 2 offspring/sex/litter and analyzed the data using appropriate statistical models, *e.g.*, ANOVA with factors of group and sex, optimally with litter as a block factor in a randomized block design in which treatment and sex are fixed effect factors within blocks and with within subject factor(s) for trial, day, or interval where the same subjects is assessed on the same parameter repeatedly; and (b) those that failed to control for litter effects, tested too few litters and/or offspring, were under-powered and prone to Type I error, and therefore should be given less weight.

Some studies included appropriate down-stream statistical analyses and some did not. In general, most studies used factorial ANOVA or MANOVA models, but the follow-up methods used varied significantly. Most ANOVA models should be further analyzed by some method that deconstructs interactions and all experiments should include a-posteriori group comparison methods than control for multiple comparisons while holding alpha-constant. There was much use of the Fisher Least Significant Difference (LSD) test for post hoc group comparisons among the above reviewed studies, which is only appropriate when there are not more than three groups. In a number of studies there were only three groups therefore this method is acceptable in those cases. However, there are a number of studies where there were more than three groups and the LSD or Protected Least Significant Difference (PLSD) tests were used despite their drawbacks. The use of Tukey tests in the absence of significant F-tests was also reported in several studies. This approach requires clearer justification and reporting of which results used Tukey T-Tests irrespective of a significant F-test and which were used after a significant F-test as a follow-up in experiments where both approaches are used.

The Radial-arm maze (RAM) has been used extensively in the experiments reviewed. A significant concern is that none of those reviewed above controlled for the response pattern known as chaining, *i.e.*, where the animal learns a strategy such as entering each adjacent arm successively or approximately successively by, for example, learning to always turn right or always turn left. When this occurs, working memory is not assessed. What is assessed is more likely habit formation, which is learning, but more rudimentary than working memory which is closely tied to higher cognitive functions such as attention and executive functions and hence assesses higher order processes. It is difficult, therefore, in all the RAM data in this group of reviewed experiments to determine what the chlorpyrifos-related effects were measuring. There is no doubt that chlorpyrifos has effects on RAM performance, but they may not be working memory effects unless it can be shown that more rudimentary forms of learning were not utilized. As for the reference memory effects, these are more likely to be as they appear but even

here one cannot be sure without better test procedures that have an inter-trial interval delay before each new trial to ensure that animals are not entering non-baited arms because the cost of doing so does not outweigh the overall retrieval of the rewards regardless of the small cost of running down an empty arm. One experiment used spatial delayed alternation in a T-maze (Maurissen *et al.*, 2000). This experiment appropriately imposed a delay between the sample and test trials, which is the appropriate way to test working memory. This paper also showed an appropriate short-term memory decay function that was dependent on the length of the delay interval. This is a valuable internal control to prove that working memory was assessed. Some delay between arm choices in the RAM is similarly needed, even if a full decay function is not demonstrated. This would ensure that trial-dependent memory is being assessed rather than some other strategy. Not one experiment that used the RAM in this group of experiments imposed this basic requirement rendering interpretations difficult at best.

There were many other test method issues among these studies that raise further concerns, including that methods were used that have no known neurotoxicological significance. The functional significance of increases or decreases in time in open arms of the elevated-plus maze is unclear as a toxicological end point. Which change represents an adverse outcome: an increase or a decrease? Or is any change from control regarded as an adverse effect? What is the meaning of greater or lesser social investigation of a stranger mouse in neurotoxicological terms? What does it mean that a female mouse has a more upright posture when presented with a male intruder? Is it more adverse that she stands more or stands less in such a defensive posture? With no validation as to the neurotoxicological significance of anxiety tests, depression tests, or social interactions, such outcomes should be regarded as exploratory, and hypothesis-generating, rather than evidence of toxicity.

The bidirectional neurodevelopmental changes that were found in many of the studies presents challenges to interpretation. These effects included a slower habituation trend in female animals in one study, but not another study with the same doses also administered during a prenatal period, and a few transient effects in some of the cognitive tests that were also observed to occur in opposite directions. The lack of specificity of the direction of the neurobehavioral findings is problematic. In some cases isolated markers of certain behaviors were determined to be statistically significant, but these findings were sometimes not supported by other studies reporting no effects (or effects in the other direction) in similar dose ranges using similar routes of exposure. It is difficult to reconcile effects that are bidirectional in given domains of neurodevelopment, or inconsistent across doses and sex, and in the absence of specific hypotheses. In addition, the lack of dose-response among test outcomes (especially those that took doses below 1 mg/kg/d such as Braquenier, *et al.*, 2010, that only found positive responses in the middle dose group) is not reassuring.

## **Prenatal Studies**

This study (Abou-Donia *et al.*, 2006) was a prenatal study of chlorpyrifos. They used timed-pregnant Sprague-Dawley rats from Charles River (evidence of conception = E1).

The authors' indicate that dams were treated transdermally on E4-20 with chlorpyrifos in a 70% ethanol vehicle or were given ethanol vehicle alone. Dams were assigned to groups as follows: 5 mg/kg of chlorpyrifos, 5 mg/kg of nicotine, 5 mg/kg of chlorpyrifos + 5 mg/kg of nicotine, or saline. It is stated that 2 M/2 F per litter were sampled for testing. No statement of whether litter was included in the statistical model was provided; therefore, presumably it was not; hence there was no control for litter effects. Pairwise comparisons were performed using Fisher's LSD test, but they had 4 groups and this test is not appropriate for >3 groups. No litter culling to standardize litter size was done; hence, postnatal rates of growth between litters was not equalized. On P90, rats were tested for Beam walking, inclined plane, and forelimb hang time. Results: They report no CFP effects on beam walking; a female-only effect on the inclined plane (females slipped at lower angles than controls as the plane was tilted). Hang time: Both sexes in the chlorpyrifos group had shorter hang times than controls. Strengths: Used a transdermal route of exposure; they exposed animals from shortly prior to implantation to near-term. Weaknesses: Only one dose of chlorpyrifos was used (estimated to be equivalent to ~1 mg/kg/day). Used 2-way ANOVA but in the results they give no F-values; they provide no indication if reported effects were main effects or interactions; and they moved from ANOVA to LSD tests with no sorting of interactions (although which reported effects were from interactions and which main effects is unclear). They slightly oversampled per litter, but the most significant weaknesses are the small group sizes that results in an under-powered design and they did not analyze the data by litter. Additionally, they used timed-pregnant dams from the supplier for use in a prenatal study for which treatment started very early on E4. To do this, they would have had to have purchased rats that were plug-positive (without confirmation) and received them within 1-2 days of mating, leaving no more than 1-2 days to acclimate to their vivarium before treatment began.

This study (Icenogle *et al.*, 2004) used timed-pregnant SD Charles River dams for a prenatal study, again raising concerns about shipping stress. They dated pregnancy as counting evidence of conception = E1. They treated on E9-12 (the rationale for these days was not given, but seems to be a very narrow period of exposure). They randomized among dams on P1 and then again "every several days" thereafter, introducing unknown stressors in the experiment for both the dams and the offspring. They culled to 10 pups per litter then selected no more than 1 M and 1 F per (artificial) litter and assigned 10 litters to each treatment group. Groups: 0, 1, 5 mg/kg/day chlorpyrifos administered s.c in DMSO on E9-12. They tested offspring for T-maze spontaneous alternation. The T-maze was elevated with 1.5 cm curbs to prevent falling off the edge, and animals were given 5 trials/day with 30 seconds post-choice confinement. The test was given for 5 successive days. The apparatus is non-standard in the field and is designed more akin to the elevated plus maze (EPM), which is a test designed to induce anxiety so that it can be measured. The T-maze for spontaneous alternation is not intended to induce anxiety and for that reason is normally an enclosed maze. The elevated T-maze used in this study probably tests alternation AND anxiety, but one cannot determine how the resulting measurements can be attributed to memory versus anxiety. They also tested rats in a figure-eight locomotor activity monitor for 1 h with data recorded in 5 min intervals and repeated the test three times each spaced one week apart and in a 16-arm radial-arm maze

(RAM) with 12 arms baited daily, 4 arms never baited; the test sessions were 10 min or until 12 baited arms had been entered; they tested twice per week for 18 sessions; then gave a scopolamine (muscarinic antagonist) challenge at doses of 0.04, 0.08, 0.16 mg/kg or to separate animals a drug challenge of mecamylamine (nicotinic antagonist) at doses of 1.25, 2.5, 5.0 mg/kg. They also tested acoustic startle response with prepulse inhibition (ASR/PPI). They first conducted ASR-only trials and later intermixed ASR and PPI trials. Lastly, they tested animals in the elevated plus maze (EPM). For this they used a standard method (standard size apparatus for the typical 5 min. test). Note: this test was given AFTER all preceding tests whereas in most labs it is given before other tests based on the fact that this test is sensitive to prior experience. The statistical approach was MANOVA; interactions were further analyzed and pairwise comparisons were made by Fisher's PLSD. Findings: Spontaneous alternation: chlorpyrifos decreased shortened choice latency on early trials; no effects on alternation but failed to show whether they got alternation rates typical of this test to establish validity in their laboratory. Figure-eight test: They obtained a significant treatment x interval interaction. The chlorpyrifos 5 mg/kg showed faster habituation on two out of four of the last 5-min intervals, with one interval with higher activity in the chlorpyrifos 1.25 mg/kg group; they also note that the linear trend in this analysis was significant for treatment group, but it is noteworthy that the effects observed on this test were very small even if significant. RAM findings: The data were blocked into three sessions per block for analysis; hence, there were six blocks for the repeated measure factor. They found increased reference memory errors in block 1 and increased working memory errors in blocks 1 and 3 in the CFP 5 mg/kg group only with males and females combined. They found no effects on RAM performance after mecamylamine challenge. They found an effect of scopolamine challenge that was complex: Scopolamine increased errors with increasing scopolamine dose in controls but in the chlorpyrifos 1 mg/kg group it increased baited arm errors at lower doses more than in controls but less than in controls at the highest scopolamine dose. For the chlorpyrifos 5 mg/kg group scopolamine increased errors more than in controls after saline but less than in controls after scopolamine. They report no ASR/PPI effects and no EPM effects for time in the open (the principal index of anxiety in this test); the CFP 5 mg/kg group crossed center more than controls, however, suggesting a slight increase in activity. Strengths: Groups sizes were adequate and the data were analyzed by (artificial) litter. The factorial MANOVA models were appropriate. They tested two doses of chlorpyrifos and used many standard methods. Weaknesses: Used timed-prenatal females for a prenatal study thereby introducing prenatal shipping stress. RAM: they ran trials continuously with no intertrial interval (ITI) delay to ensure that working memory was being assessed. While there is no doubt they obtained RAM effects, it cannot be distinguished as to whether these were working memory or habit learning effects. The small reference memory effect they obtained is more likely to be a real reference memory change, but this cannot be certain without observation of the animals' performance or a method to ensure an ITI delay. In this test, a control for chaining (moving sequentially around the maze from one arm to the next or every other arm) or for random selection is important. Without a confinement period in the center between each arm choice, there is no way to rule out that rats obtained the food without relying upon working memory instead of another strategy. For reference memory findings, what one needs to determine is whether rats found the cost of entering even

empty arms insufficiently aversive so that is was more effective to check them all versus remembering which ones were unbaited. There is a basis for this concern in the data. Reference memory errors among controls improved from about 6.5 to about 5 across the 18 sessions. Given that there are 4 unbaited arms, the data suggest that even controls never actually acquired memory for the unbaited arms. If they had, one would expect well under 4 reference memory errors per trial block, assuming they blocked the data by averages rather than by sums; but the paper is moot on how the blocked data were formed. For these reasons, center confinement between trials is an important control especially for the working memory assessments. The authors would have greatly strengthened their experiment by either using center confinement between trials or having an observer map the problem solving strategy during testing to see that the animals were using alternate strategies. The chlorpyrifos vehicle used was DMSO, which was raised as a significant concern in response to another charge question.

In this experiment (Billauer-Haimovitch *et al.*, 2009), heterozygous HB/Igb mice were used and bred in-house (conception = E1). Mice were treated on E9-18 with chlorpyrifos at doses of 1, 3, 5, 10, 20 mg/kg given s.c in DMSO. Half the litters were fostered, half were not; they report no differences in outcome in preliminary analyses and state that the data were therefore pooled for subsequent analyses. They sampled 1 M/1 F for testing per litter. At P75, mice were tested in the Morris water maze (MWM). The pool was 87 cm diameter, the platform 8x10 which is a search ratio of 74:1 (Note: this is less than optimal for a test of spatial navigation in mice). They ran 2 blocks of 4 trials per day for 4 days. After completing hidden platform trials, they ran cued trials with the platform made visible above the water line. Data were analyzed by MANOVA, log transformed, with Tukey *a-posteriori* tests. They report no differential effects between males and females, so they combined sexes for presentation (but in the statistical analyses ). They report that in the chlorpyrifos20 group all animals died. They found a chlorpyrifos main effect on MWM latency, which they report was significant for the chlorpyrifos 1 and chlorpyrifos 3 mg/kg groups but not for the chlorpyrifos 5 or 10 mg/kg groups. In a second experiment they used only the chlorpyrifos 3 mg/kg group and found somewhat larger MWM latency effects and these were reversed by nicotine treatment prior to each daily test session. In a third experiment with chlorpyrifos 3 mg/kg, they again found MWM latency effects and these were reversed by cell implantation from neonatal cells grown in neurospheres. Found no speed differences in MWM on trials where latency was significantly increased. They found no chlorpyrifos effect on developmental reflexes (surface righting, startle emergence, age of fur appearance, day of pinna unfolding, or day of eye opening). Strengths: They used the MWM which is one of the most well-validated spatial learning/reference memory tests in neuroscience; they used 14-34 litters per group. A major strength was that they replicated the MWM effect in three separate experiments. Weakness: It is unclear why effects were seen in the MWM at chlorpyrifos 1 and 3 mg/kg but not at 5 and 10 mg/kg. The vehicle used was DMSO.

In this experiment (Turgeman *et al.*, 2011), heterozygous HB/Igb mice were used and bred in-house; Conception = E1 and mice were treated on E9-18. The exposure was CFP 3 mg/kg given s.c in DMSO. Half the litters were fostered, half not; they reported no differences in outcome between fostered and non-fostered litters; therefore, they pooled

data across this factor for later analyses. They tested 1 M/1 F per litter. At P75, mice were tested in a MWM (with the same 87 cm diameter, platform 8x10, and 74:1 search area as in the study by Billauer-Haimovitch *et al.* (2009) and has raised the same concern as above). Testing was in two blocks of four trials/day for four days. After hidden platform trials they ran cued trials. Data were analyzed by factorial MANOVA, log transformed, with Tukey *a-posteriori* tests. They report no M/F differential effects, so they combined sexes for presentation (but not statistically, they note). They report that chlorpyrifos 3 mg/kg increased MWM latency which was reversed by stem cell transplantation, just as in the previous experiment by Billauer-Haimovitch *et al.* (2009) (the stem cell methods were also identical). They found no effect on swim speed, but this experiment included no cued trials to ensure the absence of proximal cue learning problems. They also ran no probe trial to assess reference memory. Strengths: Found clear MWM effects, used 13-25 litters per group, they tested only 1 sex per litter. Weaknesses: Tested only one dose of chlorpyrifos, used a small maze, and did not conduct a test of reference memory or include cued trials as a control or include a reversal component to verify that the effects were hippocampally-dependent. The vehicle used was DMSO.

This study (Laviola *et al.*, 2006) used heterozygous Reeler KO mice on a C57BL/6 background. The mice were bred in-house by het x het crosses and conception = E0. chlorpyrifos-oxon (chlorpyrifos-O) was tested, not chlorpyrifos. chlorpyrifos-O was given on E14-16 (the rationale for this embryonic period was not given but seems very narrow for a compound whose exposure would be expected to be chronic). The compound was delivered by implanted osmotic minipump at a rate equaling 5 mg/kg/d. Offspring were tested on P3, 7, 11 for ultrasonic vocalizations, wire mesh hang time when the mesh was gradually rotated 180 degrees until animal was hanging upside down, and surface righting. As adults (>P70), they tested locomotor activity (30x30 cm apparatus) for 45 min., then administered 2 mg/kg scopolamine and retested for another 45 min.; 1 week later they re-tested the same animals in another apparatus (40x30 cm) for 10 min then removed them and gave a high dose (10 mg/kg) of amphetamine and retested for 50 min.; movements were video recorded and scored later using the Noldus Observer system to rate specific behaviors. Within litters, they had 3 genotypes (KO, het, WT). Treatment and genotype were regarded as between subject factors and other factors as within-subject factors. They used factorial ANOVA with pairwise comparisons by Tukey tests. On P7, the lowest ultrasonic calls were found in KO mice; they were intermediate in hets, and highest in WT but no such pattern was found in chlorpyrifos-O-treated mice; in chlorpyrifos-treated mice they observed high call frequencies in all genotypes. For angle to fall on the wire screen grasping task, they found no genotype effect in controls but fall angle was higher in Reeler chlorpyrifos-O exposed groups and it decreased from WT to het to KO, but even KOs were better than Control KO mice. They reported a chlorpyrifos-O effect on righting only in KO mice on P 7 and 11 (longer latency) but this effect was seen only on the worst of the three daily test trials; no effect was seen on the best or intermediate test trial given each day. In the 30x30 cm activity chambers, they found control KO mice were hyperactive, an effect which according to the text chlorpyrifos-O exposure 'normalized' but they did not show these data. After they gave scopolamine, all groups showed the expected hyperactivity response, but for all

genotypes a main effect was seen of greater hyperactivity in the chlorpyrifos-O groups but only when the 45 min. test session was subdivided such that the difference occurred only in the first 25 min. of the test session and after subtraction of pre-drug activity data (in an effort to adjust for the innate hyperactivity of the KO mice). In the 30x40 cm test chamber, no differences during the 10 min habituation prior to amphetamine challenge were seen. Post-amphetamine, KO chlorpyrifos-O treated mice showed increased activity that was greater than in KO controls, but for stereotypy, the opposite occurred, *i.e*, chlorpyrifos-O exposed mice were less stereotypic than KO controls; similar but weaker trends were seen in WT and hets. Strengths: This is the only study among the neurobehavioral papers to test the oxon. Weaknesses: N's were given for progeny but not by litter; the number of litters used was vague: They state that 12 breeding pairs were used, so presumably 6 litters were treated with chlorpyrifos-O and 6 with the vehicle, but the Panel was given no information on how the genotypes were distributed among the litters within treatment groups. There does not appear to be any control for litter effects. Sex was not described as a factor in ANOVAs. Only one dose was tested. The study was under-powered given only six litters per group with each litter subdivided by genotype and sex. Given the expected Mendelian ratios, it would be expected that half of the offspring would be hets, leaving 25% as KO and 25% as WT. Given that C57BL mice typically have 6-8 pups per litter, the sample sizes per genotype per sex per litter would be quite small and not likely to be evenly distributed.

In this study (Venerosi *et al.*, 2010), they tested the effects of chlorpyrifos 6 mg/kg given by gavage on E14-17 to CD-1 mice (bred in-house), but the date of inferred conception was not provided. They treated 18 litters with oil and 16 litters with chlorpyrifos. Shortly after birth, they culled to 4M/4F. Testing began on ~ P90. They evaluated the offspring in a Light/Dark (L/D) test for 5 min and also did observer scoring. They tested the mice in the Forced swim test (FST) either with an injection of saline or of the selective serotonin reuptake inhibitor (SSRI) fluvoxamine (30 mg/kg) 30 minutesebefore the test. They scored immobility, struggling, and swimming and testing was during the dark cycle. Females were grown to adulthood and bred, and on P8, a stranger male intruder was put in their cage (pups were removed) and scored for aggression. Data were analyzed by ANOVA or Mann-Whitney using litter as a factor; ANOVAs were followed by Tukey tests but in some cases in the absence of a significant F-test. They do not define the reason for this but cite a reference that it may be used in the absence of a significant F-test. They also state that they used non-parametric tests in some cases including analyzing interactions by Mann-Whitney U tests. It is unclear how it is used to detect interactions. They report finding a significant U-test for females for time in the tunnel connecting the two sides in the L/D test, with chlorpyrifos-exposed females spending more time in the tunnel than controls. For the FST, they report an interaction between chlorpyrifos x fluvoxamine in which chlorpyrifos eliminated the increased swimming induced in controls given fluvoxamine; this pattern was repeated for immobility time, *i.e*, fluvoxamine reduces immobility in controls, but this effect was dampened in chlorpyrifos-treated animals enough that the change in this group was not significant. In the intruder test, they also reported a chlorpyrifos x fluvoxamine interaction. They found that fluvoxamine reduced the duration of attack and increased inactivity in controls, but this pattern was largely eliminated in the chlorpyrifos exposed

group. Strengths: chlorpyrifos was given in oil, not DMSO. They used litter in the analysis and tested 6-13 offspring per group. Weaknesses: They had only one dose of chlorpyrifos; they do not say which outcomes they assessed by direct Tukey tests and which by ANOVA followed by Tukey comparisons. The meaning of the intruder test outside of basic research or as an index of neurotoxicity is not known.

In this study (Haviland *et al.*, 2010), they used 0, 1, 5 mg/kg chlorpyrifos given s.c. in DMSO to Swiss-Webster on E17-20 where conception was termed E0.5 (but after stating this, they never used half days again, so it is unclear of E0.5 was rounded to E0 or E1). Mice were bred in-house and litters were culled to 8 and balanced for sex. Testing began on P60 using a modified 8-arm RAM with a T at the end of each arm with food always in the left arm of the T (they name it the Foraging maze) and they compared animals tested in this maze to a group tested in a standard 8-arm RAM. They tested animals for 3 sessions per week for 3 weeks for a total of 9 sessions. For some reason that they do not explain, in the RAM they baited 6 of the 8 arms on each trial, but for the Foraging maze they baited 4 of the 8 arms on each trial. Strengths: Testing two groups in two mazes that presumably assess the same functions is a strength (potentially). They analyzed the data by litter, but the results are not well described. They state that RAM data in standard terms of errors by type (entries into baited versus unbaited arms), but they did not use an ITI delay to ensure that working memory was being assessed. Data were analyzed by ANOVA, but no mention was given of how pairwise comparisons were done, but it was probably by LSD since they used the LSD for pairwise comparisons for their thyroid assay data. They report finding only reference memory errors in the RAM, but all groups showed poor learning; they do not show working memory data at all and the reference memory differences are scattered across the test sessions and between doses non-systematically. In the foraging maze, they report their findings differently. It is noteworthy that they state that the entrance of each of the baited arms were marked with a 0.5 cm radius peg (or 1 cm in diameter) (Note: this is an extraordinary procedure; the provision of an obvious cue to which arms are baited and which arms is not, defeats the purpose of the test which is to remember which arms are baited and which are not; cuing the rat at the entrance to each arm provides evidence only that the rat can recognize the cue, requiring little memory). They report the data as the proportion of correct choices out of total choices, such that perfect performance results in a score of 1.0. Clear improvement toward 1.0 was seen across test sessions and chlorpyrifos caused a slower increase toward 1.0 at both doses in females but not in males. In males, chlorpyrifos exposure cased a more rapid increase toward 1.0 at 1 mg/kg but not at 5 mg/kg. Strengths: Use of two learning tests and the effort to develop an improved version of the RAM are noteworthy. Weaknesses: The vehicle used was DMSO. Poor learning in the RAM, performance on the Foraging maze not shown in terms of working and reference memory errors, and the use of an in-maze cue to which arms were baited seems to defeat the purpose of the test; findings were not dose-dependent or consistent across sessions. No ITI delay was used. It is not clear what the foraging maze adds to a more standard RAM. Why would a mouse learn better if it goes down a cued arm and turns left rather than going down an unmarked arm to find food when the arm is straight?

This study (Levin *et al.*, 2002) used timed pregnant SD Zivic-Miller (ZM) rats; date of inferred conception was not stated. They tested 0, 1, 5 mg/kg chlorpyrifos given s.c. in DMSO administered on E17-20. Litters were culled to 10 and randomized to dams. The rationale for randomizing the pups is not stated, nor is there any published data this reviewer is aware of that establishes that this cancels out litter effects although it clearly randomizes within litter genetic factors, but whether that improves or confounds outcomes is unknown. Moreover, the degree of stress this induces in the pups or the dams is unknown. Testing was done during the dark cycle. Litters were culled to 10 M/10 F per group per artificial litter. Spontaneous alternation used an elevated T-maze with no walls (see above) and was tested on P28-42, 5 trials per session (with confinement) for 5 daily sessions. They tested rats in a figure-eight locomotor system on P28-42 for 1 h/session, and did this three times each spaced one week apart; also used a 16-arm RAM with testing on P57-91. The testing was for 3 days/week for a total of 18 sessions. Their procedure was to use 4 unbaited arms and 12 baited arms. Each session lasted up to 10 min or until all baits were taken. They continued RAM testing after this on P98-119 with either a scopolamine challenge (muscarinic antagonist) at doses of 0.04, 0.08, 0.16 or mecamylamine (nicotinic antagonist) challenge at doses of 1.25, 2.5, 5.0 mg/kg given prior to each test session. Data were analyzed by ANOVA with factors of treatment and sex as between factors and test interval or day as within-subject factors. Main effects were taken as significant if they occurred at $P<0.05$ but interactions were followed-up if $P<0.10$. Pairwise comparisons were by Fisher's PLSD. They found no effects on spontaneous alternation frequency but report finding that chlorpyrifos decreased latencies to choose one of the arms (Note: latency to choose is not a typical outcome measure in this test). For the figure-eight test, they found no effects on the omnibus ANOVA, but then analyzed the first session and report finding a treatment x sex interaction in females in which they habituated slower than both chlorpyrifos groups. They used a trend analysis in which the linear trend was different in both chlorpyrifos groups compared with controls, but the effect was small and linearity did not appear to be a good fit to the data. In the RAM, they report a significant main effect of treatment and a treatment x error type and a treatment x error type x sex interaction. They followed these up and found female working memory effects at chlorpyrifos 1 mg/kg but not at chlorpyrifos 5 mg/kg. They report the same pattern for reference memory errors but less pronounced. After scopolamine, it was again found that female chlorpyrifos1 animals showed differential effects on RAM on working memory errors (not on reference memory errors or latency). For females, they found the slope of improvement across trials was lower in the chlorpyrifos1 group than for controls or the chlorpyrifos5 group for total errors using a $P<0.07$ trend to justify the trend analysis. When only working memory errors were analyzed, the female chlorpyrifos 1 effect was not seen. No effects were seen in males. For reference memory errors, there was again a significant linear trend for females, but the effect was that the CFP 1 mg/kg females made fewer errors than controls but this was because scopolamine caused controls to make more errors whereas it did not cause this in the chlorpyrifos 1 mg/kg females, which might make sense except that a higher dose of scopolamine did not cause controls to make more errors, making these data largely uninterpretable. Strengths: Sample size was adequate and they used litter by selecting only 1 male and 1 female per (artificial) litter. They used two doses of chlorpyrifos. Weaknesses: They did not use an ITI during RAM testing to

ensure that working memory was being assessed. They use an unorthodox elevated T-maze to assess spontaneous alternation. They used ZM rats. They used timed-pregnant rats for a prenatal study introducing a potential prenatal stressor. They provide no rationale for the short E17-20 exposure. Given that environmental exposure to chlorpyrifos might be chronic this short exposure window does not appear to be a very appropriate choice.

In this experiment (Ricceri *et al.*, 2006), CD-1 mice were gavaged with chlorpyrifos at doses of 0, 3, 6 mg/kg/d in peanut oil on E15-18 (in-house breeding with date of inferred conception = E0). At birth, litters were culled to 4M/4F. Within each litter, one male/female pair was treated postnatally on P11-14 with 0, 1, or 3 mg/kg chlorpyrifos, creating permutations totaling 9 pre/post-natal treatment group combinations. On P70, they tested males for 20 min. in an open-field under red light, video recorded the animals, and later used the Noldus system to categorize behavior. On P75-80, males were tested for 20 min. in a novel cage with a stranger male. On P90, females were given 3 pups for 20 min. to observed induced maternal behavior. On P120, mice were tested for 5 min. in the EPM. Data were analyzed by litter using ANOVA models, but the statistical section is difficult to follow because the run-on sentence says "prenatal treatment as block with respect to postnatal treatment, sex, and repeated measure as within-litter treatment factors, postnatal treatment, and sex as fixed-effect factors within litter, and repeated measure as fixed factor within subjects." From this, it is somewhat difficult to determine whether sex was treated as a between or within factor, but the analysis does indicate that litter was handled as a blocking factor within the ANOVA and it appears that prenatal treatment was a between factor and postnatal treatment a within factor. Variables where the same subject was tested repeatedly were handled as repeated measure factors. Significant F-tests were followed up using Tukey with Bonferroni correction. Results: In the open-field, they found increased activity in the chlorpyrifos6 males (females were not tested). In the test with a stranger mouse, the principal finding was in the postnatal chlorpyrifos3 group that showed increased attack behaviors against the stranger; they also found that the prenatal chlorpyrifos6 group showed a significant increase "upright postures" during this test. In the test of induced maternal behavior, the postnatal CFP 1 & 3 groups showed decreased licking frequency but increased licking duration, along with increased crouching frequency over the pups and for longer intervals but decreased pup sniffing. In the EPM, the significant finding was in the postnatal chlorpyrifos3 group that showed increased time in open arms. Strengths: They controlled for litter effects, and then did factorial ANOVAs; they controlled for multiple comparisons; they used adequate numbers of litters; and they used sound behavioral methods. The complex prenatal x postnatal treatment design could be strength to the extent that it identifies critical periods of exposure but could be a weakness as it makes the experiment logistically complex and difficult to manage; the study included multiple doses, 2 of which were prenatal and 2 of which were postnatal. Weaknesses: The choice of the narrow exposure windows is not well justified and appears arbitrary. Changing the dose levels used for the prenatal and postnatal exposures adds another complication. The fact that significant outcomes appear only in some dose groups treated prenatally and in some treated only postnatally, but not in those treated both pre- and postnatally which cumulatively had greater chlorpyrifos exposure, is difficult to reconcile. The effects were

not generally dose-dependent even within the prenatal, postnatal, or prenatal-postnatal combination groups. More fundamentally, the interpretation of the anxiety and social interaction tests as indices of neurotoxicity is uncertain. While these tests are interesting and identify areas for further investigation, they are difficult to use in risk assessment until they have an established neurotoxicological basis. For example, is it more adverse to show a modest increase in anxiety or a decrease? It is known that in humans anxiety/stress is an inverted U-shaped function. Low stress and anxiety leads to poor performance whereas high stress and anxiety interferes with performance; moderate stress and anxiety produce optimal performance. This is true in all mammalian species. Chronic stress and anxiety follow the pattern; many studies have shown that moderate developmental stress in rodents leads to increased cortical thickness and greater arborization and improves learning. Anxiety tests, such as the EPM, have a well-validated basis in the context of antidepressants. They are valid when used as intended to assess the effects of acute or subchronic exposure to SSRI, tricyclic and atypical antidepressants. They are increasingly used in gene targeting studies where candidate genes suspected of involvement in fear, anxiety, and stress are being assessed, but they have never been validated in neurotoxicology. This is also the case with social interaction tests. These tests are still being developed and their meaning in basic neuroscience research, as for example, in genetic models of autism spectrum disorders (ASD), are not yet established. Trying to interpret such methods in neurotoxicology is premature. Again, one must pose the question: Is it worse or better that the females given pups in this experiment, lick them more or less if they were exposed pre- and/or post-natally to chlorpyrifos? Unfortunately, no one can answer this question based on currently available data.

This experiment (Venerosi *et al.*, 2006) is identical in design to that of (Ricceri *et al.*, 2006). The Ns are the same as are all the major experimental design features. In terms of test outcomes, they conducted a social recognition test in which females were placed in single cages for three days and then introduced to stranger females for three min.; 45 min. later, they were re-exposed to stranger mouse-1 for another 3 min. and 45 min. later exposure to stranger mouse-2 and ultrasonic calls were recorded. Vocalizations in controls went down on retest-same and up on retest-different, whereas vocalization in the prenatal CFP3 group changed slightly, and in the CFP6 group it changed dramatically, causing retest-same to go up and retest-different to go up more than in controls. Postnatal CFP largely reversed the pattern. Social investigation: prenatal CFP increased social investigation, had no effect on retest-same or retest-different (latter not shown in figures). Here again, neither social interaction induced vocalizations nor social investigation of other animals has a known neurotoxicology interpretation. Strengths: This experiment has the same strengths as Ricceri *et al*. (2006). Weaknesses: It has the same weaknesses as Ricceri *et al.* (2006).

**Postnatal Studies**

This study (Dam *et al.*, 2000) comes from the Slotkin lab and has most of the features of this lab's previous work. As before, Sprague-Dawley Zivic-Miller rats were used. As before, litters were culled after birth to 10 pups and randomized across dams on P1 and

every 3 days thereafter until weaning. chlorpyrifos was administered s.c. in DMSO on P1-4 at doses of 0 or 1, and other animals were treated with chlorpyrifos on P11-14 with 0 or 5 mg/kg/day. Prior to weaning, offspring were tested for surface righting on P3-4 and on the inclined plane (20 degree angle) on P5-8 for those treated on P1-4. Open-field activity was manually tested in a large 100x100 cm field on P21 and P30 for 5 min. each time. Results: They found delayed surface righting and inclined plane rotation times in the chlorpyrifos1 females treated on P1-4, but not in males. In the open-field, they report that males in the CFP1 group had decreased square crossings and rearing frequency with no change in grooming frequency. In the P11-14 groups, they report no change in P21 or P30 in open-field line crossings, but chlorpyrifos males showed increased rearing at P30 and no other changes. Strengths: They used adequate sample sizes and analyzed the data taking litter, albeit artificial litter, into account. Weaknesses: They sampled 2 offspring per sex per litter, so there was slight over-sampling. The artificial litter technique remains unverified as a technique, and it could introduce stress on the pups and the dams being shuffled every three days. The 5-minute open-field test is generally regarded as inadequate by current standards, including those from 2000 when this study was published. The Zivic-Miller rat is less than ideal for behavioral studies. The vehicle used was DMSO. The findings were not dose-dependent.

This study (Levin *et al.*, 2001) was in collaboration with the Slotkin lab and has many of the common experimental design features noted above from this group. They used Sprague-Dawley rats from Zivic-Miller. Litters were culled to 10 pups, randomized and re-randomized every several days. Offspring were treated with CFP on P1-4 with 1 mg/kg, or on P11-14 with 5 mg/kg s.c. as in the previous study by Dam *et al.* (2000) dissolved in DMSO. However, this experiment shows the influence of the Levin lab: Adult offspring were tested for spontaneous alternation in the elevated T-maze referred to previously, including significant concerns about using a non-standard way of conducting this test. Locomotor activity was tested in the figure-eight system and RAM testing was as reviewed above with pharmacological challenges given after initial learning. All testing was done during the dark cycle. Spontaneous Alternation was conducted on P28-42, 5 trials with confinement after an arm choice for five sessions; figure-eight testing was done on P28-42, 1 h per session, 3 sessions spaced one week apart. The RAM was the 16-arm system tested on P57-91 with testing conducted three 3 days per week for a total of 18 sessions (4 arms unbaited; 12 arms baited; 10 min per session). RAM with drug challenge was conducted on P98-119 with scopolamine at doses of 0.04, 0.08, 0.16 or mecamylamine at doses of 1.25, 2.5, 5.0 mg/kg. ANOVA models were treatment and sex as between factors, and interval or day as within factors at $P<0.05$ except interactions which were taken as significant at $P<0.10$; the method of doing pairwise comparisons was not indicated in this paper. Results: They reported no effect on spontaneous alternation frequency and a small effect on latency in chlorpyrifos-exposed males on this test. In the figure-eight test, no effects were seen in the chlorpyrifos P1-4 exposed group, but reduced habituation slope was noted in the P11-14 CFP group. In the RAM test, the P1-4 CFP group showed effects on working and reference memory errors in males in the first block of trials but not thereafter and in females across blocks; no effect of chlorpyrifos exposure was seen after P11-14 exposure on learning the task. Treatment with 0.16 mg/kg of scopolamine increased reference memory errors in P11-14

chlorpyrifos-exposed males, with a larger effect in females that occurred at the lower scopolamine doses but not at the highest dose (0.16 mg/kg) of. Strengths: They used adequate sample sizes, litter was accounted for in the analyses, there was no litter over-sampling, and the behavioral methods were mostly sound. Weaknesses: Not a dose-response study; the vehicle for chlorpyrifos was DMSO; the use of pup randomization; issues concerning interpretation of the RAM data given the absence of an ITI delay interval; and the use of the elevated T-maze for spontaneous alternation.

This study (Aldridge *et al.*, 2005) also comes from the Slotkin lab. This time they used Sprague-Dawley CD rats from Charles River. They obtained timed pregnant rats and culled litters shortly after birth to 10, randomized pups among dams and re-randomized them every several days. Pups were treated with 1 mg/kg chlorpyrifos on P1-4 by s.c. injection in DMSO. No more than 1M/1F per artificial litter were sampled for a total 9M/9F per treatment (hence 36 rats were used altogether). Rats were tested during the dark cycle starting. Tested consisted of the EPM on P52-53, a two-bottle sweetness preference test on P54, and starting on P64 RAM learning with ketanserin challenges on weeks 16-17 with the drug given 20 min prior to testing at doses of 0, 0.5, 1.0 or 1.5 mg/kg (5HT2 antagonist). RAM was tested for 18 sessions (as above). Results: In the EPM, they report that CFP-exposed males had increased time in open arms (indicative on reduced anxiety) with no effect in females. In the sweetness preference test, CFP exposure reduced preference for the sweet choice in both sexes from about 4:1 to about 3:1. In the RAM, chlorpyrifos caused treatment x sex interactions on both working and reference memory. For working memory, chlorpyrifos exposure increased working memory errors in males and decreased these errors in females; similarly reference memory errors were increased in chlorpyrifos-exposed males and were decreased in females. While several of these changes from the interaction were individually short of being statistically significant, when errors types were combined, the male and female changes in errors were significant. Ketanserin had no effect on errors of either type in controls, but increased both error types in chlorpyrifos exposed rats at all doses for working memory and at the high and low doses for reference memory. Strengths: Adequate sample sizes, analyses that took litter into account, well-conducted behavioral methods. The use of ketanserin to show an effect of a 5-HT2 antagonist tested at multiple dose levels of the challenge drug was a major strength of this study and is a finding worthy of future investigation. Weaknesses: They used only one dose of chlorpyrifos, they used DMSO as the vehicle, and they did not use a delayed ITI during RAM testing making it difficult to determine if working memory or habit learning was actually what was affected.

In this study (Ricceri *et al.*, 2003) CD-1 mice (bred in-house, conception = E0), were culled to 5/5 M/F per litter and 30 litters were used. chlorpyrifos was administered s.c. in DMSO at doses of 0, 1 or 3 mg/kg on P1-4 or P11-14. On P1, 5 and 11 ultrasonic vocalizations were recorded; on P10 homing behavior to home cage scent was tested; on P25 locomotor activity was tested; on P35 mice were tested in a box divided into white and black compartments; social interaction with stranger mice was tested at P45; and at P60, passive avoidance (males only) were tested for up to 10 trials to remain in light side for 2 min. with a 24 h retention test. They sampled only 1 mouse per sex per litter and

analyses took litter into account. Results: No effect of chlorpyrifos on ultrasonic vocalizations or pup homing to home cage scent were obtained, or on locomotor activity but the authors noted a p<.06 trend in the P11-14 chlorpyrifos3 mg/kg group to be more active. In the white/black box, effects of P1-4 chlorpyrifos exposure were significant on 1 out of 5 test intervals (interval-2), whereas in the P11-4 chlorpyrifos exposed animals activity changes were significant on 1 out of 5 test intervals (internval-4). Several borderline effects and several significant interactions on different social interaction measures were found in the CFP groups but a clear pattern was not evident. There were no significant CFP-related effects on passive avoidance acquisition or retention. Strengths: Sample sizes were adequate and litter was taken into account. Behavioral methods were appropriately conducted. Weaknesses: Dissolved chlorpyrifos in DMSO, effects were not dose-dependent nor exposure period-dependent. Most of the effects were small.

In this study (Venerosi *et al.*, 2008) CD-1 mice (bred in-house, conception = E0) were treated S.C. with CFP at 0 or 3 mg/kg dissolved in peanut oil on P11-14. Litters were culled 4M/4F. Mice were evaluated in a social interaction test at P40-45. Female offspring bred and after delivery tested for nest building on P1-7 and other maternal behavior on P1 and later tested in the light/dark test of anxiety. They also did put retrieval test and a test for maternal aggression. These authors did a power calculation and sample size determination; they controlled for litter effects and used mixed model ANOVAs for most data analyses, but for some data they used non-parametric methods. They had 15 litters using a split-litter design. Results: they found no effects on ultrasonic vocalizations or social investigation. They found females exposed to chlorpyrifos did not build nests as well or defend their territory as much against a stranger male mouse, and took less time to emerge from the dark side of the light/dark box, but no other measure on this test was affected. Strengths: This was one of the most rigorous experiments in terms of sample size, control for litter effects and statistical methods for analyzing the data. Weakness: The relevance of the tests as indices of neurotoxicity are entirely speculative as none have been validated in this context, or in any context as strong evidence of developmental abnormality no matter what the independent variable.

The study by (Johnson *et al.*, 2009) used Sprague-Dawley CD IGS rats (Charles River) bred in-house. Twenty litters were used in a split litter design. Litter size was adjusted to 12-14 balanced for sex, with 7 groups per litter to the extent possible. Exposure was on day P1-5, 6-13, 14-20 as follows: Controls received oil from P1-20, and the low dose group received chlorpyrifos at 1.0 mg/kg from P1-2, but the mid and high doses groups received escalating doses: the mid dose received chlorpyrifos 1.0, 2.0, and 4.0 mg/kg during each of the aforementioned exposure ages, and the high dose received doses of 1.5, 3.0, and 6.0 mg/kg, respectively. The remaining three groups were exposed to methyl parathion at doses of 0.2 mg/kg throughout, or escalating doses of 0.2, 0.4, and 0.6 mg/kg/day (mid dose) or 0.3, 0.6, or 0.9 mg/kg/day (high dose) in oil by gavage. Results: The authors' report no effects of physical landmarks of development (pinna unfolding, fur appearance, day of eye opening, or day that incisors erupted) and no effects on early reflexes (surface righting, air righting, startle emergence, cliff avoidance, or inclined plane). Ad adults, they tested rats in a 12-arm RAM (8 baited, 4 unbaited

arms). They found no significant working memory effects in females but a significant increase in working errors in the high dose chlorpyrifos males across sessions and at lower doses in final week only. For reference memory, the female mid and high dose chlorpyrifos groups made fewer errors, whereas for males in the mid and high dose groups made significantly more errors. Strengths: Used a split-litter design and controlled for litter effects, had adequate sample sizes, included multiple doses of chlorpyrifos and tested two OPs (methyl parathion), had a strong statistical approach. Weaknesses: Did not include an ITI delay in the RAM test. Overall, this was one of the stronger studies.

In this study (Carr *et al.*, 2001) Sprague-Dawley CD rats (Charles River) were used and bred in-house. Rats were assigned to four groups with whole litters assigned to each group with a total of 5 litters per group with two offspring tested per sex per litter, *i.e.*, final numbers were 10 per sex per treatment group. Rats were gavaged on P1-21 every other day with corn oil, or corn oil containing a lower dose CFP 3 mg/kg, a mid-dose P1-5 of 3 mg/kg, P7-21 of 6 mg/kg, or a higher dose of P1-5 of 3 mg/kg, P7-13 of 6 mg/kg, and P15-21 of 12 mg/kg. The offspring were tested in an open-field on P10 and P12 for 3 min. each time and on P14, 16, 18, 20, 25, and 30 for 6 min each time). Statistically, they used a general linear model ANOVA and they set a significance level at $p<0.01$. The pairwise method used was not mentioned. They found reduced locomotion at P25 and P30 at the mid and high dose levels in both males and females. Strengths: They accounted for litter effects in the design and used appropriate statistical methods. They included three doses levels of chlorpyrifos plus control. Weaknesses: The sample size of five litters per group made the study under-powered and they slightly over-sampled per litter by using two per sex per litter.

This study (Jett *et al.*, 2001) used Long-Evans (Charles River) rats and obtained timed-pregnant animals. They culled litters to 10 and randomized pups among dams (no mention of sex balancing). Offspring were treated with chlorpyrifos on P7, 11, 15 at doses of 0.3 or 7 mg/kg given S.C. in oil to entire litters. A major concern is this sentence: "2 or more litters were used for random selection of pups used in behavioral studies". This suggests that severe over-sampling from a few as two litters were used. Another group was treated with chlorpyrifos (same doses) on P22 & 26). They tested offspring in the MWM on P24-28 hence in postweaning treatment group one dose was given in the middle of the testing regimen. The MWM was 90 cm in diameter and the goal platform was 25 $cm^2$; hence search ratio was 245:1 (which is within the range typically used for mice). They tested for 5 days, 2 trials on day-1, then 1 trial per day on days 2-5 with a probe trial 30 min after last training trial. They gave cued trials on day-5 the method used was not described. Statistically, they used ANOVAs but the details are not provided. They report a main effect on MWM latency for both CFP exposed groups but the method of pairwise comparisons is not given, and they report a high dose effect on the probe trial for time spent in the target quadrant. There is no mention of path length; no differences were reported on cued performance or on swim speed. Their final Ns were (M/F): Control 10/10, low dose 10/9, high dose 9/8. For the postweaning chlorpyrifos treatment, they report a treatment main effect which they report sorting by day by an unspecified statistical method. Looking at the figure, most of the effect

appears to be on days 3-5; also both CFP groups spent less time in target quadrant on the probe trial. They report no speed differences and state that no cued differences were found but the data are not shown. For the postnatal experiment the Ns are (M/F): Control 4/3; low dose 4/3; high dose 4/4. Strengths: They did most of the procedures in the MWM that should be included, such as assessing swim speed and cued performance. Weaknesses: There appears to be no control for litter effects and the number of litter used was as low as two, indicating a severely under-powered and potentially fatally flawed design. This is unfortunate because this is one of the few studies to test a lower dose of chlorpyrifos (0.3 mg/kg). Also, giving the probe trial for the MWM shortly after the last training trial provides somewhat ambiguous information. Changes may be attributed to either working or reference memory since the interval was too short to rule out working memory as a principal contribution.

**Pre- and Postnatal Studies**

This study (Maurissen *et al.*, 2000) used Sprague-Dawley CD rats (Charles River) bred in-house. They used 20 litters per treatment group (conception = E0). Treatment was by gavage on E6-P10. Doses given were 0, 0.3, 1.0, 5.0 mg/kg/day, but note that the postnatal exposure was to the dams, not the pups. On P4 litters were culled to 5/5 M/F. They used different subsets of pups per litter for different tests as follows: Set-1: Brain morphometry on P11; Set-2: Delayed spatial alternation on P22-24 and again on P61-90 (but using only 8/sex/group from 16 litters rather than 10/sex/group from all 20 litters); Set-3: Locomotor activity on P13, 17, 21, 60, and ASR on P22 and 61 (used 1 M/1 F from all 20 litters/group for these tests); Set-4: Developmental landmarks; body weight, and on P65-70 brained dissected and fixed for neuropathology. For the delayed spatial alternation test they used 3 delay intervals at each test age. For locomotion they used a 40 x 25 photocell system. For ASR they gave 50 trials with acoustic signals of 120 dB with ITI = 10 s. Data were analyzed by ANOVA with litter taken into account. P-values were considered significant at $P<0.02$. Where treatment main effect occurred, follow up was by a stepwise approach by first removing of high dose and re-analyzing the data, if still significant, then removal of the mid dose and reanalyze, *etc*. Interactions were followed up using simple-effect ANOVAs. Results: They report delayed vaginal patency at the high dose; delayed pinna detachment and prenuptial separation at $p<0.03$ and $P<.05$ neither of which reached their $P<.02$ cut-off. However these would be more commonly regarded as significant. They found no significant effects on the delayed spatial alternation test. They found no significant effects on locomotor activity. They found a trend on ASR latency at $P<.03$ but no effect on startle amplitude (the principal measure on this test). Strengths: This study has the most robust sample size in this entire group of 21 articles. They controlled for litter effects and used appropriate statistical methods. They delayed spatial alternation test was a particular strength and in their Fig. 7 they show the working memory decay as a function of the length of the delay interval, proving that they are measuring working memory. They included control plus 3 dose levels of chlorpyrifos and this is one of the only studies to test a low dose (0.3 mg/kg chlorpyrifos). They also avoided the use of DMSO as the vehicle. Weaknesses: The delayed spatial alternation test used much reduced sample sizes compared to the study as

a whole (8 per sex per group rather than 20). While this is of some concern, there were no trends in the results suggesting that a latent effect might have been missed. For the ASR, they performed a simple startle habituation test rather than a PPI procedure which is more informative. Note: It may be significant given the discussion at the meeting that they used an oral rather than subcutaneous route of exposure. Given the first pass metabolism of chlorpyrifos, this may have implications for the total amount of exposure the rats received.

This study (Braquenier *et al*., 2010) used CD-1 mice (Charles River) but the number of dams was not given nor was how conception was dated (a relevant factor in studies with prenatal exposure). Doses of chlorpyrifos were 0, 0.2, 1.0, 5 mg/kg/d given on E14-P14. For the postnatal CFP exposure the compound was given to dams (not pups). Litters were culled to 4M/4F with one female per litter used for testing. They tested offspring for locomotor activity (5 min/session) for 8 days (scored manually). They also did a 5 min light/dark test (mice started on the light side). EPM was assessed (standard 5-min procedure). Data were analyzed by ANOVA with follow-up by Dunnett's test. Results: They found no significant effects on locomotor activity. For the light/dark test, they found no effect on the percent of time spent in the dark compartment but a significant decrease in the percent of time in the light compartment in the chlorpyrifos1 group, but not in the chlorpyrifos0.2 or 5 mg/kg groups. Side transitions showed trend (P<.08) which they followed-up anyway and report a significant reduction in the CFP1 group by Dunnett. In the EPM they report a trend at p<.10, did Dunnett follow up tests anyway and found reduced time in open arms in the CFP1 group but not the other dose groups. They also found that the percentage of arm entries into open arms was significantly decreased in the CFP1 group. Total transitions were not significantly affected. Strengths: Tested three dose levels of chlorpyrifos, including a low dose (0.3 mg/kg), and gave the compound in oil rather than DMSO, and used standard methods. Weaknesses: No indication that litter effects were accounted for in the design or statistical analyses. Effects were found only at the 1 mg/kg dose and were not dose-dependent, they did follow-up tests on many trends that were not statistically significant and declared these follow-up effects to be significant findings.

**Summary**

Among these 21 reviewed articles (which include more than 21 experiments), many effects are reported at chlorpyrifos doses ranging from at 1.0 to 7.0 mg/kg (and in one study 10 mg/kg). However, these are dose levels known to significantly inhibit cholinesterase in RBC; therefore, based on these 21 studies, cholinesterase inhibition is an adequate threshold as no credible evidence of neurobehavioral effects below 1.0 mg/kg were found.

Three studies tested doses < 1.0 mg/kg chlorpyrifos. Two studies used 0.3 mg/kg and one used 0.2 mg/kg (Jett *et al.*, 2001; Braquenier *et al.*, 2010; Maurissen *et al.*, 2000, respectively). Of these, two found no effects at these lower doses *(Maurissen et al., 2000;Braquenier et al.*, 2010). Only one study found effects at 0.3 mg/kg (Jett *et al.*, 2001); however, this study contains serious methodological flaws which are of sufficient

magnitude to cast serious doubt on the credibility of the findings. While the data from Jett *et al.* (2001) raise the possibility of neurobehavioral effects at 0.3 mg/kg/d, their data require replication in a study that is properly designed, adequately powered, and appropriately analyzed. Until such time as new data at such lower doses become available, it is concluded that no dose <1.0 mg/kg in any neurodevelopmental behavioral study shows evidence of adverse effects (or of any effects, even including those outcome measures whose effect is indeterminate or unknown).

In addition, effects of chlorpyrifos at 1.0 mg/kg are difficult to interpret because of methodological limitations, inconsistencies, and variation in study design, sometimes lack of control for litter effects, oversampling issues, behavioral methods used, and lack of dose-response findings.

Above 1.0 mg/kg, the data show somewhat more consistency, but even here, dose-response experiments are the exception. At 5.0 mg/kg of chlorpyrifos, reduced body weight is sometimes seen, and at doses above 5.0 mg/kg, increased mortality may occur along with other evidence of toxicity. Given this, a significant gap in the literature of dose-response studies exists in the range downward toward 0.2 mg/kg and extending up to and including doses previously tested of 1.0-2.0 mg/kg that is needed in order to determine what, if any, dose-effect curve for neurobehavioral effects occurs in this range.

It appears that neurobehavioral outcomes are more sensitive to prenatal and prenatal-neonatal exposures than to neonatal exposure alone . This implies that prenatal exposure may be the exposure period contributing to this observation, but unfortunately, most of the pre- and neonatal studies are not entirely informative because the neonatal exposure was to the dam rather than directly to the progeny. This makes it unclear what the dose to the offspring actually was . More studies, especially dose-response studies, in the lower dose ranges with exposure from implantation to the end of major neurogenesis (approximately P20) are needed, again with doses below 1.0 mg/kg and with concomitant measurement of maternal, fetal, and neonatal cholinesterase activity.

Exposures in many of the existing studies are for only a narrow interval during gestation or the neonatal period. Prenatal exposures should be from E6-20 to 21 for rats, and E6-18 or 19 in mice in order to span most of early brain development (equivalent to human first and part of the second trimester). And for neonatal treatment, exposures should be from shortly after birth to approximately P20 (equivalent to latter half of the second and all of the third trimester equivalent brain development for humans). If the critical period or most sensitive period is within this range, then such comprehensive exposure should cover the entire span of CNS development that represents the species being modeled, *i.e*, human beings.

Exposures in many of the existing studies are for only a narrow interval during gestation or the neonatal period. Prenatal exposures should be from E6-20 to 21 for rats, and E6-18 or 19 in mice in order to span most of early brain development (equivalent to human first and part of the second trimester). And for neonatal treatment, exposures should be from shortly after birth to approximately P20 (equivalent to latter half of the second and

all of the third trimester equivalent brain development for humans). If the critical period or most sensitive period is within this range, then such comprehensive exposure should cover the entire span of CNS development that represents the species being modeled, *i.e*, human beings.

In the prenatal studies, the use of timed-pregnant females shipped from breeders is problematic for behavioral studies because maternal stress, even if regarded as equivalent across dams assigned to the treated and control groups, introduces a variable that has the potential to interact with the independent variable. Were maternal stress to interact with chlorpyrifos, it would confound the outcome and make a result difficult to interpret (which is exactly what is found in the 21 reviewed studies). Since no one has tested for this, it is currently impossible to rule it out.

Many studies use diurnal and some nocturnal testing. If additional dose-response studies are undertaken, this factor should be held constant so that results can be better compared.